

浙江大学

本科周报告

姓 名:	赖梓林
学 院:	计算机科学与技术
系:	计算机科学与技术
专 业:	计算机科学与技术
学 号:	3170104684

目录

第一章 摘要	3
第二章 正文	4
1.U-Net.....	4
1.1 UNet 的简介	4
1.2 UNet 的结构	4
1.3 UNet 的优点	6
1.4 UNet 的策略	8
1.5 UNet 的优化	9
2.结构生物学相关	10
2.1 X 射线晶体学	11
2.2 断层扫描	11
2.3 冷冻电镜单颗粒分析	11
2.3.1 简介	11
2.3.2 条件	12
2.3.3 基本步骤	12
2.3.4 技术优点	13
2.3.5 技术局限	13
2.3.6 相关知识	14
2.3.6.1 傅里叶变换	14
2.3.6.2 功率谱密度	15
2.3.6.3 卷积相关	16
2.3.6.4 图像滤波	17
2.3.6.5 中央截面定理	17
2.3.6.6 雷登变换	18
2.3.6.7 缺失锥问题	18
2.3.6.8 朝向求解	18
3.参考资料	19

摘要

报告结合了第一周关于典型的全卷积网络结构：U-Net 的相关基础知识，讨论了更深入的相关知识内容。另外，大致讨论了结构生物学中冷冻电镜的几种常用的观测手段：简述了断层扫描以及 X 射线晶体学的相关内容，对冷冻电镜单颗粒分析进行了更深入的讨论。

二、正文：

1. U-Net：

1.1. U-Net 的简介：

(1.1 部分未修改，和第一周报告内容一致。以下内容和第一周重复部分文字以灰色标注。)

U-Net 结构是由德国 Freiburg 大学计算机部门开发的一种全连接神经网络结构。开发时主要着重于解决生物医学领域图像分割问题。但近期逐渐被广泛用于各种图像分割相关领域。

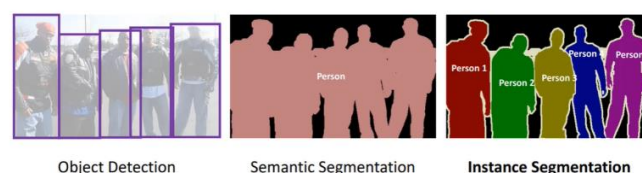


图 不同的图像处理任务

U-Net 比较特别的点是包含了一条收缩路径(contracting path)和一条扩展路径(expanding path)，且通过数据增强，U-Net 可以更有效率地使用标注样本。在较小的训练数据集的训练下，U-Net 仍有良好的表现，并且耗时极低。

U-Net 在比赛中的表现非常好，准确率远超历届最高水准和同届的第二名，并且对于不同的分割应用都有良好表现。

1.2. UNet 的结构：

UNet 首先属于全卷积神经网络，不包含全连接层。

UNet 具有一条收缩路径，和与之相对的扩展路径。整体的结构形似“U”，所以称为 UNet。在收缩路径上实施下采样，图像的分辨率不断下降，而在扩展路径上实施上采样，分辨率上升。

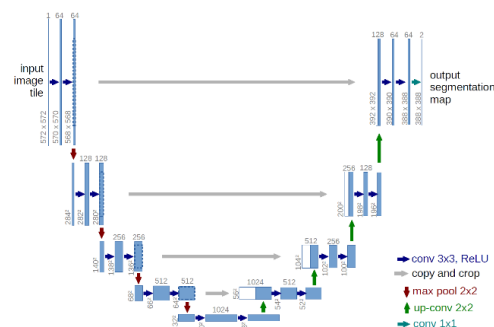


图 UNet 的具体结构

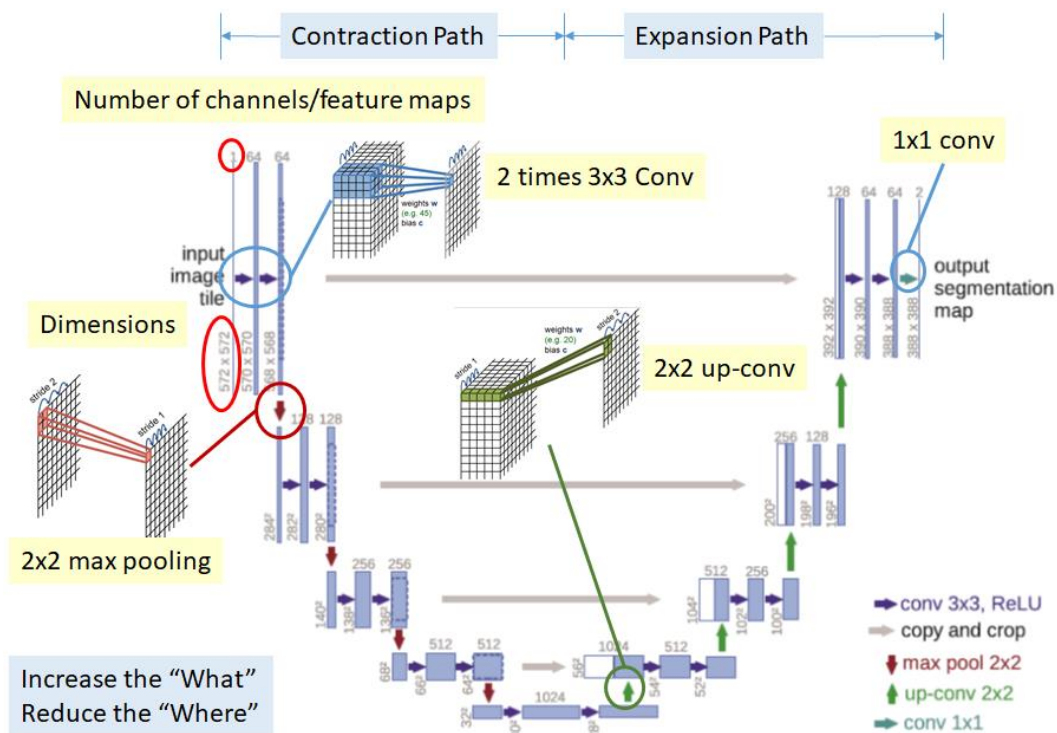


图 UNet 的具体结构解释

Unet 结构包括了卷积神经网络的核心组件卷积层和池化层，除此之外，添加了分割层，拼接层和上采样层（也属于卷积层）。

其中细节部分，池化层部分使用最大值池化层，在原论文中未采用填充策略，但在有些讨论中指出启用填充策略后能取得更好的表现。

从最左上角的结构开始，收缩路径最上层连续进行两次卷积核为 3×3 的卷积操作。由于未启用填充策略，每次卷积操作后，图像的长度和宽度都减少两个像素，从 512×512 大小的图像得到 64 张 508×508 的特征图。之后进行两个操作，一是将结果拷贝至扩展路径的最上层便于后面的拼接操作，二是对结果进行池化操作（卷积核为 2×2 的最大化池化）后，传入收缩路径的第二层。重复

这个操作直到收缩路径的最后一层。在收缩路径上，分辨率层层减少，特征图数量逐渐增大。

在收缩路径的最后一层，对所得结果进行上采样操作，扩大分辨率，上采样后的结果与收缩路径倒数第二层的结果拼接，作为这一层扩展路径的输入，经过两次卷积核为 3×3 的卷积操作后，结果传入扩展路径下一层，与之前收缩路径拷贝过来的特征图拼接。重复这个操作直到扩展路径最上层后，与收缩路径不同的是，扩展路径最后额外进行了一次卷积核为 1×1 的卷积操作，所得结果即为整个结构的输出。

收缩路径每一层：

原图像 / 上一层结果 \rightarrow 两次 3×3 卷积 \rightarrow 拷贝至扩展路径
 \rightarrow 2×2 池化 \rightarrow 传入下一层

扩展路径每一层：

上一层结果 + 收缩路径拷贝结果 \rightarrow 两次 3×3 卷积 \rightarrow 拷贝至扩展路径
 \rightarrow 2×2 上采样 \rightarrow 传入下一层

1.3. UNet 的优点：

因为并不存在一种对于所有问题都有良好表现的结构，需要根据问题决定特定的结构。在计算机视觉领域，卷积神经网络在图像识别、图像分类的问题中表现良好，而 UNet 则侧重于图像分割。

在普通的卷积神经网络中，数据不断通过卷积层、池化层，分辨率的下降使得偏底层的数据，如边界和边缘等逐渐模糊、丢失。最后只剩下特征信息。这样无法达到图像分割的目的。

随着卷积层的深入，我们获得特征信息，失去语义信息。而在这种情况下，因为忽视了局部和整体关系，只关注于局部的特征信息，会出现歧义。



图 歧义问题

在 U-Net 的收缩路径上，我们逐渐获得更多高级的特征并且降低了特征图的分辨率。

在扩展路径上，上采样操作和连续卷积操作恢复特征图的分辨率。这样的后果是，高级的特征保存了下来，但浅层的特征就会丢失，就是说对于图像整体来说，我们更明白特征“是什么”，但是会忘记特征“在哪里”。

对于上面的示例图来说，我们可能逐渐能辨别局部特征（黑色绒毛），但是对于特征在整体中的定位（这个局部和附近局部的关系等）信息就会慢慢丢失，会造成上面的歧义问题。

所以 U-Net 在扩展路径上，每一层的输入并不只是上一层上采样后的结果，而是上采样后的结果拼接上相同等级的收缩路径的特征图，这一步操作就提供了“在哪里”特征，即取得了特征和整体的把控。

最后一步的 1×1 卷积操作将 64 张特征图卷积为 2 张，即细胞特征图和膜特征图，这一步操作的结果即为最后的输出图像结果。

1.4. U-Net 的策略:

文章中提到 U-Net 的良好表现也得益于它充分使用了数据增强策略。

对于显微图像来说，主要需要的是平移和旋转不变性，以及对灰度变化和图像变形的鲁棒性。文章提到，在数据增强步骤中，对训练样本实施随机弹性变化，可能是以较少训练样本训练分割网络的关键。

他们在 3×3 的网格上施加平滑化变形，以 10 个像素为标准的高斯分布进行采样位移。之后通过双三次插值算法计算每个像素的位移，随后在收缩路径结尾的丢弃层(Dropout Layer)进一步执行隐式的数据增强。

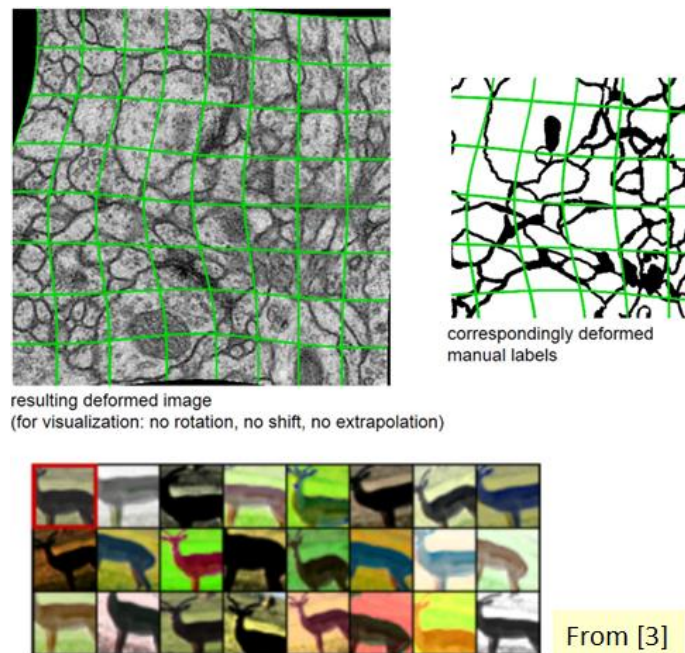


图 图像弹性形变

另外，正如前面小节提到的，U-Net 并没有采用零填充策略，每次卷积操作后，图像的分辨率会下降两个像素。因在结构中进行了多次卷积操作，最后的输出图像分辨率(388*388)远小于输入图像的分辨率(572*572)，网络上对于这一点的讨论很多，有些人在自己的项目中启用了填充策略得到了更好的效果，有些人认为不启用填充策略的原因是为了避免填充的 0 对结果造成影响。

U-Net 采用了重叠图块策略(Overlap Tile Strategy)，这种策略通常用于大图像的无缝分割。在这种策略下，对局部图像的标签预测需要考虑更大的一个局部的图像，对于在图像边缘的局部来说，使用镜像填充来代替空的部分。

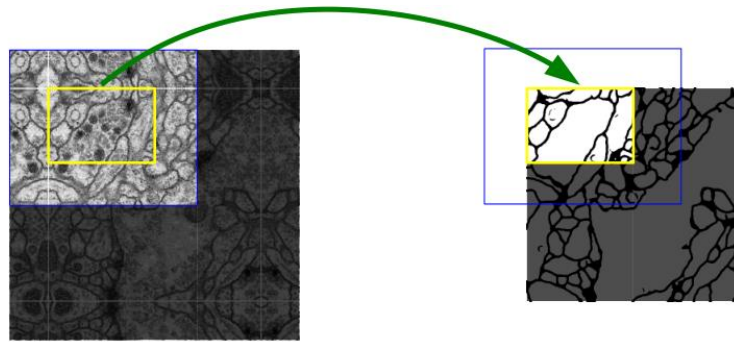


Fig. 2. Overlap-tile strategy for seamless segmentation of arbitrary large images (here segmentation of neuronal structures in EM stacks). Prediction of the segmentation in the yellow area, requires image data within the blue area as input. Missing input data is extrapolated by mirroring
<https://blog.csdn.net/mieleizhi0522>

图 重叠图块策略

另外，在样本中常会出现一种情况，两个物体紧紧贴在一起，这样的物体在网络训练中可能会被合并，所以为了使分割效果更好，U-Net 使用权重图 (Weight Map) 来协助判定两个相邻的物体。

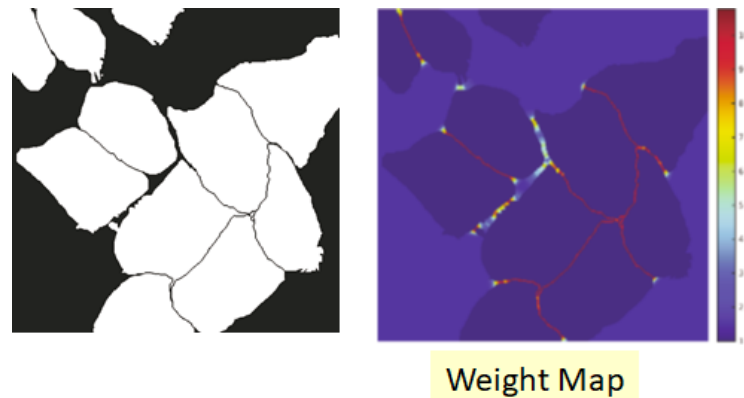


图 权重图

1.5. U-Net 的优化:

在优化问题上，我们首先要明确一点，优化要基于面向的问题来考虑，在网络上，有许多人针对不同的领域提出了对 U-Net 可能的优化。这一小节中，我们主要关注其中一个领域的优化方向，并从中获得新的思路。

在 towardsdatascience 网站上，研究者 Sik-Ho Tsang 提出了在牙科 X 射线图像分割问题中，对 U-Net 的几点改进。

首先，在原 U-Net 网络结构中，与输出层相连的是 1×1 卷积层，使得结果从 64 张特征图转化为 2 张，这样做的原因是，在原问题中，只需要分辨细胞特征和膜特征即可完成图像分割。

而在牙科 X 射线图像分割问题中，牙齿需要明确地被分割为不同的牙齿部分，而不是当作一个牙齿整体来看，研究者 Sik-Ho Tsang 指出，牙齿需要被分割为七个不同的部分以达到分割的目的。下图为七个部分的说明。



所以原 U-Net 最后输出的两张特征图就不足以完成对牙科图像的分割，因此，在最后一步中，需要修改配置，使 1×1 卷积的结果为 7 张特征图。

另外，在重叠图块策略中，原 U-Net 对图像边缘的局部的无像素部分采用镜像策略，而在牙科图像中，这位研究者采用零填充策略，他指出对于牙齿来说，镜像策略没有任何意义。

从他的优化方向上我们可以明白，大部分的优化都取决于你要处理的特定问题。确定要分割成大概多少部分，以此决定最后一层 1×1 卷积后取得的特征图数量。确定你的样本是否高度对称，镜像策略对样本来说不可行，才能决定重叠图块策略的细节部分。

除此之外，损失函数部分也有许多优化的可能，许多讨论者针对不同的问题提出了很多优化方向，在这里不深入讨论。

2. 结构生物学：

在生物学，很多最根本，本质的问题可以通过“看”来解决，由于大多细胞都很小，无法直观地观察或研究其结构等，就出现了一些不同的方法来“看”各种蛋白质、细胞甚至是复合物等。

接下来的小节中讨论几种在三维电镜下常见的观测方式。

2.1 X 射线晶体学：

经过 100 年的发展和成熟，X 射线晶体学已成为一种常规方法，这种方法需要蛋白质被打包在一个稳定的晶体中，由于在晶体中，每个个体高度有序地排列，用光源照射晶体时，个体地衍射信号互相叠加，增强了信号。但对一些无法简单晶体化的目标，例如较大的蛋白质或者蛋白质复合体来说，难以结晶化，无法使个体高度有序地排列，衍射信号是混乱的，此时就需要另一种方法来解决这个问题。

2.2 断层扫描：

2.3 冷冻电镜单颗粒分析：

2.3.1. 简介

冷冻电镜技术与 X 射线晶体学不同，它可以处理无法结晶化的情况。在实现上，单颗粒冷冻电镜并非通过晶体衍射确定结构，而是计算并组合多个独立大分子的图像来取得三维结构信息。它基本步骤是取得单颗粒、对齐、分类归类、做平均。一个典型的流程是：取得显微照片、估算衬度传递函数、挑选单颗粒、单颗粒排序、分析二维图像、三维重构、密度优化。

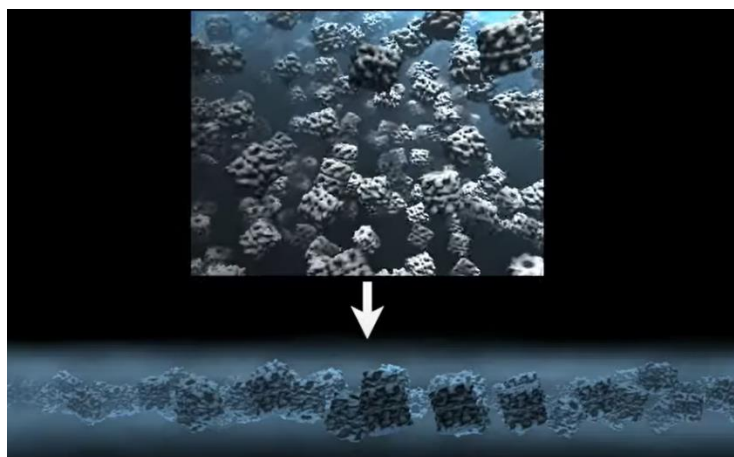


图 无序到有序

随着技术的发展，冷冻电镜技术也从早期的形态学观察工具发展成了三维结构体系观察工具，从早期的低分辨率发展到现在的原子级别分辨率。早期只能取得高分辨率的静态结构，而现在已经可能取得大分子复合物的动态结构

信息。并且在过去，冷冻电镜只能拍摄单张照片，而现在可以拍摄一段时间的影像，可以更好地处理、研究。

2.3.2. 条件

要完成单颗粒分析，首先需要有多目标（颗粒）的拷贝，且它们在溶剂中随机滚动。随后我们在电镜格栅上，将它们铺成非常薄的薄膜，然后冷冻。在冷冻的瞬间，这些颗粒将保持他们最后的朝向，所以我们就取得了大量的在三维空间不同朝向的颗粒。

随后，我们需要做的是，找到每个颗粒图像与其他颗粒图像对齐所需的参数和方向，并将它们平均。在平均后，我们在取得更高的信噪比，解释更高解析度(resolution)的细节的同时，也完成了三维重构。

2.3.3. 基本步骤

抛去细节部分，冷冻电镜三维重构大概包含以下步骤：首先，分子嵌入一层薄 amorphous ice 中，随后电子枪以光速射出电子，电子束途径样本，透镜(lens)。假如电子束经过样本中密度较大的一个点，则会分散开，此时经过孔(aperture)的筛选之后，分散较远的电子会被忽略。同时，因为有些电子发生非弹性散射，能量较少，会被能量过滤器(energy filter)过滤掉。最后由特殊设计的高科技相机截取图像。每个粒子的图像都是分子二维投影，之后由计算机在图像中选出相同方向的颗粒并根据方向分类。此时，每一个类中的颗粒都应有相同方向，将它们叠加，对每个类别重复采取相同操作。此时将每个方向的颗粒经过复杂的过程叠加起来，就形成了三维结构。



图 简要流程

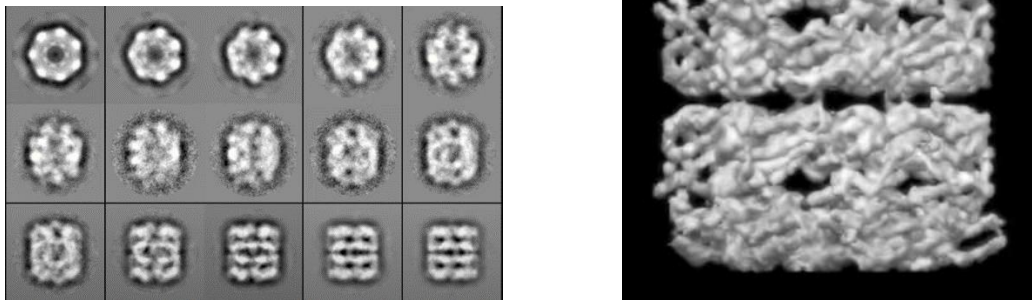


图 对多个方向的投影进行三维重构

2.3.4. 技术优点

冷冻电镜技术的好处是，它不需要限定样本为晶体，可以处理低对称对象，也可以处理局部不均匀的样本，适用范围更广。同时，在上面的一张图片中可以看到，因为颗粒在冰中一颗一颗排列，我们可以对比颗粒，排除异质或不纯的颗粒，挑出我们需要的目标颗粒。并且因为电镜照片同时包含了振幅(amplitude)和相位(phase)信息，通过冷冻电镜的结构测定不存在X射线晶体学中的相位问题。并且X射线难以集中，会产生样本损伤的问题。另外对样本量的需求也较小。并且，许多步骤可以自动化，节省了许多时间。随着单颗粒分析的发展，研究者提出可能最终可以在体内实施这项技术，因为颗粒周围的物质的影响是有可能在平均步骤中被消除的。

2.3.5. 技术局限

单颗粒分析技术也有它的局限性，如果想要取得目标颗粒清晰清楚的图像，可能会对那个颗粒造成过多伤害，甚至毁掉那个颗粒。

所以为了保证样本不受损伤，我们必须控制电子的数量。这样做的结果是最后只能取得信噪比较低的图像。并且在平均步骤也只能使用这些信噪比低的图像。

同时，平均步骤需要的一些不同的对齐方向和参数会对三维重构产生重大影响，如果参数出现误差，重构结果将变得模糊。

为了解决结果模糊的问题，我们可以在平均步骤中使用更多的颗粒来平均，在早期的单颗粒分析中，由于算力、技术等限制，平均步骤使用大约几百到一千的颗粒，而随着科技发展，可以使用的颗粒的数目上限不断上升。

在一些研究中，研究者使用百万到千万数量级的颗粒进行重构，随着颗粒数目的上升，重构结果的解析度也开始上升。但同时，他们发现随着颗粒数目的上升，解析度逐渐也达到一个上限，此时再增加颗粒数目对解析度的提升也越来越小。

而在最近期的研究中发现，由于直接检测器(direct detector)和运动校正(motion correction)的出现，颗粒图像的质量获得了提升，解析度好到足以建立原子模型级别的三维重构仅需要十万以内的颗粒图像就可以完成。

另外的局限性是由于生物学中的构象异质性产生的，这点不作讨论。

	Advantages	Disadvantages
Visible light	Not very damaging Easily focused Eye wonderful detector	Long wavelengths (~400 nm)
X-rays	Small wavelength (Angstroms) Good penetration	Hard to focus Damage sample
Electrons	Small wavelength (pm) Can be focused	Damage sample Poor penetration
Neutrons	Low sample damage Small wavelength (pm)	How to produce? How to focus?

图 对不同的放射线上的对比

另外，对于 X 射线晶体学来说，可以用 R-free value 来评估总体的质量，而对于冷冻电镜单颗粒分析来说，目前没有一种简单的质量评估标准。这也是单颗粒分析存在的一个问题。

2.3.6. 相关知识

2.3.6.1 傅里叶变换

对同一个物体来说，我们可以用不同方式，以不同角度去刻画理解。傅里叶变换就是实空间的物体在频率空间重的刻画。

因为任何复杂图形都可以看作由一系列正弦波组成，所以，傅里叶变换就是将复杂的波形分解成一系列简单正弦波。正弦波又可以用振幅和频率表示，

所以可以以振幅和频率作为坐标轴，表示这一系列的正弦波，这一系列操作的结果就是将实空间中的物体的各点密度转化为频率空间的信息。其中频率最高的也称为 Nyquist 频率。

在三维重构中，最终的目的是得到在空间上每一点的密度值。也就是取得由不同空间频率信息叠加的情况。

电镜成像的过程就是两次傅里叶变换过程，这也是冷冻电镜单颗粒分析使用傅里叶变换作为手段的关键原因。这也是

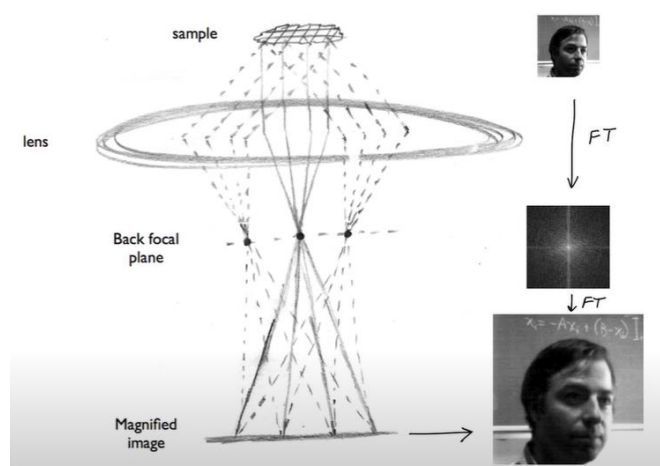
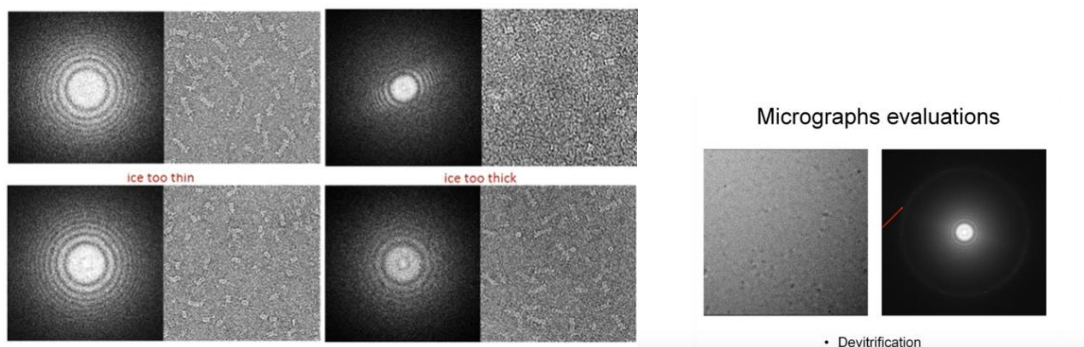


图 理想中的电镜

2.3.6.2 功率谱密度

根据傅里叶分析，任何物理信号都可以分解成一些离散频率或连续范围的频谱。对特定信号或特定种类信号（包括噪声）频率内容的分析的统计平均，称作其频谱。

在功率谱中没有相位信息，理想的功率谱呈复数同心圆环状。如果在某个方向丢失，则可能出现样品漂移。如果谱中圆环（Thon ring）是椭圆形状。则可能存在像散。如果没有圆环则是没有高解析度信息。如果出现黑白相间的斑块，则是样品有重结晶问题。



2.3.6.3 卷积

由于前面的章节中有讨论卷积神经网络，在这一小节就只提出几个关键点。实际上，卷积的过程就是图像的每一点变成一个卷积核，再叠加。

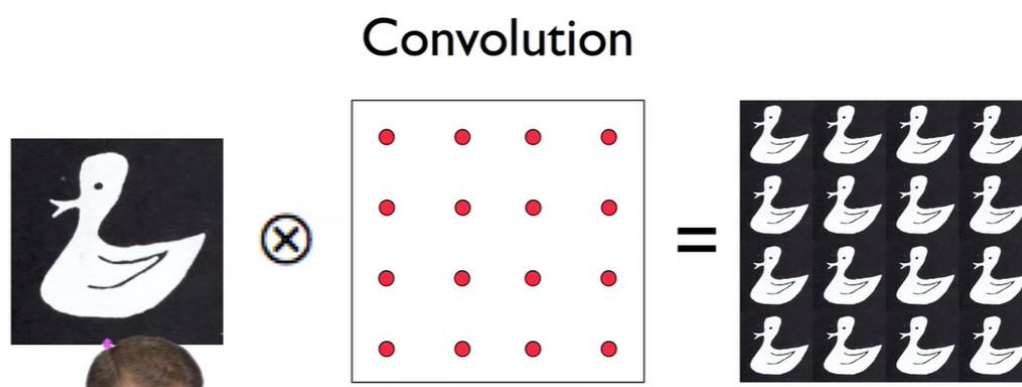


图 卷积过程

而在电镜成像过程中，理想中、完美的电镜应当涉及两次傅里叶变换，最后成像结果即为理想结果，但由于成像过程的傅里叶变换中，每个傅里叶变换组成部分在成像形成上的强度不同，有些更大，而有些完全丢失。所以卷积核即为衬度传递函数，它描述不同频率分量在像平面的强度。

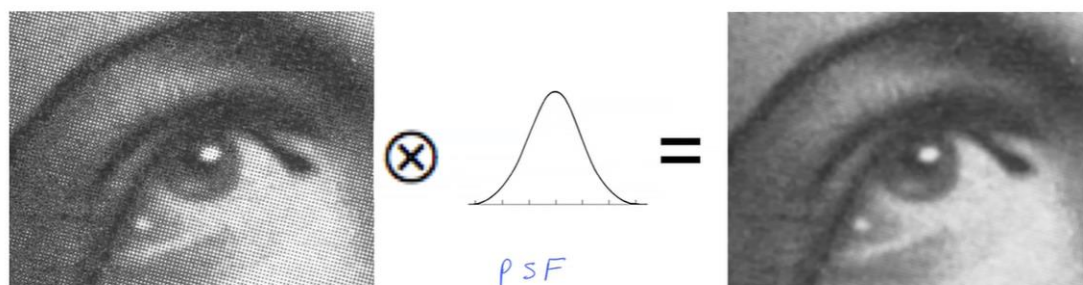


图 衬度传递函数作为卷积核

通过卷积理论可知，卷积可以根据傅里叶变换转化为乘积。

2.3.6.4 图像滤波

图像滤波是整个过程中重要的一环，可用的滤波器包括：低通滤波器、高通滤波器和带通滤波器。

其中低通滤波器保留低频信息，在实际处理中尤为重要。经过低通滤波器，图像对比度上升，取得信噪比更高的信息。

但重要的一点是，在图像处理过程中，高频信息和低频信息都是需要的，高频信息决定了最终图像的解析度上限，而低频提供的对比度提升粒子可视化的效果。这两种信息决定在平均阶段我们如何选择颗粒的集合，以及如何对齐这些图像，和确定对于我们所需要的解析度来说，所需要的颗粒图像数。

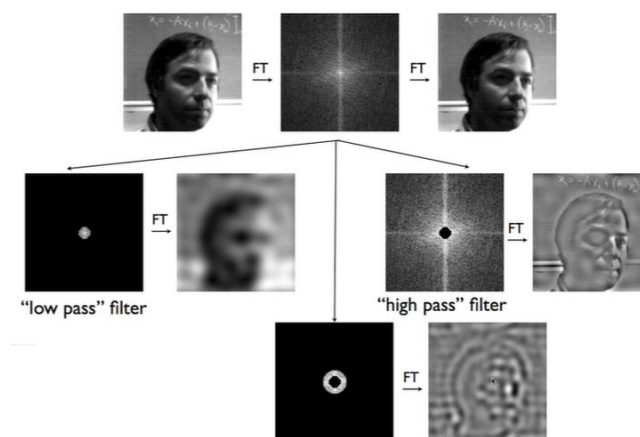


图 几种滤波器

2.3.6.5 中央截面定理：

中央截面定理是通过二维图像计算三维结构涉及的理论之一。因为对于一个物体来说，它的三维傅里叶变换和投影之间有确定的数学关系。它的投影的傅里叶等于它的三维傅里叶变换在垂直投影方向上一个截面。

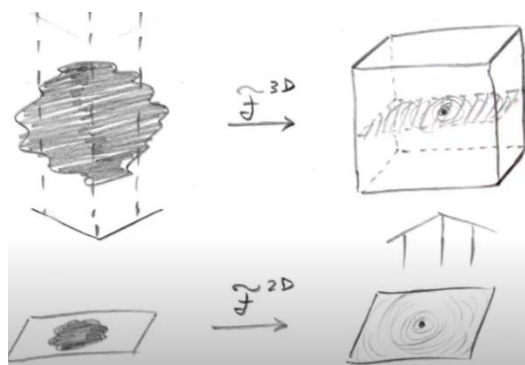


图 中央截面定理示意图

由这个定理我们能知道，如果取得了目标物体的各个角度的投影，我们就可以先取得对各个角度的投影的傅里叶变换，然后用这些投影的傅里叶变换填满三维空间，最后进行反向傅里叶变换来取得三维结构。

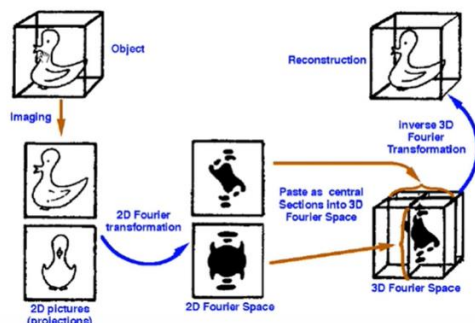


图 二维图像到三维结构

2.3.6.6 雷登变换(Radon Transformation):

在我们讨论的单颗粒分析问题中，雷登变换意味着如果各个方向投影方向是连续且满的，那么反向投影后，最后的三维图像密度和原密度是相同的。

雷登变换也是断层扫描的核心数学原理，在数学上和中央截面定理等价。

2.3.6.7 缺失锥问题(Missing wedge problem):

三维重构时中，由于采得投影不能将整个空间填满。在缺失方向会产生拉长效应。

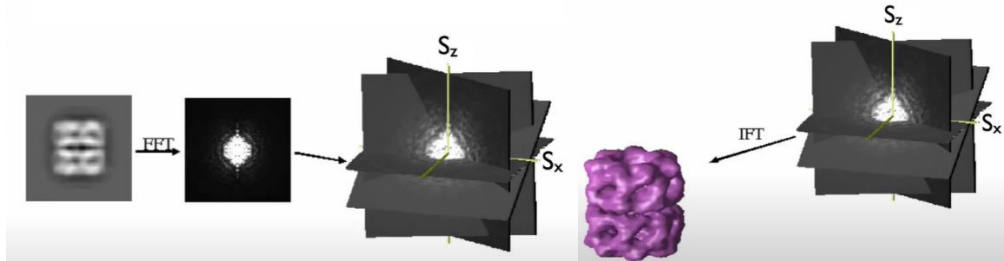
2.3.6.8 朝向求解:

在单颗粒分析中，各个颗粒的朝向确定是十分关键的，朝向的误差会导致最终重构结构模糊。我们通过等价线方法来解决朝向确定问题。

如果两个颗粒的投影都可以进行傅里叶变换的分析。那么根据中央截面定理，他们一定是在三维空间中通过原点的某一个截面，他们对应的傅里叶变换都是三维傅里叶变换的一个截面，而且这个截面通过原点。于是两平面会有共

同点，则可知一定有共同线（等价线）。任何两个投影的傅里叶变换结果间一定各自存在一条相同的线。

但若只有两个平面，还需要判断自由度，所以如果有三个以上平面间的等价线，朝向就可以三个截面间相对空间位置求解出来。



3. 参考资料：

1. U-net 源码分析（Keras 版本） <https://blog.csdn.net/jieshaoxiansen/article/details/82591864>
2. U-net 论文解析 <https://blog.csdn.net/mieleizhi0522/article/details/82025509>
3. An Overview On U-Net Architecture <https://medium.com/datadriveninvestor/an-overview-on-u-net-architecture-d6caabf7caa4>
4. An Introduction to different Types of Convolutions in Deep Learning <https://towardsdatascience.com/types-of-convolutions-in-deep-learning-717013397f4d>
5. Minaee, Shervin & Boykov, Yuri & Porikli, Fatih & Plaza, Antonio & Kehtarnavaz, Nasser & Terzopoulos, Demetri. (2020). Image Segmentation Using Deep Learning: A Survey.
6. Review: U-Net (Biomedical Image Segmentation) <https://towardsdatascience.com/review-u-net-biomedical-image-segmentation-d02bf06ca760>
7. Keras U-Net starter <https://www.kaggle.com/keegil/keras-u-net-starter-lb-0-277>
8. Visualization of U-Net <https://www.youtube.com/watch?v=uJ63xbVPMKs>
9. CSCI E-89 Deep Learning - Final Project - Image Segmentation <https://www.youtube.com/watch?v=cm6GmGkbhBo>
10. Image Segmentation using U-Net - Part1~6 <https://www.youtube.com/watch?v=azM57JuQpQI>
11. An in-depth look at image segmentation in modern deep learning architectures through the UNet <https://www.youtube.com/watch?v=NzY5IJodjek&t=1280s>
12. A 3 minute introduction to CryoEMh <https://www.youtube.com/watch?v=BJKkCOW-6Qk>

13. What is Cryo-Electron Microscopy (Cryo-EM)? <https://www.youtube.com/watch?v=Qq8D0-4BnIY>

14. Getting Started in Cryo-EM with Professor Grant Jensen

15 ~ 19. Part 2 ~ 6:- G. Jensen

https://www.youtube.com/watch?v=fEyLh9HqsWU&list=PL8_xPU5epJdctoHdQjpfHmd_z9WvGxK8-&index=18

以及学姐提供的相关资料和论文。