

用于学习高光谱显微镜仪器噪声的方法

摘要

拉曼光谱作为一种先进的分子指纹检测技术，凭借其卓越的光谱分辨率能力，能够同时捕捉生物或化学进程中多种物质的动态变化。该技术深入揭示了分子的振动与转动信息，为分子官能团及化学键的精确鉴别分析提供了有力工具。然而，拉曼光谱测量的质量常受限于其固有的低散射效率，导致所得光谱易受噪声严重干扰。针对这一挑战，本文探索了一种创新的噪声学习方法（Noise Learning, NL），该方法通过学习像素空间频域内的噪声，从而估算出每台仪器的固有噪声分布。随后，通过从原始噪声频谱中剔除这些估算出的噪声成分，实现了噪声的有效抑制。值得注意的是，NL方法在训练阶段采用了模拟生成的干净光谱与仪器噪声数据，这一策略规避了传统深度学习中对大规模标记数据集的依赖，实现了从样本依赖向仪器依赖的转变。实验结果显示，NL方法显著提升了数据的空间分辨率和光谱分辨率。无论是在块状硅、二维材料还是活细胞等复杂样本中，通过NL方法训练的UNet网络均能够大幅提升拉曼光谱的信噪比，为去噪领域的研究与应用开辟了新的道路。

关键词：去噪；高光谱图像；仪器噪声学习；Unet；拉曼光谱

1 引言

当用波长比试样粒径小得多的单色光照射气体、液体或透明试样时，大部分的光会按原来的方向透射，而小部分则按不同的角度散射开来，产生散射光。在垂直方向观察时，除了与原入射光有相同频率的瑞利散射外，还有一系列对称分布着若干条很弱的与入射光频率发生位移的拉曼谱线，这种现象称为拉曼效应。由于拉曼谱线的数目，位移的大小，谱线的长度直接与试样分子振动或转动能级有关。因此，与红外吸收光谱类似，对拉曼光谱的研究，也可以得到有关分子振动或转动的信息。目前拉曼光谱分析技术已广泛应用于物质的鉴定，分子结构的研究谱线特征。然而，拉曼光谱测量的质量常受限于其固有的低散射效率。具体地，对于大多数分子，每一百万个入射光子中的1个光子能够有拉曼散射 [8]。即使有等离子体增强，拉曼信号仍然容易受到噪声的影响，特别是在纳米级高光谱拉曼成像中，当对信号有贡献的分子数量很小并且随着成像面积减小到几个纳米，以及需要弱光照明或快速数据采集的活细胞成像时，光谱信号被噪声干扰的更加严重。因此，去噪成为拉曼光谱后续应用的一个关键前处理步骤。

2 相关工作

2.1 基于模型的去噪方法

高光谱图像可以被看作是灰度图像的堆叠，所以逐带去噪是该领域中最早出现的方法，它通过对每个光谱波段进行单独处理来去除噪声。代表性的方法包括字典学习法（TDL）、三维块匹配算法（BM3D）、非局部均值滤波（NLM）和加权核范数最小化（WNNM）等。这些方法在一定程度上能够去除噪声，但由于忽略了光谱之间的相关性，即频谱依赖性，它们往往不能充分利用HSI的特性。为了克服逐带去噪的局限性，研究者们提出了基于手工先验的模型驱动方法。这类方法通过构建数学模型来精细表示HSI数据的先验知识，从而更有效地去除噪声。常用的平滑图像模型是全变分（TV）模型，它能够在保留图像边缘细节的同时去除噪声。为了利用光谱的物理特性，研究者们还引入了空间光谱全变分（SSTV）来平滑光谱域的噪声。然而，全变分模型往往会产生阶梯效应和过度平滑，导致图像失去原有的纹理特征。为了优化去噪效果，研究者们对稀疏表示（SR）进行了深入的开发。例如，Peng等人提出了一种创新的方法，该方法在计算所有波段的梯度映射子空间稀疏度之前，将全变分正则化和低秩正则化相结合，从而实现了更精细的图像去噪。Zhang等人则进一步提出了基于群的稀疏表示方法，该方法通过考虑光谱波段之间的群结构，提高了稀疏表示的准确性和效率。同样，Lu等人也开展了自适应的稀疏表示频谱空间特征的研究，为HSI数据的去噪提供了新的思路。此外，低秩（LR）正则化也是HSI去噪方法中常用的技术。Zhang等人设计了将HSI数据表示为二维矩阵的低秩矩阵恢复（LRMR）框架，并利用高效的优化算法进行求解。由于HSI在空间域中通常呈现分段平滑的特性，而在频谱域中则表现为低秩子空间结构，这使得低秩正则化在HSI去噪中具有良好的适用性。随后，He等人进一步将低秩矩阵与全变分模型相结合，提出了全变分正则化低秩矩阵分解（LRTV）方法，实现了对HSI数据的更为精确的建模和去噪。然而，传统的矩阵技术往往导致HSI数据的空间细节丢失，为了克服这一问题，研究者们开始探索张量建模方法。例如，Wang等人提出了全变分正则化低秩张量分解（LRTDTV）方法，该方法采用Tucker分解来捕捉光谱波段之间的全局相关性，并结合各向异性SSTV正则化来保持空间和光谱域的分段平滑结构。另外，Chang等人提出了基于超拉普拉斯正则化的单向低秩张量恢复（LLRT）。He等人也提出了非局部与全局结合模型（NGmeet），这是一种高效的秩自适应交替最小化方法。Zhuang等人则提出了一种全局和非局部低秩分解（GLF）方法，该方法基于低秩张量展开的最优奇异值阈值，能够更有效地表示HSI数据中的低维子空间结构，这些方法都进一步提升了HSI数据去噪的性能。模型驱动方法建立在统计先验建模的基础上，具有良好的可解释性和泛化性。它们能够通过引入适当的正则化项来精细表示HSI数据的先验知识，从而实现对噪声的有效去除。然而，这类方法也存在一些局限性：首先，超参数的选择往往依赖于经验规则，这使得它们难以优化。其次，由于优化算法需要大量的迭代，大多数基于模型的方法在效率上并不是最优的。此外，手工设计的先验常常不能精确地捕获数据的先验结构，这在一定程度上限制了去噪效果的提升。

2.2 基于深度学习的去噪方法

近年来，深度学习在高光谱图像去噪领域取得了显著突破。基于模型的去噪方法往往依赖复杂的物理假设和手工设计的先验知识，而深度学习则能够利用其强大的图像表示能力，

自动生成噪声HSI与干净HSI之间的复杂映射关系。最早用于HSI去噪的深度学习方法为DnCNN，它利用二维卷积挖掘空间信息，逐带消除噪声。随后，Chang等人提出了一种HSI-DeNet，通过学习多通道二维滤波来捕捉光谱间的相关性。MemNet引入了一个存储模块，用于保留用于去噪的长期特征信息。Yuan等人在此基础上引入了基于滑动窗口策略的HSI残差网络结构，通过多尺度特征提取和多层次特征表示，提出了一种多尺度光谱空间去噪方。为了进一步利用空间-光谱相关性，Dong等人设计了一个三维U-net架构，将三维卷积滤波器分解为一维和二维滤波器，通过三维卷积提取空间光谱相关特征，减少了参数数量和计算复杂度。Xiong等人提出了一种端到端网络SMDS-Net来学习HSI的低秩和稀疏。Sidorov和Hardeberg基于深度图像先验，开发了一种高光谱图像恢复方法Deep HS Prior。此外，为了更全面地利用HSI中的信息，研究者们还探索了混合方法。例如，QRNN3D将三维卷积与准递归网络相结合，从全局角度整合不同波段的信息，捕捉光谱的全局相关性并消除因果依赖。GRN-net则通过两个推理模块仔细提取全局和局部空间光谱特征，实现混合噪声的去。同时，T3SC结合了稀疏编码原理和深度神经网络，提出了一种模型导向的频谱空间网络。TRQ3DNet在QRNN3D模型的基础上增加了Uformer块，以增强捕获远程空间依赖的能力进一步提高了去噪性能。与基于模型的方法相比，基于深度学习的去噪方法无需设计复杂的手工先验，而是直接关注端到端数据，利用网络对噪声和干净的HSI进行非线性映射。尽管深度学习在HSI去噪方面取得了显著进展，但其泛化能力、可解释性和性能稳定性仍有待进一步提高。更重要的是，深度学习学习的是噪声HSI与干净HSI之间的复杂映射关系，需要干净HIS作为GT去指导模型的训练，但实际情况中这个干净HIS是很难获取到的，所以此种方式的训练精确度也会受到影响[6]。本文复现了一种用于学习仪器噪声的方法，该方法利用深度学习去得到不同数据中对应的仪器噪声，然后从原始数据中减去模型预测的仪器噪声，就得到了干净的数据，是一种间接去噪的方法。

3 本文方法

3.1 模型总体框架

模型总体框架如图1所示。首先，原始拉曼光谱数据通过离散余弦变换（DCT）转换到频率域，得到DCT系数。接着，这些系数被输入到一个预训练的AUNet模型中，该模型通过学习从DCT系数中预测并去除噪声。去噪后的DCT系数经过逆离散余弦变换（IDCT）回到时域，生成预测的噪声。最后，通过从原始光谱中减去预测的噪声，得到一个去噪后的光谱，这个光谱的峰值更加清晰，有助于提高后续分析的准确性。整个流程结合了信号处理技术和深度学习的优势，有效提升了拉曼光谱数据的质量。

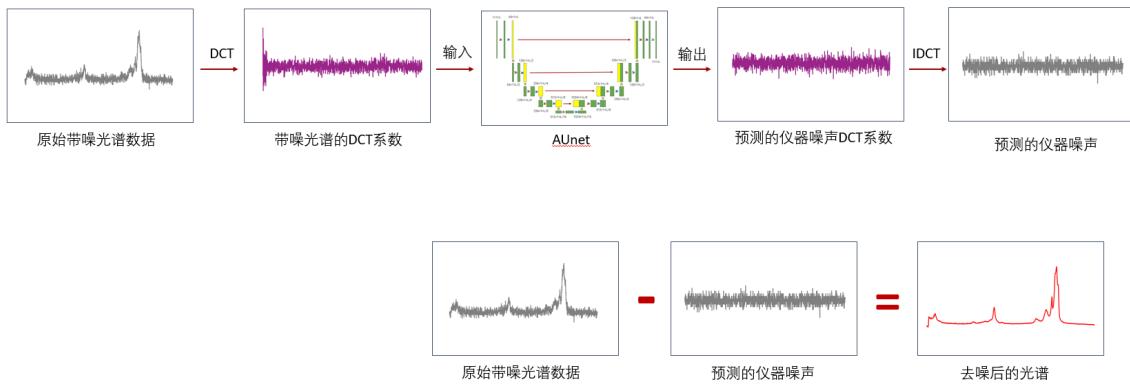


图 1. 模型总体框架图

在NL方法中，伪Voigt函数首先用于生成GT拉曼光谱，模拟低信噪比光谱数据由实验测量的仪器噪声和生成的GT数据相加生成。在NL中，将模拟的低信噪比光谱数据作为模型输入，实验测量的仪器噪声作为GT指导模型的训练，去训练深度学习模型以拟合仪器噪声的模式。为此，我们首先执行离散余弦变换（DCT）将低信噪比谱（x）变换到像素空间频率域，然后将DCT系数输入到预训练的AUnet模型。接下来，通过对预测的DCT系数执行逆DCT（IDCT）来获得预测的仪器噪声（n）。从低信噪比数据中减去预测的仪器噪声，就可以得到高信噪比的光谱。这样，低信噪比谱的DCT系数及其对应的工具噪声形成匹配数据集以执行监督训练。在NL中，使用噪声频谱及其相应的仪器噪声的离散余弦变换（DCT）系数数据对来训练网络。网络的输出系数，然后处理的逆DCT（IDCT）操作，以重建预测的仪器噪声。DCT和IDCT可以描述为：

$$Y(k) = \sqrt{\frac{2}{N}} \sum_{m=1}^N X(m) \cos\left(\frac{(2m+1)k\pi}{2N}\right) \quad (1)$$

$$X(m) = \sqrt{\frac{2}{N}} \sum_{k=1}^N Y(k) \cos\left(\frac{(2k+1)m\pi}{2N}\right) \quad (2)$$

该网络使用MSE损失函数进行优化，可以描述为：

$$loss = \frac{1}{m} \sum_{i=1}^m |y(i) - x(i)|^2$$

AUnet模型是基于1-D U-net的主干建立的，由于其出色的性能，1-D U-net已被广泛应用于各种应用 [2]。最重要的是，我们进一步利用通道和空间注意力模块来改进模型，这些模块已被证明在提高网络学习能力方面是有效的。简而言之，每个特征在输入到下一层之前都通过注意力模块进行了细化。AUnet模型如图 2所示。

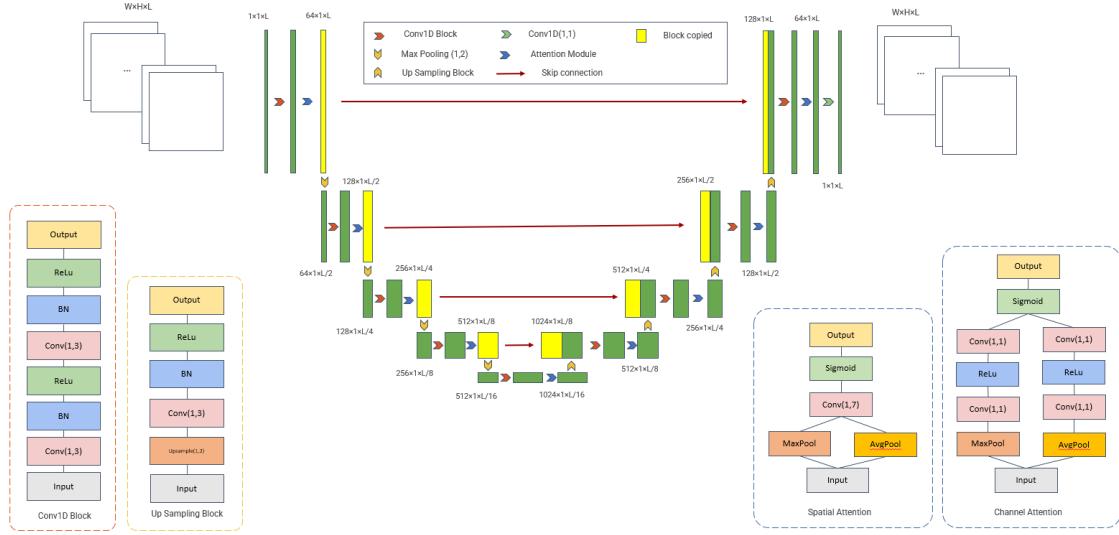


图 2. AUnet模型结构图

ChannelAttention通过初始化自适应平均池化和最大池化层来捕捉输入数据的全局信息，然后利用一个共享的多层感知机（MLP）处理这些池化后的特征，MLP由两个一维卷积层组成，中间夹着一个ReLU激活函数，用于学习通道间的依赖关系，最后通过Sigmoid激活函数输出通道注意力权重，这些权重与原始输入相乘以增强重要特征。而SpatialAttention则通过计算输入数据的平均值和最大值来获取空间信息，并将这两个值在通道维度上拼接，然后通过一个一维卷积层和Sigmoid激活函数来学习空间注意力权重，这些权重用于加权原始输入，从而突出重要的空间位置。这两种注意力机制分别增强了模型对输入数据特征和位置的敏感度。

3.2 数据集生成方法

与传统的监督学习（CSL）不同，NL利用基于先验的光谱生成器，使用拉曼信号线轮廓的先验知识产生高信噪比的光谱。同时，NL提出了一种数据生成方法，该方法使其能够用估计的仪器噪声生成匹配的低信噪比数据。这样，NL仅在生成的高低信噪比光谱数据集上进行训练，而无需从实际样本中获取大量光谱。由于NL学习了仪器噪声，它的性能是仪器依赖的。CSL与NL的示意如图 3所示。

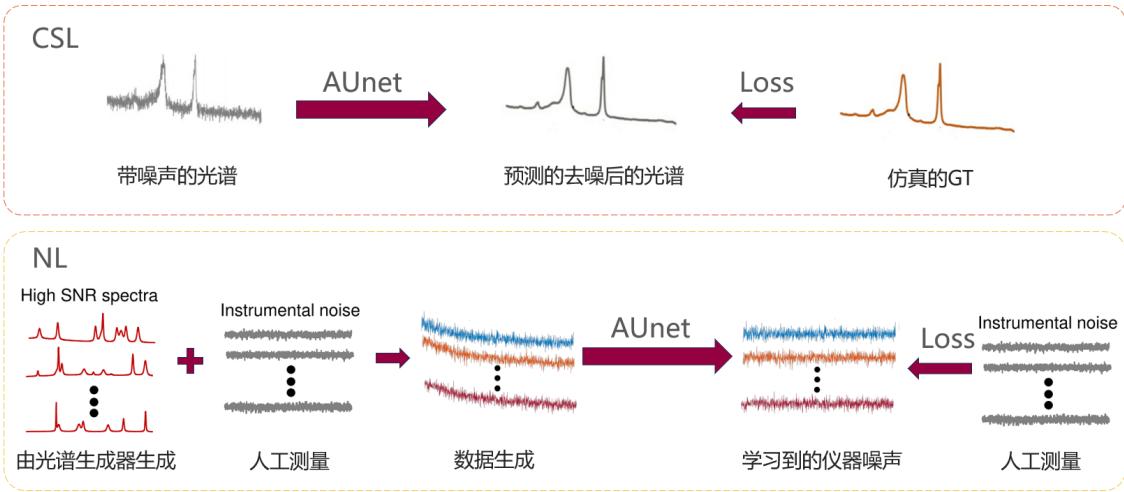


图 3. CSL与NL方法对比

在典型的CSL去噪算法中，模型训练需要对样本分别获取高信噪比和低信噪比的数据对，因此，在这样的数据集上训练的深度学习模型的性能取决于样本。作为替代方案，NL引入了基于先验的光谱生成器和仪器噪声估计方法来生成数据集。拉曼仪器的仪器噪声（包括光子散粒噪声、电子读出噪声和暗电流噪声[7]）使用拉曼非活性样品（这里使用平坦的Au膜）测量，然后通过基于奇异值分解（SVD）的方法从测量的光谱估计。通过这样做，仪器噪声近似地被认为是加性的。这是合理的，因为加性噪声源是仪器散粒噪声的主要贡献，特别是在低信噪比条件下。具体地说，由探测器和仪器背景产生的散粒噪声是仪器噪声的主要来源。因此，所提出的方法致力于估计这种噪声的贡献和预期的拉曼（光致发光）信号从测量的光谱。尽管人们认为噪声是随机的，但我们发现每个拉曼仪器在像素空间频域中都表现出独特的、仪器特定的、统计上稳定的噪声模式，这构成了使用深度学习在频域中拟合仪器噪声的基础。

3.2.1 仪器噪声估计

拉曼显微镜中的仪器噪声是通过从光滑的Au膜获得的光谱间接估计的，该Au膜不显示可观察到的拉曼峰。注意，Au膜仍然可以发射光致发光（PL）信号并反射光，这需要从Au光谱中消除。为此，利用奇异值分解（SVD）[4]来估计测量的Au光谱的这些光谱成分。频谱被排列为二维矩阵，然后通过奇异值分解（SVD）将其分解为：

$$S_{m \times n} = U_{m \times n} \Sigma_{m \times n} V_{m \times n}^T = [u_1 \cdots u_m] \cdot \begin{bmatrix} \Sigma_k & 0 \\ 0 & 0 \end{bmatrix} \cdot \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix}, \quad \Sigma_k = \begin{bmatrix} \sigma_1 & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & \sigma_k \end{bmatrix}$$

其中m表示由CCD检测器的有效读出像素确定的每个拉曼光谱中的数据点的数量，n表示矩阵中光谱的总数。之后，可以通过去除上述方程中的背景分量来估计仪器噪声。这可以通过手动检查左奇异矩阵（ u_1, \dots, u_m ）的每个列向量来完成，或者简单地移除第一个分量（ $\sigma_1 \cdot u_1 \cdot v_1^T$ ），因为它贡献了大部分背景。通过将J表示为背景分量指数集，仪器噪声可以估计如下：

$$\text{Noise}_{m \times n} = S_{m \times n} - \sum_j \sigma_j \cdot u_j \cdot v_j^T$$

拉曼显微镜仪器噪声测量与分析方案如图 4 所示。(a)一种测量光滑 Au 膜仪器噪声的装置方案。(b)Au 薄膜的噪声谱。(c)噪声谱的二维矩阵的奇异值 (SV)，其通过奇异值分解 (SVD) 获得。去除与背景相关的 SV (BG，红色)，并将剩余 SV (黑色) 相加以获得仪器噪声谱。(d)重建的仪器噪声谱。(e)用快速傅立叶变换 (FFT) 或离散余弦变换 (DCT) 对每个仪器噪声谱进行变换，以获得每个仪器在频域中的平均功率谱。功率谱示出了清晰的图案，其可以进一步分离成白噪声 (WN，即，功率谱的基线) 和固定模式噪声 (FPN，即特定频率的噪声源)。

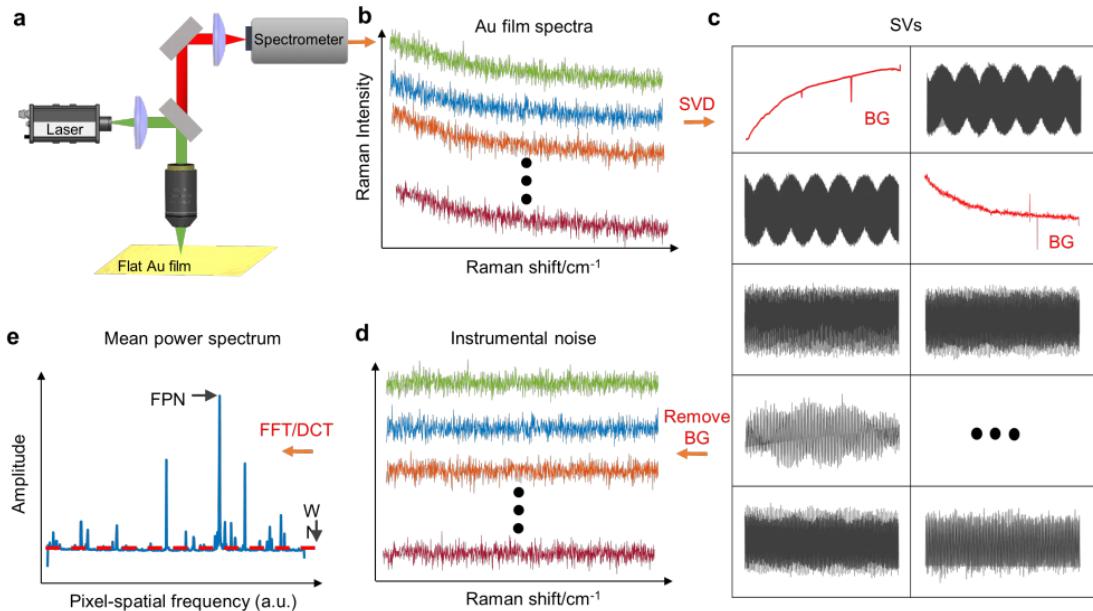


图 4. 拉曼显微镜仪器噪声测量与分析方案

3.2.2 高信噪比数据生成

由于分子与相邻分子的相互作用以及导致峰展宽的各种因素，拉曼峰的线型不能简单地用高斯或洛伦兹函数拟合。在这里，使用 Voigt 函数拟合更合适，Voigt 函数是高斯和洛伦兹函数的卷积。为了简单起见，我们使用 Voigt 函数的伪版本来生成拉曼峰 [3]：

$$V(x, \mu, w) = \rho \times G(x, \mu, w) + (1 - \rho) \times L(x, \mu, w)$$

$$G(x, \mu, w) = \frac{1}{\sigma \sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \quad \sigma = \frac{w}{2\sqrt{2 \ln 2}}$$

$$L(x, \mu, w) = \frac{1}{\pi} \cdot \frac{w/2}{(x - \mu)^2 + (w/2)^2}$$

其中 μ 为峰值位置， w 为半高宽， G 为高斯部分， L 为洛伦兹部分， ρ 为高斯函数对总 Voigt 函数的贡献， $\rho = 0.6785$ 。这样就可以产生如下的多峰拉曼光谱：

$$R = \sum_{i=1}^N S(i) = \sum_{i=1}^N A(i) \times V(x, U(i), W(i))$$

其中 A 表示信号的幅度， U 表示峰位置， W 表示半高宽， N 表示拉曼峰的总数。同样，基线也是使用相同的方法生成的。这些参数在给定的范围内随机提取，应在特定的拉曼仪器上

进行优化。例如，当训练Horiba仪器的AUnet模型时，拉曼峰的数量被设置在1-10的范围内，FWHM被设置在5 - 200 cm⁻¹的范围内，峰位置被设置在1 -1600 cm⁻¹的范围内。

4 复现细节

4.1 与已有开源代码对比

在本研究中，我们严格遵循论文作者所开源的实验源代码，依据论文详述的实验需求，搭建并配置了相应的实验环境 [5]。与此同时，对实验数据集进行了系统的整理与预处理，以确保数据的质量与可用性，为后续模型训练奠定坚实基础。在此基础上，我们运用所配置的环境与整理好的数据，成功复现了论文中的验证实验，实现了对原研究实验流程的有效再现。

进一步地，针对复现模型展开深入剖析。经研究发现，源代码中所采用的模型结构为一维卷积神经网络，该模型在处理数据时，仅以每个像素点的光谱数据作为输入，这种方式在一定程度上忽视了数据中的空间相关性信息。然而，在实际的应用场景中，空间相关性对于精准建模与预测起着不可忽视的作用，其包含的位置关系、邻域特征等信息对于提升模型性能具有潜在价值。

鉴于此，为优化模型性能，后续改进工作可考虑将模型架构由一维卷积调整为二维卷积神经网络。二维卷积能够同时兼顾空间与光谱两个维度的信息，通过在两个维度上的滑动卷积操作，有效学习空间维度上的位置特征以及光谱维度上的光谱特征之间的映射关系，有望显著提升模型对复杂数据的表征能力与泛化性能，从而推动相关研究向更精准、高效的方向发展。

4.2 实验环境搭建

本实验基于PyTorch框架构建与优化了模型，并在64位Windows 10操作系统上进行了模型的训练与测试。该实验平台包括一颗Intel(R) Xeon(R) CPU E5-2683 v3 @2.00GHz处理器，以及32GB的内存容量。为了显著提升网络训练的速度，利用了NVIDIA GeForce RTX 3060显卡进行GPU加速，其显存容量为12GB。

实验选择了Python 3.8作为编程语言，并安装了与之兼容的PyTorch 1.8.1版本。为了充分发挥GPU的性能优势，实验配置了CUDA 11.3和cuDNN 8.2.1，这两个库为深度学习应用提供了高效的GPU加速功能。通过这样的软硬件配置，确保了在实验过程中能够充分利用计算资源，快速而准确地完成模型的训练与测试。

除了PyTorch和GPU加速库外，实验还安装了一系列必要的依赖包，包括einops、h5py、lmbdb、matplotlib、numpy、opencv python、Pillow、scikit image、scipy、six、tensorboardX、thop、timm、torchnet、torchstat和torchvision等。这些依赖包为模型的构建、数据处理、可视化等方面提供了必要的支持。

4.3 模型训练

用NL方法训练好的模型可以很好地拟合由特定拉曼仪器测量的不同噪声水平的原始拉曼光谱的噪声分布。为了验证这一点，我们获得了商业拉曼显微镜（LabRAM HR-Evolution）

的噪声光谱与不同的积分时间（0.1-0.5 s）在一个平坦的Au膜，而保持其他条件不变。我们训练通过NL方法的AUnet模型只有12500个噪声光谱（每个积分时间2500个光谱），并且GT拉曼光谱在每个训练步骤中使用所提出的基于物理的方法随机生成。

5 实验结果分析

5.1 评价指标

实验使用了三个关键的性能指标：信噪比（Signal-to-Noise Ratio, SNR）、峰值信噪比（Peak Signal-to-Noise Ratio, PSNR）以及结构相似性指数（Structural Similarity Index, SSIM）。在数值方面，SNR、PSNR和SSIM的值越高，反映出降噪效果越显著，图像质量越好 [9]。SNR应该等于信号与噪声的功率谱之比，但通常功率谱难以计算，有一种方法可以近似估计图像信噪比，即信号与噪声的方差之比。首先计算图像所有象素的局部方差，将局部方差的最大值认为是信号方差，最小值是噪声方差，求出它们的比值，再转成dB数，最后用经验公式修正，计算方法为：

$$SNR = 10 \lg \frac{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} g(i, j)^2}{\sum_{i=0}^{M-1} \sum_{j=0}^{N-1} [g(i, j) - f(i, j)]^2}$$

PSNR是最普遍和使用最为广泛的一种图像客观评价指标，然而它是基于对应像素点间的误差，即基于误差敏感的图像质量评价，由于并未考虑到人眼的视觉特性，因而经常出现评价结果与人的主观感觉不一致的情况。PSNR的计算表达式为：

$$MSE = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2$$

$$PSNR = 10 \log_{10} \left(\frac{(2^n - 1)^2}{MSE} \right)$$

其中，MSE表示当前图像 X 和参考图像 Y 的均方误差（Mean Square Error, MSE），H、W 分别为图像的高度和宽度，n 为每像素的比特数，一般取8，即像素灰阶数为256。

SSIM也是一种全参考的图像质量评价指标，它分别从亮度、对比度、结构三方面度量图像相似性，计算过程为：

$$SSIM(X, Y) = l(X, Y) \cdot c(X, Y) \cdot s(X, Y)$$

其中：

$$l(X, Y) = \frac{2\mu_X\mu_Y + C_1}{\mu_X^2 + \mu_Y^2 + C_1}$$

$$c(X, Y) = \frac{2\sigma_X\sigma_Y + C_2}{\sigma_X^2 + \sigma_Y^2 + C_2}$$

$$s(X, Y) = \frac{\sigma_{XY} + C_3}{\sigma_X\sigma_Y + C_3}$$

并且：

$$\mu_X = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W X(i, j)$$

$$\sigma_X^2 = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - \mu_X)^2$$

$$\sigma_{XY} = \frac{1}{H \times W - 1} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - \mu_X)(Y(i, j) - \mu_Y)$$

其中 $C_1 = (K1 * L)^2$, $C_2 = (K2 * L)^2$, $C_3 = (K3 * L)^2$, 一般地, $K1 = 0.01$, $K2 = 0.03$, $L = 255$ 。

这些公式用于衡量两幅图像之间的结构相似性。以下是各参数的解释:

$\text{SSIM}(X, Y)$: 结构相似性指数, 用于评估两幅图像 X 和 Y 的相似程度。 $l(X, Y)$: 亮度比较函数, 比较两幅图像的平均亮度。 $c(X, Y)$: 对比度比较函数, 比较两幅图像的对比度。 $s(X, Y)$: 结构比较函数, 比较两幅图像的结构信息。 μ_X 和 μ_Y : 分别是图像 X 和 Y 的均值。 σ_X 和 σ_Y : 分别是图像 X 和 Y 的标准差 (方差的平方根)。 σ_{XY} : 图像 X 和 Y 的协方差。 C_1 、 C_2 、 C_3 : 常数, 用于避免分母为零的情况。

在实际应用中, SSIM值越接近1, 表示两幅图像越相似。

5.2 AUnet与其他模型对比

由于深度学习模型是使用随机生成的数据进行训练的, 因此来自真实样品的拉曼光谱的量化性能是未知的。我们进一步使用一种更现实的方法来生成匹配的低SNR和高SNR拉曼光谱, 使用相同的拉曼仪器从真实的2D材料样品中获得的原始数据。石墨烯、MoSe2和WS2薄层的代表性结果如图所示。可以看到, SG和小波方法可以提高谱的信噪比, 但存在伪影和欠平滑, 相比之下, AUnet很好地恢复了GT光谱, 即使原始光谱的SNR非常低或为负。为了统计量化这一观察结果, 实验计算了来自三个2D材料样本的测试数据集的MSE和SNR (总共37500个光谱, 每个样本12500个光谱)。如图 5所示, 与传统的去噪方法相比, 使用AUnet可以大大改善MSE和SNR。特别在模拟的石墨烯拉曼光谱上, 平均SNR可以提高到22.3 dB (约10倍), 并且与原始数据相比, 平均MSE可以降低超过149倍。在所有情况下, AUnet处理的光谱都能很好地再现原始光谱, 并大大增强了SNR。这表明NL的良好适用性, 即使相同的AUnet模型, 没有重新训练, 用于不同的样本。

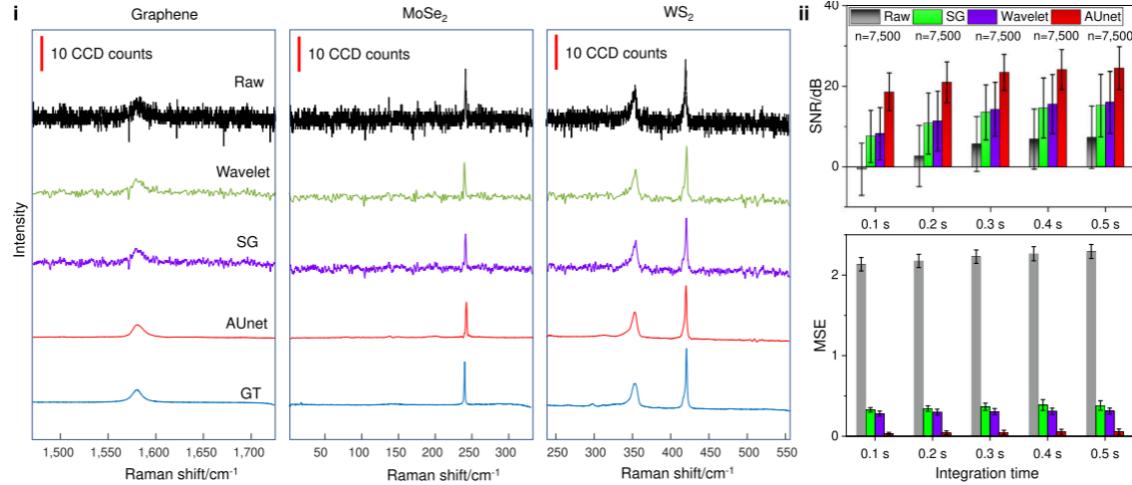


图 5. AUnet与SG、Wavelet的性能比较

5.3 NL方法与CSL方法对比

在展示了NL在模拟拉曼光谱恢复上的能力及其相对于传统去噪算法的优势之后，我们进一步验证它是否适用于从真实的样品上获得的光谱，以及它是否可以成为CSL方法的上级的替代方案。为此，我们通过CSL方法训练了一个1-D U-Net模型，该数据集由同一个拉曼显微镜产生。训练数据集包含12500个低信噪比和高信噪比光谱对，这些光谱对是通过图6中描述的方法获得的，实验在Cr/Si光栅样品上进行数据集的采集，光谱范围为289.81–604.13cm⁻¹。原始数据模拟生成高信噪比拉曼光谱分为两个步骤：(1)将原始成像数据堆叠成三维(3D)数据立方体，对每个图像切片进行二维移动均值滤波。(2)采用自适应低秩矩阵近似(ALRMA)[1]方法进一步抑制噪声，得到高信噪比的GT数据。

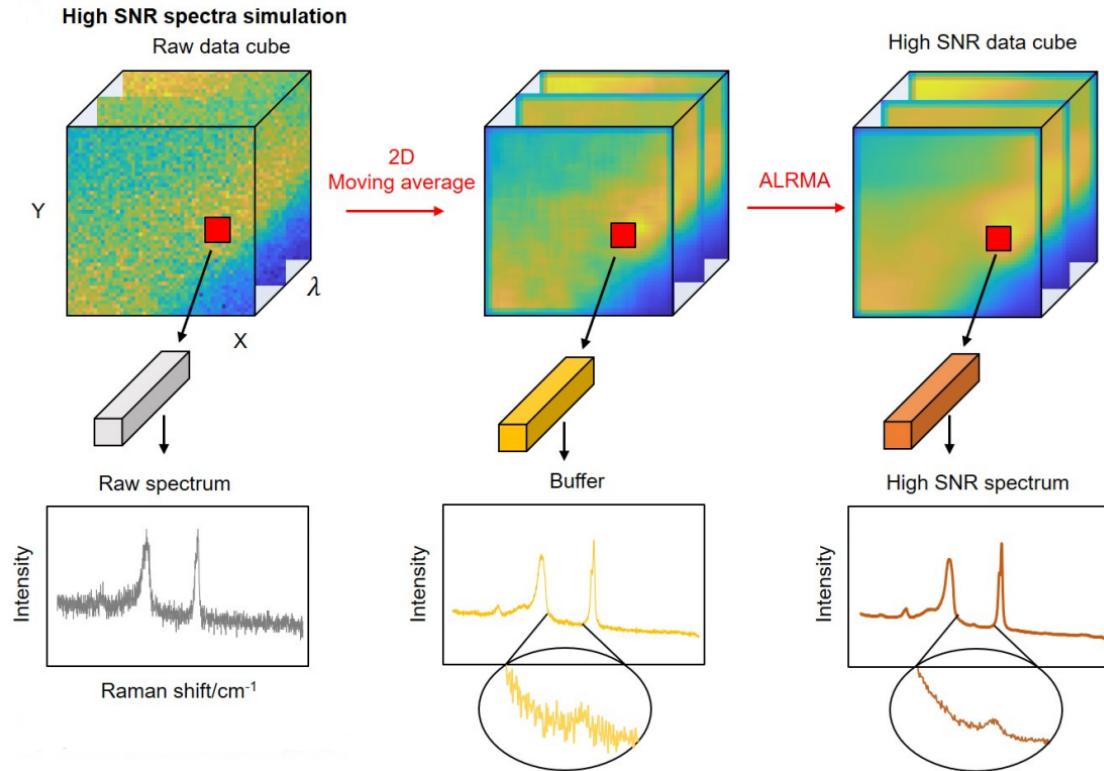


图 6. CSL方法获得GT数据方式

训练好的CSL模型然后被用于恢复从不同样品中得到的原始光谱，包括“可见”和“不可见”样品，而这里使用的NL模型与第一个实验中使用的NL模型相同，没有进一步的训练。可以预期的是，CSL模型在训练阶段“看到”的数据上表现良好。在520.6 cm⁻¹处的三个Si特征峰的拉曼图像如图7所示。与原始光谱图像相比，通过CSL或NL方法逐像素去噪后的恢复图像显示对比度大大提高，并且与亮场图像匹配良好。图7右侧显示了Cr和Si点位置的代表性原始、CSL和NL恢复光谱，表明CSL和NL对该特定样品都很有效。通过比较恢复的光谱与图7左侧中白色矩形标记的小区域上平均的局部平均光谱，可以验证这一点。

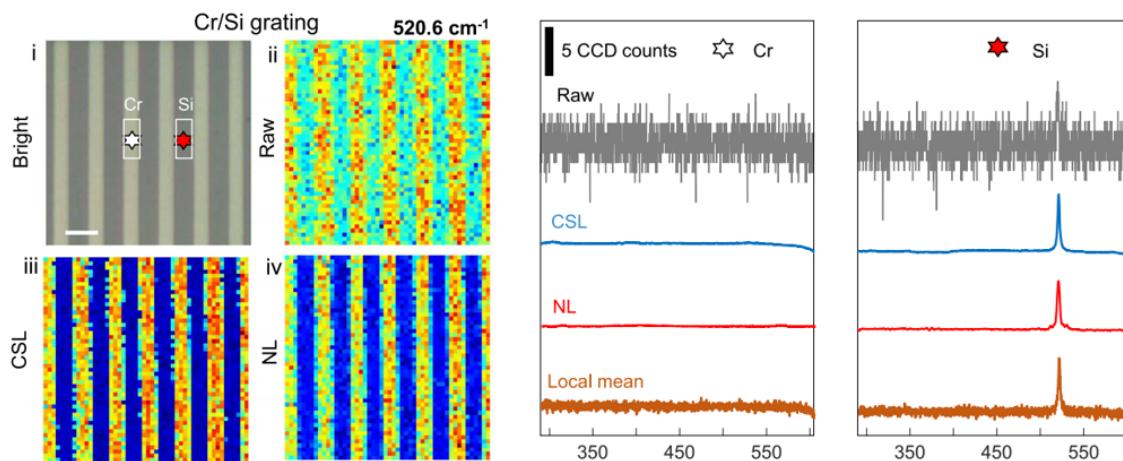


图 7. NL和CSL在Cr Si光栅上的去噪效果

我们进一步将在Cr/Si光栅样品上训练的CSL模型应用于“看不见”的样品，如石墨烯（放置在Si/SiO₂衬底上）和HeLa细胞，发现其性能严重恶化。具体地，通过CSL处理的G带

(1581 cm^{-1}) 的拉曼图像几乎不能区分4层 (4 L) 石墨烯区域与衬底。但在NL处理的拉曼图像中可以很容易地分辨出这种微小的差异。此外，图 8右侧中的拉曼光谱表明，对于8 L和4 L石墨烯区域，通过与局部平均光谱进行比较，NL可以很好地恢复本征峰特征，相比之下，CSL光谱显示出明显的伪影和过度平滑时。

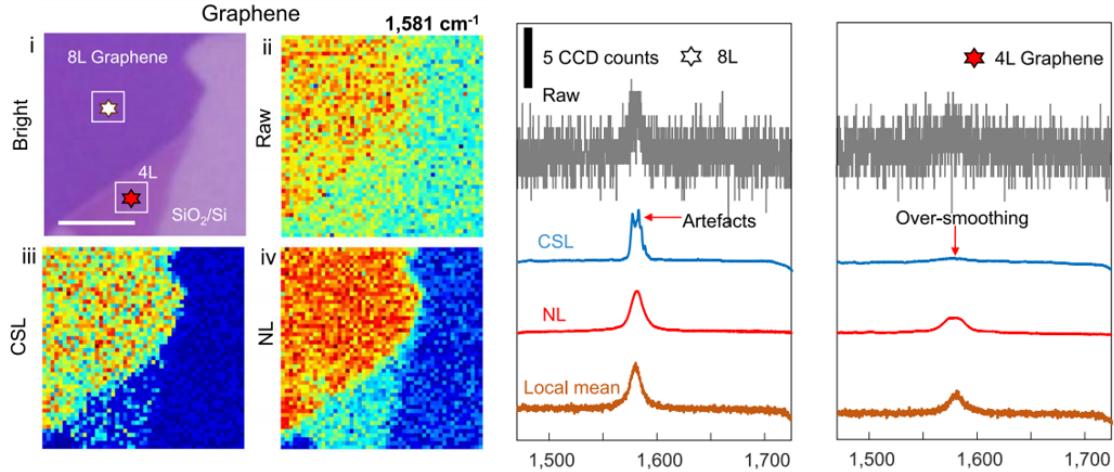


图 8. NL和CSL在石墨烯上的去噪效果

然后在HeLa细胞样品上进行类似的实验。在这种情况下，NL处理的拉曼图像显示出优于CSL图像的SNR和对比度,如图 9所示。特别地，CSL光谱在处理细胞外部和细胞内部获得的拉曼光谱时存在更多的伪影和欠平滑问题。相反，NL光谱与局部平均光谱（白色矩形标记的 11×11 像素的平均值）吻合得很好，这在 $2800\text{-}3000\text{ cm}^{-1}$ 附近的CH丰度带中可以清楚地看到。这些结果可以证明，NL方法允许将单个深度学习模型应用于恢复由同一拉曼显微镜获得的各种拉曼光谱，这是通过CSL方法无法实现的。

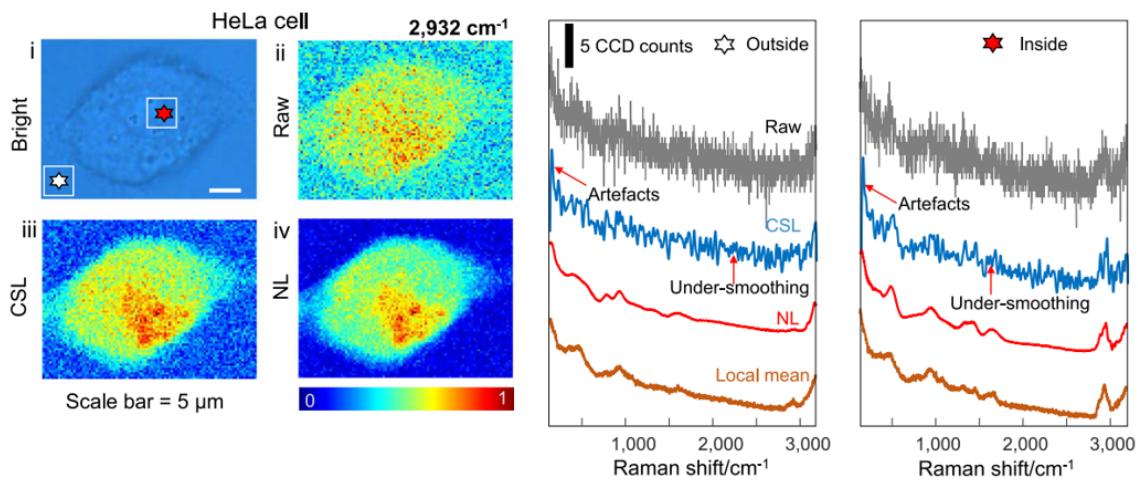


图 9. NL和CSL在HeLa细胞上的去噪效果

为了证明NL在不同仪器上的可推广性，我们在另一个商业线扫描拉曼显微镜 (Raman-11, Nanophoton Inc) 上训练第二个AUnet模型。AUnet模型通过NL方法训练，数据量为目标显微镜的400000个仪器噪声谱。然后使用训练好的AUnet模型来恢复HeLa细胞的线扫描荧光和拉曼成像以及用同一仪器获得的2D材料的光致发光成像。活细胞荧光成像的一个关键问题

是荧光团在长期成像过程中由于有限的光子预算而可能发生的光漂白。同时，激光照射会对易碎的生物样品产生光毒性。因此，低激光功率是优选的，但以降低SNR为代价。在此，我们证明了这种限制可以通过基于AUnet的降噪技术来基本消除。我们强调，我们的经NL方法训练的AUnet模型可直接应用于同一仪器所获得的不同种类的光谱。为了证实这一点，我们采用了相同的模型对HeLa细胞的无标记拉曼成像进行去噪，实验结果如图 10所示。在指纹区的两个谱带（1445和1612 cm⁻¹: 脂类和蛋白质的CH₂和CH₃变形，1656 cm⁻¹: 脂类的C=C伸缩）和在CH丰富区的两个谱带（2867 cm⁻¹: 脂类的CH 2对称伸缩，2915 cm⁻¹: 脂类和蛋白质的CH谱带）的拉曼图像被呈现出来 [10]。这些波段的原始图像是有噪声的，不可分析的，尽管激光功率很高（c.a.50 mW/线）和长积分时间（5s/线）。在用AUnet逐像素去除噪声之后，这些拉曼图像的对比度大大提高，并且可以清楚地观察到细胞内脂质和蛋白质的液滴状分布。这表明NL方法在下游生物分析中的潜在应用，甚至可能用于真实的分析。在同一仪器上用相同的AUnet模型获得的荧光和拉曼图像的成功去噪表明NL确实是仪器依赖性的而不是样品依赖性的。

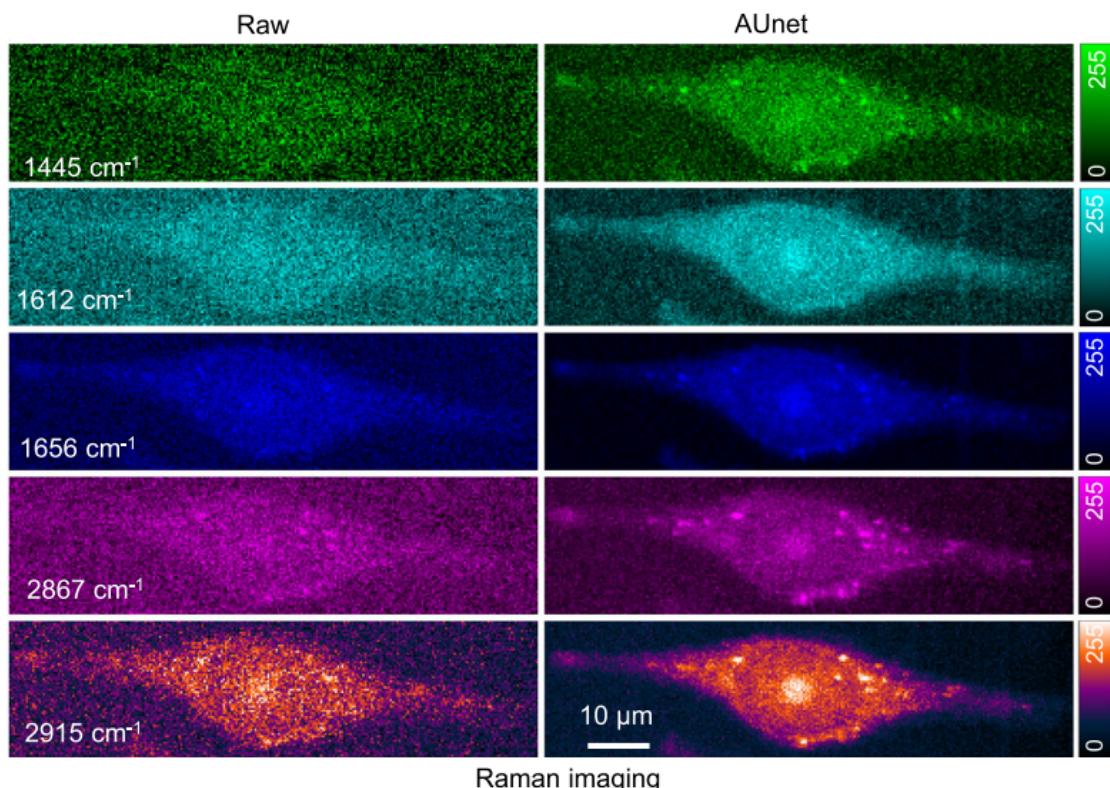


图 10. 在Raman-11上新模型的去噪效果

6 总结与展望

我们探索了一种NL方法来学习来自不同的拉曼仪器的特征噪声分布，以便仅使用估计的仪器噪声以及生成的GT数据来训练深度学习模型，而无需费力的手动标记。这使得一个单一的AUnet模型可以适用于同一仪器在不同实验条件下获得的不同样品的多种拉曼光谱。这绕过了传统的监督机器学习技术的局限性，其性能严重依赖于现有的注释数据集，并且只适用于“看到”的样本。我们已经在从体硅到2D材料和活细胞的多个例子中表明，通过NL方法

训练的AUnet可以大大提高拉曼光谱的SNR。

NL在广泛的光谱场景中增强了SNR，以及共焦拉曼、荧光和光致发光的显微成像。NL方法将深度学习的性能从依赖于样本转变为依赖于工具。这具有提高所有光学显微镜/光谱学模态的性能的潜力，而不需要通过有效的计算方法进行硬件修改，为去噪领域的研究与应用开辟了新的道路。

参考文献

- [1] M. Bebendorf and S. Rjasanow. Adaptive low-rank approximation of collocation matrices. *Computing*, 70(1):1–24, 2003.
- [2] Thorsten Falk, Dominic Mai, Robert Bensch, Ozgun Cicek, Ahmed Abdulkadir, Yassine Marakchi, Anton Bohm, Jan Deubner, Zoe Jackel, and Katharina Seiwald. U-net: deep learning for cell counting, detection, and morphometry (vol 16, pg 67, 2019). *Nature methods*, (4):16, 2019.
- [3] Masashi Fukuhara, Kazuhiko Fujiwara, Yoshihiro Maruyama, and Hiroyasu Itoh. Feature visualization of raman spectrum analysis with deep convolutional neural network. 2020.
- [4] Qiang Guo, Caiming Zhang, Yunfeng Zhang, and Hui Liu. An efficient svd-based method for image denoising. *IEEE Transactions on Circuits and Systems for Video Technology*, 26(5):868–880, 2016.
- [5] Hao He, Maofeng Cao, Yun Gao, Peng Zheng, Sen Yan, Jin Hui Zhong, Lei Wang, Dayong Jin, Bin Ren, and Nathalie Le Bot. Noise learning of instruments for high-contrast, high-resolution and fast hyperspectral microscopy and nanoscopy. *Nature Communications*, 2024.
- [6] Conor C Horgan, Magnus Jensen, Anika Nagelkerke, Jean Philippe St-Pierre, Tom Vercauteren, Molly M Stevens, and Mads S Bergholt. High-throughput molecular imaging via deep-learning-enabled raman spectroscopy. *American Chemical Society (ACS)*, 2021.
- [7] Izabella Jolan Jahn, Alexej Grjasnow, Henry John, Karina Weber, Jürgen Popp, and Walter Hauswald. Noise sources and requirements for confocal raman spectrometers in biosensor applications. *Sensors*, (15), 2021.
- [8] Xiang Wang, Sheng Chao Huang, Shu Hu, Sen Yan, and Bin Ren. Fundamental understanding and applications of plasmon-enhanced raman spectroscopy. *Nature Reviews Physics*.
- [9] Xin Wang, Guokun Liu, Mengxi Xu, Bin Ren, and Zhong Qun Tian. Development of weak signal recognition and extraction algorithm for fast raman imaging. *Analytical Chemistry*, 91(20), 2019.
- [10] Ying Xu, Xiaoli Hou, Qiaoqiao Zhu, Shijie Mao, Jie Ren, Jidong Lin, and Ning Xu. Phenotype identification of hela cells knockout cdk6 gene based on label-free raman imaging. *Analytical chemistry*, (25):94, 2022.