

## Abstract

1. **背景和动机：**摘要的开头介绍了电子显微镜（EM）在高分辨率成像组织和细胞方面的应用，强调了手动分割大规模EM数据集的耗时和费力。自动分割方法的重要性得到了强调。
2. **重点关注：**摘要指出，这篇综述的主要焦点是过去六年中在大规模细胞EM中深度学习分割技术的进展。这期间取得了显著的进展，涵盖了语义分割和实例分割两个方面。
3. **数据集和贡献：**文中提供了关键数据集的详细介绍，这些数据集对深度学习在2D和3D EM分割中的发展起到了重要作用。
4. **学习方法：**综述包括了监督学习、无监督学习和自监督学习方法，并分析了这些算法如何适用于分割EM图像中的细胞和亚细胞结构。这表明了多样化的方法。
5. **挑战与网络架构：**讨论了分割EM图像所面临的特殊挑战，如异质性和空间复杂性，并描述了一些克服这些挑战的网络架构。
6. **评估指标：**提供了一些用于评估EM数据集在不同分割任务中的性能的评估指标的概述。
7. **未来展望：**最后，摘要展望了EM分割的当前趋势和未来前景，特别是与大规模模型和未标记图像一起学习EM数据集的通用特征。

## Introduction

1. **引言背景：**电子显微镜（EM）是一种利用电子束替代可见光进行成像的技术。它可以在纳米尺度下捕捉生物样品的图像，用于研究生物组织、细胞、亚细胞结构和分子复合物。EM产生的图像可以是二维（2D）图像或三维（3D）体积，允许科学家分析各种细胞器的超微结构和它们之间复杂的空间关系。最近，EM技术有了显著的进展，不仅可以获得2D图像，还可以生成3D体积数据。
2. **体积EM（vEM）的进展：**3D EM，也称为体积EM（vEM），使得可以以前所未有的细节对细胞结构进行三维分析。这是通过采集一系列2D EM切片，然后将它们堆叠在一起来创建3D数据集的方法实现的。自动3D EM技术的使用使得可以获得大容量的数据，有些研究甚至获得了以TB（1TB = 1,000 GB）为单位的数据。例如，一项研究使用串行切片TEM（ssTEM）成像了成年果蝇的完整脑部体积，覆盖了1 mm<sup>3</sup>或10,000个体素，需要100 TB的存储容量。此外，大规模的3D EM也可以通过冷冻电子层析（cryo-ET）成像，以在生物样品的原生环境中研究细胞结构和大分子组装。

3. **EM技术的进步**：最新的EM技术，如多束扫描电子显微镜（multibeam scanning EM），实现了高速成像，比以前快数百倍，使得可以在几小时内成像整个组织切片，而不是几天。这导致了数据量的大幅增加，因此需要自动化的EM分割方法来处理这些大规模数据集。
4. **语义分割与实例分割**：自动化EM分割需要两种关键技术：语义分割和实例分割。语义分割是一种像素级图像分析任务，它将图像分割成不同的连贯区域，并为每个像素分配一个表示其语义类别的标签，例如核、线粒体等。实例分割则为每个语义类别的像素分配唯一的实例标识，以区分相邻结构。这对于精确分析细胞内部结构和相互作用非常重要。
5. **传统方法与深度学习**：以前，传统的图像分析方法和浅层学习算法用于分割EM图像，这些方法包括像统计分析、特征值分析、分水岭算法等。然而，近年来，深度学习方法已成为主导方法，因为它们能够从原始数据中提取复杂的特征，无需先验领域知识，适应不同规模和类型的数据，提供高鲁棒性和可扩展性。

## Strategy of literature search

1. **文献搜索背景**：文献搜索的目的是回答两个主要问题：1) 哪些电子显微镜（EM）数据集可以用于分析，它们存在哪些挑战，以及它们在深度学习（DL）研究中的作用是什么？2) 如何利用全/半/无监督深度学习流程来处理EM图像的语义分割和实例分割？
2. **搜索策略**：为了回答上述问题，研究人员在PubMed、Web of Science和Google Scholar上执行了文献搜索，仅检索标题中的关键词（TI），并限制搜索时间范围为2017年至2022年。搜索查询如下：TI = ((electron microscopy OR EM) AND (segmentation OR semantic OR instance OR supervised OR unsupervised OR self-supervised OR semi-supervised))，以及在Google Scholar中包含（deep learning, segmentation, electron microscopy）的标题或摘要。从搜索结果中排除了不在研究范围内的文献，如材料科学中的深度学习和基于传统图像处理方法（DL之前的方法）。然后，使用前向和后向雪球技术来汇编最终的38篇论文列表。
3. **结果总结**：根据文献搜索的结果，研究人员整理了38篇论文，根据学习技术（全监督或非全监督）、分割类型（语义或实例）、应用（2D或3D）以及底层建模骨干等方面进行了总结。这个总结的图表（图3）提供了有关这些文献的概览信息。

在讨论这些文献之前，论文还提到了关键的EM数据集和DL架构的演变，这两个关键组成部分对于推动EM分割分析的进展起到了至关重要的作用。

## 主要EM数据集的收集

1. **序列切片TEM和SEM数据集**：序列切片透射电子显微镜（ssTEM）或扫描电子显微镜（ssSEM）用于研究神经组织中的突触连接和高分辨率膜结构。这些数据集具有高度各向异性，即在 $x$ -和 $y$ -方向具有高分辨率，而在 $z$ -方向具有较低分辨率，因为它依赖于序列切片的精度。
  - **Drosophila Larvae数据集**：这是ISBI 2012挑战赛的一部分，包含两个包含30个切片的卷，用于自动神经元分割。它是自动神经网络在ISBI 2012挑战中成功用作像素分类器的里程碑。
  - **CREMI3D数据集**：包含来自成年果蝇大脑体积的三个大而多样的子体积，用于神经元分割的训练和评估。这是MICCAI 2016会议的一部分，旨在理解大脑电路的功能。
  - **SNEMI3D数据集**：包含来自小鼠皮层的100个ssSEM图像的体积。这是ISBI 2013挑战的一部分，用于分割神经结构和构建组织的3D模型。
2. **FIB-SEM数据集**：离子束切片扫描电子显微镜（FIB-SEM）提供了各向同性体素大小的高分辨率数据，适用于神经元和各种亚细胞结构（如线粒体、囊泡和高尔基体等）的自动分割。
  - **Lucchi数据集**：这是从小鼠大脑海马部位获取的FIB-SEM体积，已成为评估线粒体分割性能的标准数据集。
  - **FIB-25数据集**：包含了果蝇大脑七个列中突触回路的重建，用于研究果蝇大脑的神经电路。
  - **OpenOrganelle Atlas**：包含整个细胞和组织的3D体积，用于研究局部细胞相互作用和其复杂排列。
3. **SBF-SEM数据集**：扫描束扫描电子显微镜（SBF-SEM）用于成像神经回路和细胞，适用于各向异性切片。它还包括了与大体积EM相关的数据集。
  - **Perez数据集**：包含小鼠大脑下丘脑的图像堆栈，用于线粒体、溶酶体、细胞核和核仁的自动分割。
  - **SegEM数据集**：来自小鼠体感皮层，提供了树突、轴突和突触的分割标签。
  - **CDeep3M数据集**：包含了用于脑部线粒体、突触囊泡和线粒体等的分割标签的数据集。

4. **Cryo-ET数据集**：电子层析（ET）用于获取EM切片的3D结构，而冷冻电子层析（cryo-ET）在低温下成像生物样本。这些数据集用于分割细胞器。
- **Cellular Cryo-ET数据集**：用于细胞器如线粒体、囊泡、微管和质膜的分割和定性分析。
  - **CDeep3M-C数据集**：用于小鼠大脑中的囊泡和膜的分割。

这些数据集在EM分割研究中发挥了关键作用，用于训练和评估深度学习模型，以实现自动化的细胞和亚细胞结构分割。这些数据集的特点和挑战有助于推动EM分割技术的发展。

## 用于 EM 语义和实例分割的骨干深度学习网络的背景

### 2D 语义分割：

早期的2D EM图像语义分割采用了卷积神经网络（CNNs）的方法，首次应用于神经元膜的自动分割。这些方法通过使用滑动窗口策略，对局部区域或补丁中的每个像素进行标签预测，从而识别不同的细胞结构和膜。

然而，这种方法存在两个主要问题：

1. **冗余处理**：由于滑动窗口策略的使用，相邻补丁之间存在冗余处理，导致计算资源的浪费。
2. **上下文和定位准确性的权衡**：这种方法在处理较大的上下文时需要更深的网络，但在深度网络中存在下采样，从而导致了较深网络的定位能力下降，而较小的上下文则限制了模型的上下文理解能力。

随后，引入了Fully Convolutional Networks（FCN）和U-Net等网络架构，这些网络采用了跳跃连接（skip connections）的结构，以在上采样过程中融合深层次的语义信息和浅层次的外观信息，从而改善了语义分割的性能。这些网络可以对任何尺寸的输入图像进行像素级别的密集预测。

## 3D 语义分割：

3D EM图像的语义分割涉及将3D体积划分为不同语义区域，为体积中的每个体素分配相应的标签。由于许多EM数据集由沿z轴堆叠的2D截面组成，因此进行体积分割对于准确捕捉3D结构和连接性非常重要。这一领域面临的主要挑战包括空间复杂性、序列截面的错位或缺失以及体积各向异性。

有三种主要方法用于解决3D语义分割问题：

1. **基于2D分割的方法**：首先，对堆叠的2D图像进行2D分割，然后使用后处理技术进行3D重建。这些方法可以使用基本的分水岭算法或复杂的图割算法等。
2. **基于3D CNN的方法**：这些方法使用3D卷积神经网络（CNNs）来学习包含3D空间上下文的数据表示。例如，3D U-Net和V-Net是两种扩展的架构，它们以不同方式处理3D数据。
3. **基于图形分析的方法**：这些方法通常使用图形理论算法，如图割或最小生成树，来根据颜色、强度、边缘强度等属性将图形分成不同区域或集群，然后识别不同区域之间的边界。

## 实例分割：

实例分割涉及将图像或体积中的每个像素或体素分类为特定类别，并为每个对象分配唯一标识。实例分割方法可分为基于提案的方法和基于无提案的方法：

**基于提案的方法**（proposal-based）：这些方法使用两阶段检测网络，如RCNN、FastRCNN、FasterRCNN等。它们首先使用深度神经网络进行特征提取和目标提案生成，然后进行边界框回归和分类以获得实例分割结果。Mask-RCNN是其中的一个例子，它在FasterRCNN基础上添加了一个分支来预测每个对象实例的二进制掩码。

**基于无提案的方法**（proposal-free）：这些方法旨在将语义分割和实例分割结合在底层方法中，而无需显式的目标提案。这些方法通常包括轮廓/边缘地图的预测以及语义掩模，然后将这些信息融合以获得实例分割结果。

这些方法的发展使得在电子显微镜图像分析中能够更准确地识别和分割细胞结构和细胞器，并为研究人员提供了有力的工具来理解生物学样本的结构和组织。



## 完全监督方法

完全监督方法使用注释图像（训练数据）来学习计算模型，这些模型可以从相似分布（测试数据）中分割出未见图像中的结构。算法利用训练集来确定模型参数，从而最大限度地提高模型的泛化能力。

这些方法主要依赖于有标注的图像（训练数据）来学习计算模型，以便在未见过的图像上进行分割。以下是一些主要内容的详细信息：

### 5.1. End-to-end 学习 — 语义分割

- **End-to-end 学习：**这是一种机器学习方法，其中一个单一模型被训练来执行任务，而不需要预定义的中间步骤或特征。模型直接将输入数据映射到所需的输出，通常通过反向传播进行训练，以最大化模型的泛化能力。
- **2D CNN架构：**一些方法采用了2D卷积神经网络（CNN）架构，如FCN（全卷积网络）和U-Net，用于从EM图像中进行语义分割。这些网络能够捕获大的感受野，以及在解码过程中的低级信息，从而实现准确的分割。
- **3D CNN架构：**对于3D EM图像的语义分割，一些方法使用了完整的3D CNN架构，这些网络能够处理3D体积中的信息。这种方法通常需要更多的计算资源。
- **RDN和其他方法：**一些方法引入了创新的技术，如Residual Deconvolutional Networks（RDN），以改善分割的性能。RDN使用了残差连接和低分辨率特征映射融合等技术。
- **数据增强：**一些方法使用数据增强策略，如随机变换图像，以提高模型的鲁棒性。

### 5.2. End-to-end 学习 — 实例分割

- **实例分割：**这些方法旨在将图像中的每个像素或区域分配给特定的实例，例如细胞器。这对于将相邻的细胞器进行区分非常重要。
- **2D CNN架构：**一些实例分割方法使用2D CNN架构，具有形状先验（shape prior）等特性，以改善分割的质量。这些方法通常需要训练不同的模型来处理不同的细胞结构。
- **3D CNN架构：**还有一些方法采用完整的3D CNN架构，以实现3D细胞器的实例分割。这通常需要更多的计算资源。

这些方法的应用领域包括神经元膜的分割和线粒体分割，以及在电子显微镜图像中进行的实例分割。这些深度学习方法已经在细胞学研究中发挥了重要作用，可以帮助科学家更好地理解细胞结构和组织。

### 5.3. 集成学习：

- 集成学习是一种方法，将多个算法或模型的输出结合在一起，以提高预测性能，包括准确性和泛化性能。
- 常用的聚合方法包括像素或体素的平均值、多数或中位数投票。
- 在大脑细胞膜分割中，Zeng等人使用了集成技术，结合了多个不同版本的DeepEM3D网络，以提高性能。
- CDeep3M是DeepEM3D的云实现，用于分割细胞器。
- Guay等人提出了一种集成范式，用于血小板细胞中的细胞质、线粒体和颗粒的分割。
- Khadangi等人使用集成方法聚合了不同类型网络的输出，以提高线粒体分割性能。

### 5.4. 迁移学习：

- 迁移学习是一种框架，用于将从一个数据集中学到的知识应用到另一个数据集中，通常通过微调预训练模型来实现。
- Mekuč等人使用迁移学习来分割线粒体和内溶酶体，演示了如何在应用程序的标记训练示例有限的情况下使用迁移学习。
- 迁移学习的挑战包括过拟合，因此研究也集中在少样本学习和域自适应领域。
- Dietlmeier等人提出了一种基于超列的Few-shot学习方法，用于线粒体分割。
- Bermúdez-Chacón等人使用域自适应来实现在有限标记目标数据上的高性能。

### 5.5. 可配置性和可重复性：

- 自配置网络是一种能够根据输入数据和任务自动调整结构和参数的人工神经网络。
- nnU-Net框架是一个例子，可以根据交叉验证自动确定一些超参数。
- Franco-Barranco等人的研究揭示了不同网络在EM数据中线粒体分割中的实质性可重复性问题，帮助确定稳定和轻量级的模型。

综上所述，这一部分讨论了集成学习、迁移学习以及模型的可配置性和可重复性在电子显微镜图像分割中的应用和方法。这些技术和策略有助于提高分割性能和适应不同的数据和任务。

## 半监督、无监督和自监督方法

这一部分涵盖了半监督学习、无监督学习和自监督学习方法，以及它们在电子显微镜图像分割中的应用。以下是对这一部分的总结：

### 半监督学习、无监督学习和自监督学习：

- 半监督学习和无监督学习是两种机器学习方法，它们的主要区别在于用于训练模型的标记数据量不同。
- 无监督学习是一种处理未标记训练数据中的模式和关系的机器学习类型。在这种情况下，算法通过对相似的数据点进行聚类或分组来学习识别数据中的模式和关系。
- 半监督学习是监督学习和无监督学习的结合。它使用标记和未标记数据来训练模型。标记数据用于训练模型执行特定任务，而未标记数据用于帮助算法学习数据中的模式和关系。
- 自监督学习中，模型在从数据本身生成的标签数据集上进行训练。其目标是学习可用于下游任务（如分割）的数据的有用表示。

### 半监督学习策略：

- 常见的半监督学习策略之一是使用标签传播和自训练。该过程从对标记样本进行训练开始，然后对未标记样本进行分类。根据主动选择策略或学习的分类器，选择一些样本并将其添加到训练集中，然后多次重复此过程。这可以在归纳或转导方式下执行。
- 归纳方式是指对看不见的目标进行模型训练，以向先前训练的模型添加新信息，使其能够在新的未见数据上进行泛化。而转导方式是指根据选择的一部分标记和未标记数据对模型进行训练，以便在有限的已见目标上进行正确预测。

### 相关研究：

- Takaya等人提出了一种名为4S的半监督方法，用于神经膜的分割。他们的方法利用卷积神经网络，并通过迭代方式使用模型生成的分割图来重新训练模型。这种方法的目标是使模型只能泛化到接下来的几个切片，而不是整个体积。



- Wolny等人提出了一种半监督方法，用于线粒体的分割。他们的方法在一些图像中手动注释了少量结构，然后通过引入嵌入一致性项来训练模型，以便模型可以泛化到整个数据集。
- Bermúdez-Chacón等人探索了无监督学习，研究了线粒体分割的无监督领域自适应策略，将一个模型从一个大脑结构（源：小鼠纹状体）适应到另一个大脑结构（目标：小鼠海马）。
- Peng等人研究了领域自适应，使用对抗性学习进行FIB-SEM到ssSEM数据集的分割。对抗性学习有助于在不同组织和物种的各种数据之间进行领域自适应。
- Conrad和Narayan使用对比学习，特别是时刻对比，从未标记的CEM500K数据集中学习有用的特征表示，然后进行迁移学习。

总的来说，这一部分介绍了半监督学习、无监督学习和自监督学习的方法，并提供了几个用于分割的具体案例研究，这些方法在电子显微镜图像分割中具有广泛的应用。这些方法有助于克服标记数据不足的问题，并提高了分割的性能和效率。

## 分割评估指标

分割方法的评估是通过测量地面真实（GT）和预测（PR）分割图之间的重叠程度来进行的。以下是用于评估分割性能的常见度量标准：

### 语义分割：

- 对于语义分割，所有的GT连接组件都被视为一个对象，同样，所有的PR连接组件也被视为一个对象。这将问题简化为二进制分类。
- 典型的性能指标包括准确率（Accuracy）、精确度（Precision）、召回率（Recall）及其谐波平均，也称为F分数（F-score，或当Precision和Recall给予相同权重时称为F1分数），或者Dice相似系数（Dice similarity coefficient, DSC）、像素错误（Pixel Error, PE）、Jaccard指数（Jaccard Index, JI，也称为交并比Intersection over Union, IoU）以及符合系数（Conformity coefficient）。

### 性能度量定义如下：

- 准确率（Accuracy）：所有正确分类的像素数与总像素数之比。

- 精确度 (Precision) : 所有真正例 (TP) 与所有正预测的数量之比。
- 召回率 (Recall) : 所有真正例 (TP) 与GT中的所有正像素数量之比。
- F1分数 (F1或DSC) : 精确度和召回率的谐波均值, 用于综合考虑准确性和召回率。
- 像素错误 (Pixel Error, PE) : 1减去最大的F1分数, 用于度量最大F1分数的错误版本。
- Jaccard指数 (Jaccard Index, JI) : GT与PR之间的交集与并集之比, 也称为交并比 (IoU) 。
- 符合系数 (Conformity coefficient, CF) : 度量被错误分类的像素数与真正例像素数之比, 从1中减去以得到。
- 兰德指数 (Rand Index, RI) : 度量两个数据分组之间的相似性, 适用于评估没有分割标签的情况。RE (Rand Error) 是RI的错误版本, 用于度量GT和PR分区之间的一致性。

### 分割的其他度量标准:

- 战争错误 (Warping Error, WE) : 用于评估PR和GT之间的拓扑不一致性, 包括几何标签错误。它度量了两个分割之间的拓扑变化上限。
- 分割和合并错误度量: 对于评估分割的分割和合并错误, 可以使用*VRand*、*VInfo*以及信息变化度 (Variation of Information, VOI) 等指标。
- 平均精度 (Average Precision, AP) : 用于目标检测问题, 通过改变检测阈值系统地计算精度-召回曲线下的面积来度量性能。
- 平均精度 (mAP) : 用于多类别问题, 将所有类别的AP取平均得到。

### 实例分割度量标准:

- 聚合Jaccard指数 (Aggregated Jaccard Index, AJI) : 用于评估实例分割的综合性能。
- 全景质量 (Panoptic Quality, PQ) : 用于综合评估分割和检测性能, 度量了分割的质量和物体的准确性。

这些度量标准用于评估不同类型的图像分割方法的性能。表格中呈现了基于这些度量标准的研究论文的结果, 展示了当前的最先进性能。

## 讨论和挑战

当然，请见以下更详细的总结：

**卷积神经网络（CNNs）在EM图像分割中的应用：** CNNs是当前EM图像分割领域的主要工具，特别是全卷积网络（FCNs）和U-Net等架构，它们具有强大的特征提取能力和上下文感知能力。这些网络使用深层架构以及多尺度卷积操作，生成高质量的2D分割结果，无需额外的后处理。此外，一些新的技术如扩张卷积和空间金字塔池化也被引入，以捕获多尺度的上下文信息，提高分割效果。

**2D和3D分割的对比：** 文章对比了2D和3D CNNs在不同EM数据上的应用。3D CNNs对于体积数据的分割效果更好，尤其适用于各向同性的数据。然而，3D模型面临着计算复杂度和内存需求大的挑战。为了解决这些问题，一些方法减小了输入块的大小，但这可能导致上下文信息不足，影响分割精度。此外，3D模型需要更大规模的标记数据集来训练。

**数据标注和学习方法：** EM图像的精细标注通常需要大量的人工努力，但这在实际应用中往往难以实现。因此，研究人员正在积极探索弱监督、半监督和自监督学习方法，以减少对手工标记数据的依赖。这些方法使得模型可以在少量标记样本的情况下进行分割，同时提高了泛化性能。

**实验设计和性能评估：** 目前，EM图像分割研究中存在一些挑战。首先，多类别分割方法的研究相对较少，大多数方法专注于特定细胞器的二元分割。其次，不同方法使用不同的性能评估指标，导致很难进行公平的跨方法比较。建立标准化的性能评估协议对于解决这些问题非常重要，以便更准确地比较不同算法的性能。此外，文章还提到，生物学家通常更关心形态学特征如直径、面积等，因此未来的数据集可能需要包含这些信息。

**未来研究方向：** 文章指出，未来的研究方向包括改进3D CNNs的计算效率，探索更多的上下文信息整合技巧，以及改进性能评估的标准化协议。此外，自监督学习方法和使用自注意力机制（transformer）的潜力也被强调，它们有望为EM图像分割带来更多的创新和改进。

综上所述，这篇文章提供了深入的见解，深入探讨了EM图像分割领域的关键问题和未来方向。它强调了CNNs在该领域的重要性，并鼓励研究人员在数据标注、学习方法和性能评估方面进行更多的创新，以推动EM图像分割的进一步发展。

## 结论

在这份综述中，我们描述了卷积神经网络在大规模细胞电子显微镜（EM）图像分割中的作用。基于标记数据的先进CNN架构的端到端学习已经在语义分割任务中实现了人类水平的准确性，但在实例分割问题上仍需要努力，特别是在高度拥挤的结构情况下。尽管取得了显著的进展，但在3D分割方面仍然存在一些挑战，主要是由于增加的计算复杂性和导致重建困难的体素误分类问题。

随着我们迈向更大规模的EM数据集，获得一致的标注变得更加具有挑战性。考虑到EM领域中完全标记数据的不足将持续存在，因此半监督和自监督学习的使用将变得更加常见。以前，重点主要放在设计网络架构上，以解决与单个结构相关的特定任务，比如神经区域中的突触或线粒体的实例分割。然而，在未来几年，我们预计将会出现向更多通用分割模型的转变，利用我们综述中讨论的大规模网络和学习方法，以一种与任务无关的方式提取通用特征。这些特征可以允许无监督发现新的结构和感兴趣区域，或者可以适应特定的监督分割任务。

## 额外信息

### 8.1. 概述

卷积神经网络（CNNs）已经成为处理电子显微镜（EM）数据中的自动特征提取和分割的首选方法，其中基于全卷积网络（FCNs）的骨干网络，包括流行的U-Net架构，尤为突出。这些网络使用更深层的架构，同时整合了不同级别的图像上下文，生成有效的2D预测，随后可以简单地集成用于3D重建，无需显式的后处理过程。此外，研究人员还探讨了使用膨胀卷积或孔卷积，特别是在初始层中，以扩展卷积的感受野。最近的工作还采用了空间金字塔池化，以捕获多尺度的上下文信息，实现在更高级别获取全局信息。值得注意的是，像DeepEM3D这样的架构通过仅在初始层中使用3D操作并在中心切片上预测像素概率，在各向异性EM数据集中实现了高精度。

### 8.2. 2D和3D分割

3D CNN已经成为体积EM数据集中EM分割工作流程中的重要组成部分，通过使用体素表示相比2D CNNs提供了总体上更好的精度。对于各向同性的体素，3D CNNs的有效性尤为明显，能够精确分割多样的细胞器。混合2D-3D网络现在是解决串行切片EM中各向异性的标准方法。nnU-Net是一种自配置的EM分割方法，既适用于2D又适用于3D，为网络深度或基于数据集特征和一组经验实验进行超参数调整提供了一个良好的起点。然而，由于数据集特征和

各种细胞器的独特挑战，仍然存在一些限制。

3D网络的主要限制是增加的计算复杂性，因为操作数量随输入大小的增加呈立方增长，并且对高计算内存的需求增加。为了解决这些问题，在训练阶段已经研究了较小的块大小作为输入，导致上下文减小，从而导致不精确的预测。此外，3D网络中的大量参数需要更大规模的数据集进行有效的模型训练。与3D相比，对2D数据的训练和推理需要更少的计算资源，在计算资源有限的情况下具有优势。3D模型的另一个挑战是处理通过切片失去形状连续性的细胞器，正如Franco-Barranco等人（2022）所示。这促使了在3D模型中研究形状先验，类似于在2D模型中所做的，其中形状先验被用于正则化。

无论是2D还是3D方法，关于相邻细胞器的实例分割都存在共同的挑战，特别是当细胞器包含与它们的膜相似的结构时。心脏线粒体内的晶体就是这样一个例子。未来可能会采取的一个方向是研究如何将环绕抑制滤波有效地嵌入到CNN中，以抑制对晶体的响应，从而更好地勾画出膜。这种方法的潜力在Aswath等人（2023）的最新研究中得以展示，他们将环绕抑制滤波引入作为线粒体实例分割的后处理技术。环绕抑制已经在低级图像处理中用于抑制轮廓检测任务中的虚假笔画，显示出很高的效果。

总的来说，未来的研究应该旨在克服3D CNN目前的局限性，探索创新技术来包含更多的上下文信息，并推进形状先验的理解和应用。这些进步将有助于更健壮和准确的EM分割，从而在EM领域带来更多的洞见和发现。

### 8.3. 数据标注和学习方法

大多数EM分割的进展都是通过完全监督的方法实现的，这些方法强烈依赖于精细标注数据的可用性和质量。然而，在现代大规模EM生物数据集中，很少提供精细的标注。相反，大多数数据集是使用预训练网络和校对自动生成的粗略掩膜。这种大规模EM数据集需要使用弱监督、半监督和自监督技术来进行端到端的改进EM分割（Papandreou等人，2015; Kirillov等人，2023）。

除了手动标注，EM图像还可以使用专门的成像模态进行标记，以针对样本中的特定结构。例如，CLEM（关联光电子显微镜）用于标记以荧光探针在（亚）细胞尺度上瞄准的结构（de Boer等人，2015; Heinrich等人，2021）。其他EM模

态包括在纳米尺度上进行的分析成像，用于提供各种细胞器的元素引导识别（Pirozzi等人，2018）。这些方法可以减少人工标注的偏见，但可能需要更长的样品制备和采集时间、专门的设备以及额外的后处理以生成分割结果。



由于标签的缺乏，为每个成像实验开发新的训练过程是不切实际的。当面临未见样本时，监督方法的性能通常受到负面影响，并需要昂贵的重新设计和重新训练工作。已经采用各种方法来使用少量标记样本来适应模型，并提高CNN的泛化能力。迁移学习和自监督技术都已用于分割未见的EM数据集。这些方法使EM图像能够以最少或无注释的方式进行分割。虽然总体而言，完全监督的方法是最有效的方法，但少样本和域自适应策略的初步结果看起来很有前景，并且在分割某些细胞器时表现出可比较的性能。

自监督学习的兴起，归功于它能够从无标签数据中学习通用表示，也在EM分割领域具有巨大潜力。事实上，CEM500K的可用性对通过自监督学习获得通用表示发挥了重要作用，CEM500K是一个包含来自不同细胞类型、组织和制备方法的无标签数据的综合性仓库。通过对CEM500K上进行预训练网络的微调，结果模型展示出能够处理广泛的细胞器的能力，性能可与专门的监督网络相媲美。

自监督学习潜力的一个例子是最近的Segment Anything Model (SAM) 的突破性 (Kirillov等人, 2023)，该模型在不同成像领域的各种分割应用中表现良好。该模型具有基于变压器的架构，可以在新的图像分布和任务之间进行零样本迁移。这种方法有潜力允许对许多不同的数据集进行更多的研究，并允许并入可以用作突出显示感兴趣区域的提示的附加输入。展望未来，特别值得注意的是，在自分割中使用变压器的潜力，尤其是用于管理大规模EM数据。

#### 8.4. 实验设计和性能评估

多类别分割只有一篇论文进行了研究 (Heinrich等人, 2018)，而其余的方法都侧重于特定细胞器的二进制分割。在这个限制下，CNNs的潜力没有充分发挥出来。多类别分割的使用利用全局和局部特征以提高总体精度。全局特征有助于通过捕获类间差异和这些结构出现的空间上下文来学习上下文信息。在多种类别中训练模型有助于提高泛化能力，使其更具适应性和可转移性到新数据集。

至于性能评估指标，目前在量化EM分割的有效性方面缺乏标准化的评估协议。虽然提出了各种性能度量标准，但方法的性能未一致地与相同的度量标准进行比较。这在线粒体的情况下尤其如此。这使得很难对现有研究得出强有力的结论。因此，建立标准化的评估协议对于实现不同算法之间的公平和有意义的比较至关重要。

此外，评估方法时考虑到生物学家的具体需求和要求也是重要的。虽然计算机科学家可能会优先考虑精确的轮廓描绘，但生物学家可能更强调测量形态特性，如直径、面积或长短轴比。因此，未来的数据集可能不仅包括分割掩膜，还包括生物学家在分割后进行的形态分析的真值。此外，生物学家对组织中的常见和罕见的超结构变化都感兴趣。从生物学角度看，分割的目标不仅是识别所有细胞器，还包括某些离群值或未建模的现象，如裂变事件、细胞



破坏或疾病进展。当前的分割度量标准无法捕捉到分割结果中特定类型的错误或这些错误的空间分布。错误的空间分布提供了关于分割区域中的不准确性或不一致性的位置和程度的见解，这有助于表征组织。生物学家参与评估过程，报告与正常结构外观不符的特定错误或异常事件，对于评估和量化这些事件是必要的。