题目：基于自注意力机制进行计算光谱重建

作者：吴豪1，吴辉1，唐峰2， 郑万国1,2,\*，刘爽利1,\*，叶鑫2,\*

1.西南科技大学 信息工程学院

2.中国工程物理研究院，激光聚变中心

摘要：

目前，光谱仪可以提供超精细的分辨率和较宽的光谱范围，但是其系统结构复杂、体积庞大。且随着光谱分析的应用空间迅速增长，便携式或集成式光谱分析设备对降低物理尺寸、成本或功耗的需求优先于对高性能的需求。而新出现的一种微型计算光谱仪需要通过计算对其进行光谱重建，但现在基于传统算法的光谱重建准确性并不是很理想。因此本研究考虑通过重建矩阵的设计和使用自注意力机制算法来进一步提高光谱重建准确性。通过对仿真的透过率函数库使用互相关算法筛选出100个稀疏的透过率函数并设计成编码矩阵，然后将自注意力机制算法应用于光谱重建中，并且与商用光谱仪（型号Optics-\*\*）进行对比取得了R2 0.9780和MSE0.0019的结果，该方法对于光谱重建技术的改进和优化具有重要参考的意义。

引言：

计算光谱重建是一种新型的光谱重建方法，区别与传统的光栅光谱仪直接进行分光或者滤波片的设计方式从而直接得到物体的光谱信息不同。计算光谱系统更多通过设计不同的材料层和编码矩阵，它可以利用计算机算法来处理光学信号，从而实现光谱的重建，得到物体的光谱信息。与传统的光谱仪相比，计算光谱重建型的光谱仪具有成本低、体积小和易于操作等优点，因此在许多应用场合具有广泛的应用前景。**错误!未找到引用源。** -[6]

近年来随着深度学习的出现，能够解决一系列复杂的问题，深度神经网络能够通过训练大量的例子来学习输入和目标输出分布之间的映射[7]，对其加入自注意力机制能够更好的学习数据中更加相关的特征，注意力机制已经被广泛应用在语言识别[8]，文本分类[9][10]，言语翻译[11]，自然语言生成[12]，图像识别[13][14]，图像生成[15][16]等。然而目前应用在微型计算光谱仪上的算法通常是各种基于压缩感知改进的算法[17]。例如，在以前的研究中开发的光谱相机是基于CS的方法通过迭代获得了充分收敛的结果。而近期有研究表明，引入神经网络对计算光谱系统进行光谱重建所取得的良好结果，证明了深度学习对光谱重建有着重大的帮助，但由于只是利用全连接层本身的线性变换并没有考虑到光谱与光谱之间的关联性，导致光谱重建的准确率并不是很完美[18][19][20]。因此通过引入残差和自注意力机制来提升光谱重建的准确性和稳定性，其中残差机制可以对上一步更好的结果进行记录并保存，这样可以使得当网络层数过于深时防止其过拟合，而自注意力机制能够更好的提高光谱编码之类的相关性，从而更好提取特征进行重建。

因此为了得到更高的光谱重建精度，需要对仿真的透过率函数库进行了互相关之间的计算并筛选出相关性最弱并保留着编码效应的100条透过率函数的进行构建成一个编码矩阵，基于自注意力机制并添加了残差来对光谱数据集进行训练以得到更好的重建结果。

本研究的组织结构如下：首先，介绍了光谱相机是如何通过编码矩阵后得到的未解码的数据原理，并介绍了该数据如何通过设计的神经网络算法重建得到物体光谱信息的原理。随后，具体讲述如何利用设计编码矩阵和构建数据并且在构建数据集时通过添加不同的噪音系数以期模拟现实环境噪音，之后对重建结果进行了展示，其中分为抽样结果和整体重建结果。最后，讨论了基于自注意力机制能够取得更好的结果，并总结了工作。

结果：

光谱相机示意图

本节中演示了如何在太阳光下测量目标样本的反射光谱，简化原理图如图1所示。首先需要将设计好的编码矩阵贴在相机前方，用来测试样本的反射光的光谱信息，因此在进入相机转换前光谱强度是，然后经过编码矩阵和相机响应函数最后得到未解码的强度。其中设计的编码矩阵可以加工成和相机完全贴合的芯片，以便使得光谱测量更加便捷。

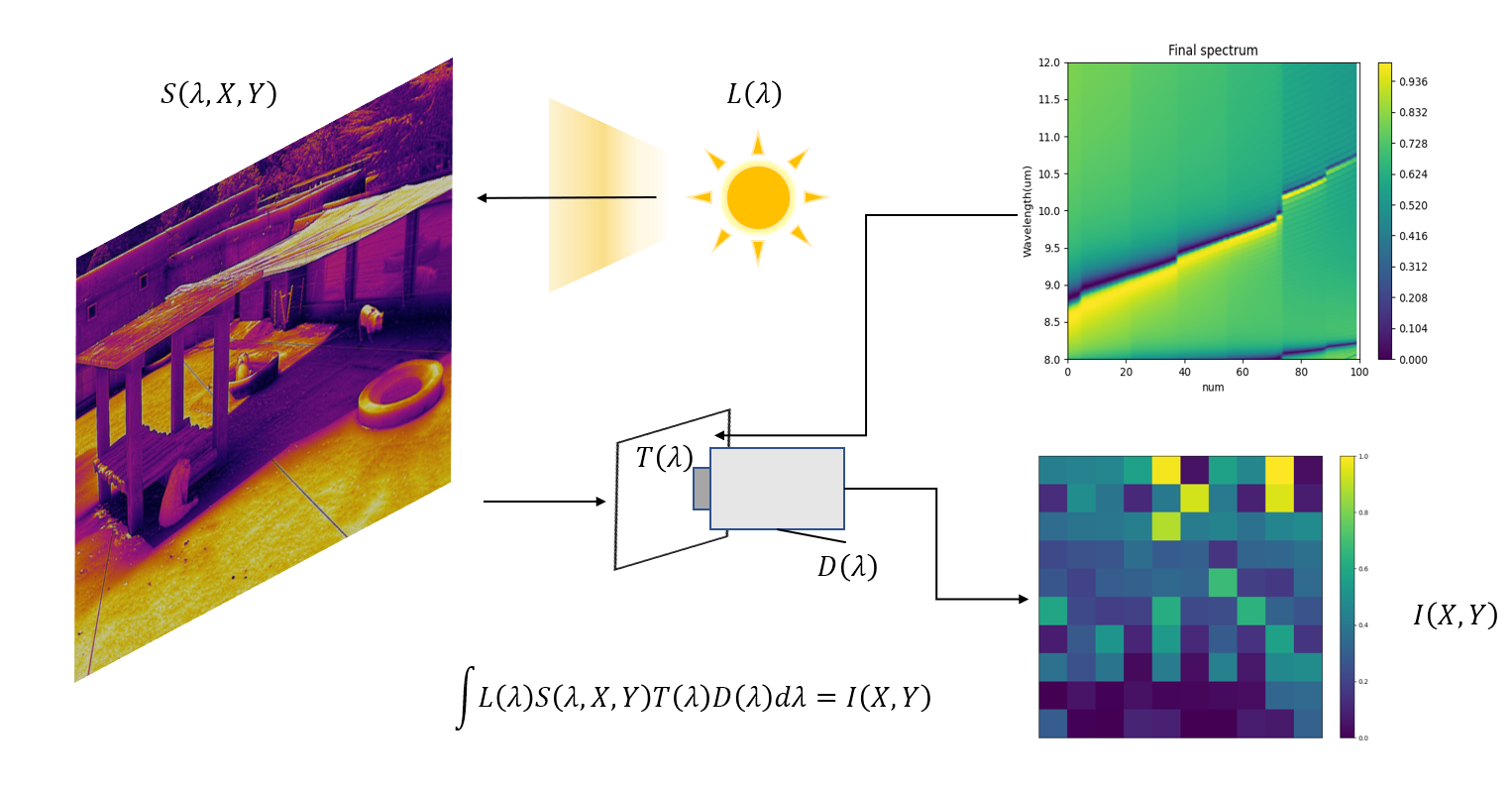


Fig. 1 目标光谱采集简易原理。当目标的光谱信号经过T编码矩阵后经过相机转换到未解码信号。

结合自注意机制的光谱重建

算法模型展示在图2的b，该模型为了从这个随机编码得到的数据得到更多的特征，使用了自注意力机制来加强特征之间的联系，以此来提取更有价值的特征便于更好的重建。自注意力机制是计算每个位置和其他位置之间的相似度，并将相似度作为权重，对特征进行加权求和，这种机制能够将注意力集中在重要的特征上，从而提高最终结果的准确性。残差机制是对神经网络层添加跨层连接，可以有效地避免梯度消失和梯度爆炸等问题，有助于提高模型的训练效果[22]。在这个框架就是利用自注意力机制和残差来进行训练，首先是对输入进的数据进行自注意力的运算，以得到更具有相关性的特征，然后在各个神经层引入残差，防止在训练的时候参数爆炸和过拟合。此外，训练的模型加入了正则化参数，进一步提高模型的鲁棒性[23][24]。一般在CS算法中，去噪很大程度依赖于先验，一般情况下在迭代过程中手动设置参数，以抵消噪音引起的偏差，虽然有一定的效果但当噪音水平发生变化后，它就不能提供令人信服的结果。

编码矩阵设计和数据集构建

先对FDTD仿真得到的滤波函数库如图3a，对其进行初筛得到一个新的滤波函数矩阵如图3b。初次筛选是计算各个滤波函数的期望和方差，保证每条滤波函数都有编码效应，因为没有编码效应的话会使得检测目标时没有光强变化以致无法解码得到光谱。最后利用互相关函数[25]对初筛后的滤波矩阵进行再次筛选从而得到了图3c的结果。

在图3中还展示了对模拟光谱数据的设计，添加了不同程度上的高斯噪音以便于模拟现实环境因素，来加强训练模型的鲁棒性。将模拟的光谱数据通过如下公式：

以此得到了对应的输入未解码的数据，其中是编码矩阵，是光谱数据，是添加的噪音水平，转换时是对每个滤波函数与光谱数据进行转换最后全部转换成如图2a所示的未解码数据。用高斯基函数模拟了2W条光谱原始光谱数据，然后对其分别添加5%，10%，15%，20%的高斯噪音，如图3e，f，以构成了10W条可训练的数据集，并将其8：2分为训练集和验证集。且对验证集是在10W条数据中随机选择以保证分布的合理性，最后拿实际测试的1000条光谱数据作为测试集来测试模型的准确率和鲁棒性。

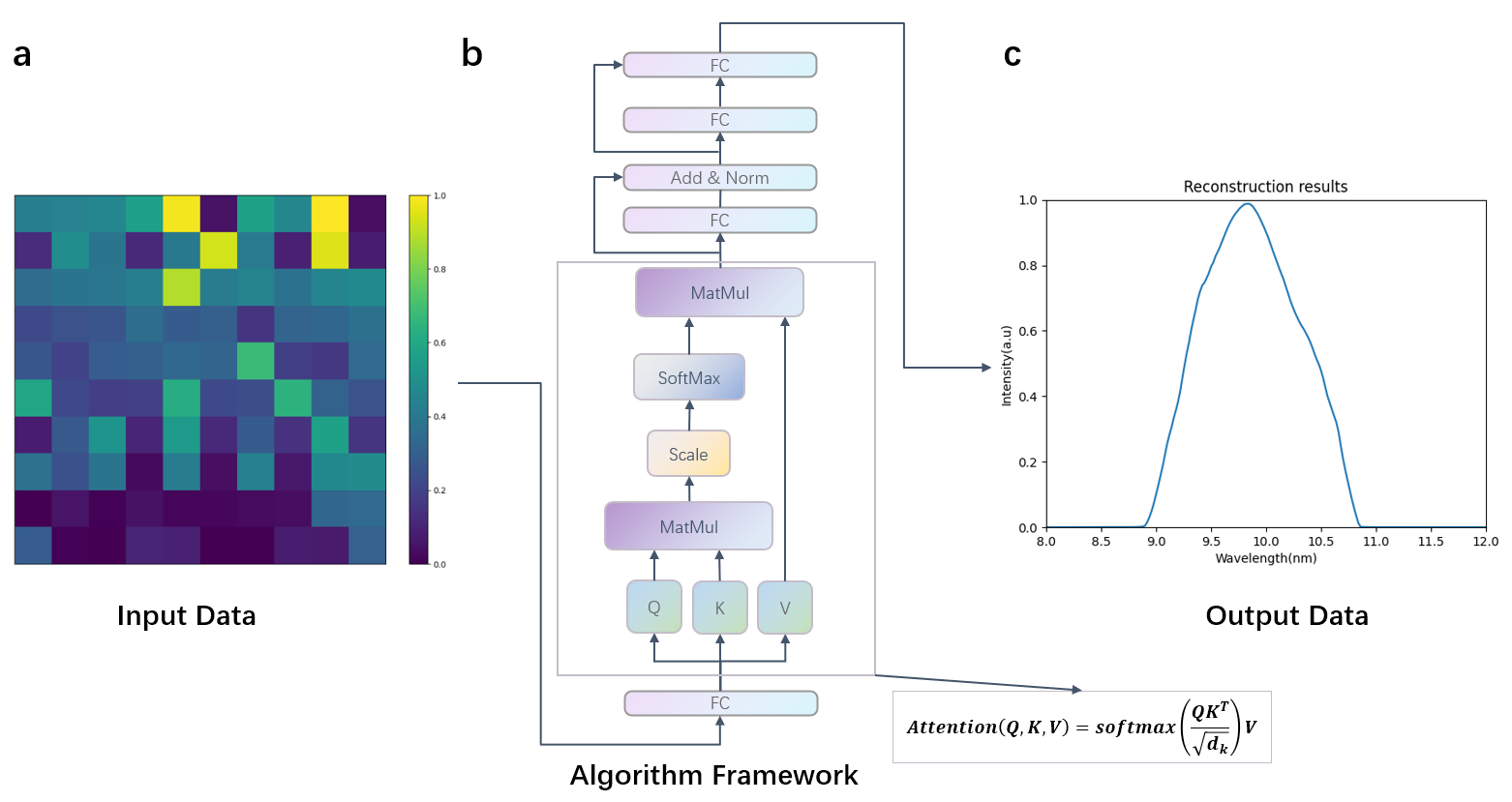


Fig. 2 基于自注意力机制的光谱重建算法原理。A 是由相机所接收到的未解码信号；B 是算法框架原理，并在旁边标注了数学原理；C 是最后重建出的目标样本的光谱信号。

重建结果展示

在这部分中，在图4a-f中展示了基于注意力算法模型重建的结果，并标注了相应的指标数值。还将测试集整体R2和MSE的平均结果展示出来。研究结果表明使用互相关算法筛选滤波函数并结合自注意力机制网络进行光谱重建，可以获得更好的重建结果，这对于光谱重建的实际应用具有重要的参考价值。

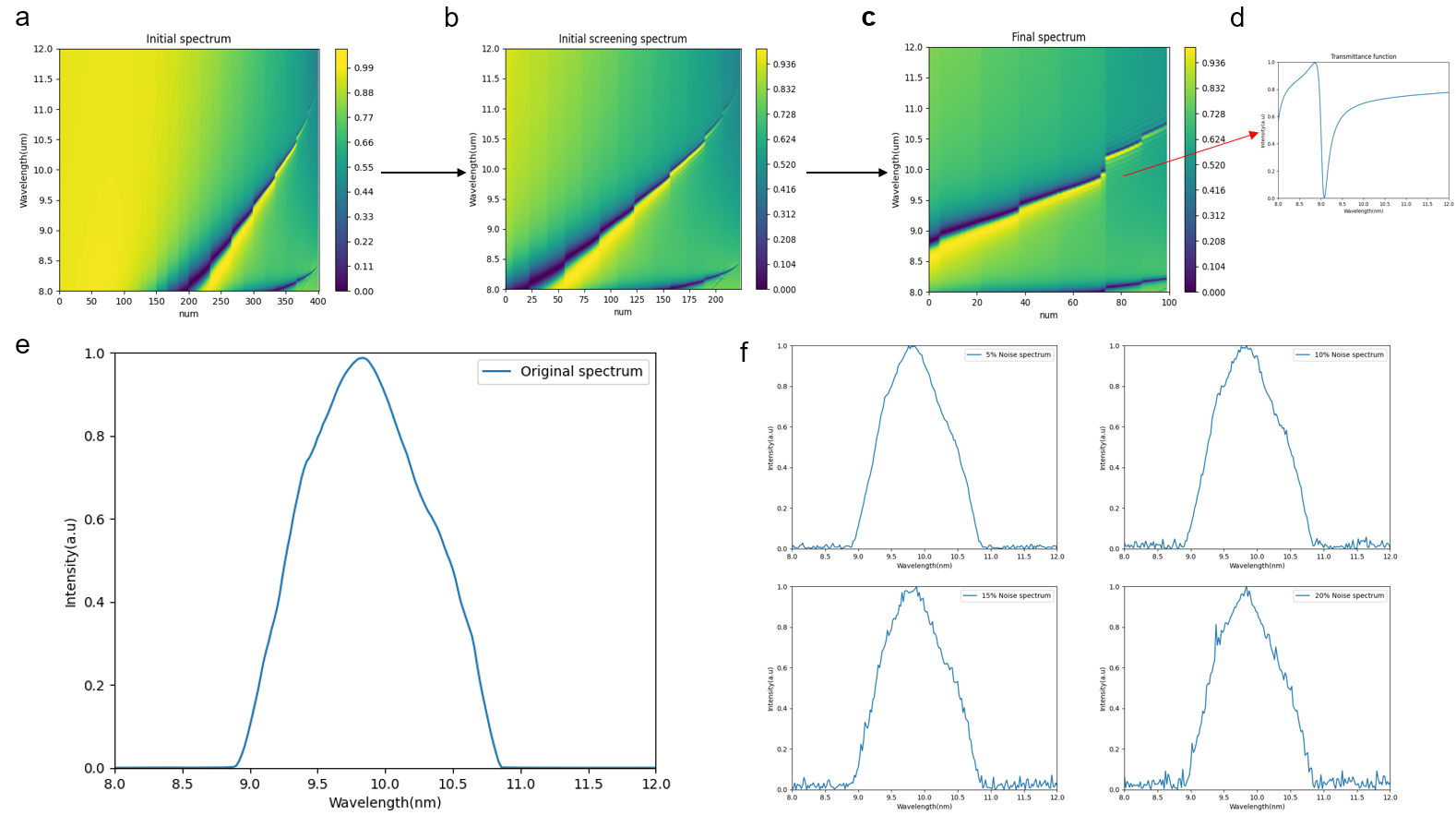


Fig3. 编码矩阵和模拟光谱数据设计。A是由FTDT仿真得到的滤波函数库，高1um，长宽一致100um-500um的硅材料；B是对滤波函数库经过初步的筛选后得到的结果；C是通过互相关函数得到的结果，D是C中其中的一条滤波函数。E是用高斯基函数模拟仿真的一条光滑的光谱曲线；F是分别添加了5%，10%，15%,20%的高斯噪音的光谱曲线。

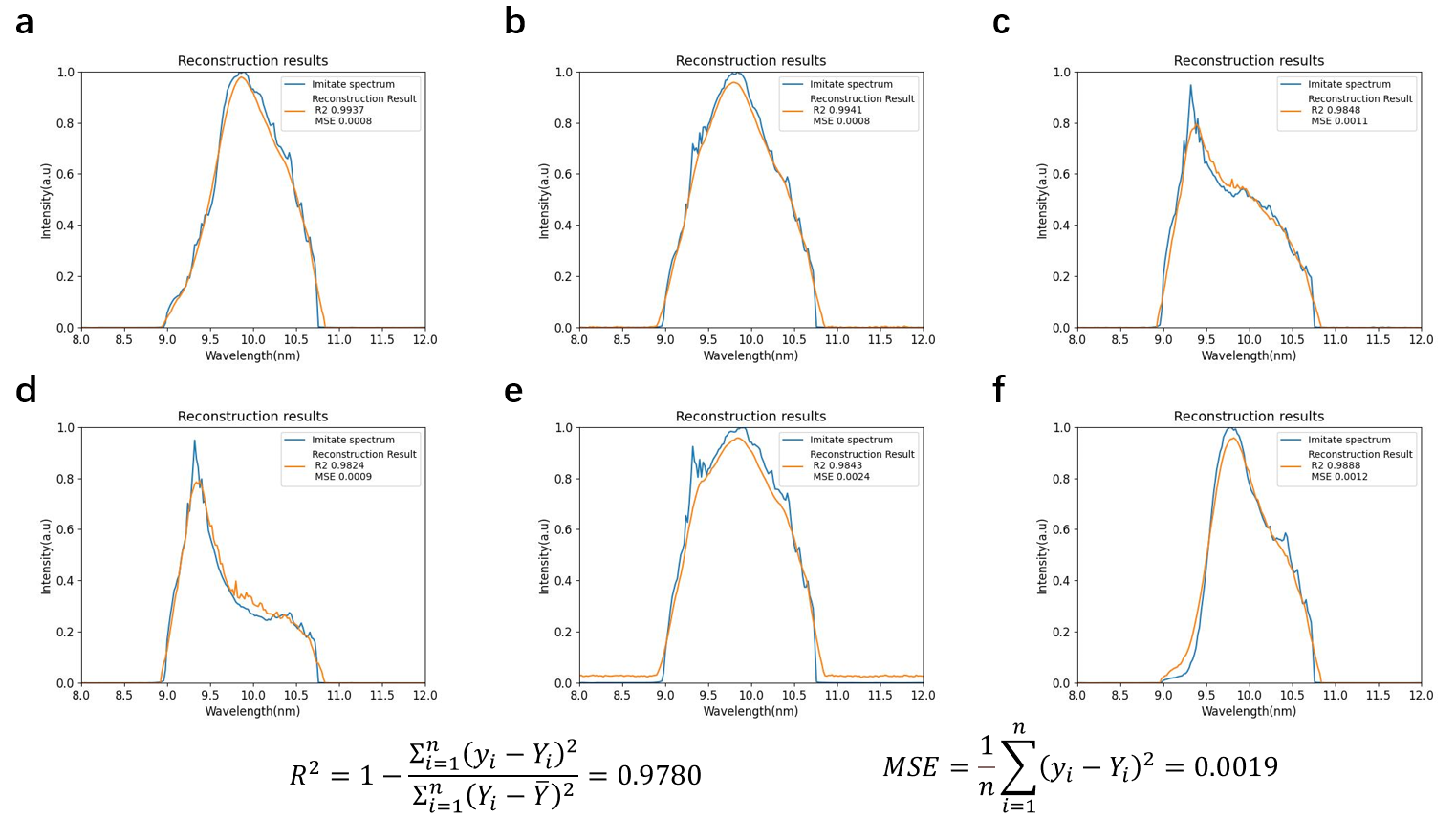


Fig4. 重建结果和测试集整体平均结果。

讨论：

计算光谱重建具有许多优点，例如成本低、体积小和易于操作等，因此在许多应用场合都具有广泛的应用前景。然而，在实际应用中，光谱重建的准确性常常受到影响。因此在本研究中，我们提出了一种新的方法来构建滤波函数编码矩阵并进行光谱重建，以提高计算光谱重建的准确性。过去的一些研究通常采用随机构建编码矩阵的方法，而基于互相关函数的筛选方法，可以确保构建的编码矩阵足够稀疏，从而更好地表达特征，方便后续光谱重建的数据处理。此外，还使用了基于自注意力机制的光谱重建方法，该方法可以加强特征之间的关联性，并对非线性数据进行拟合，从而进一步提高光谱重建的准确性，并使用加入残差机制防止训练时过拟合。实验结果表明，发现该算法相比于其他传统算法具有更高的准确性和更好的性能，能够提高光谱重建的准确性。总体来说在光谱重建中，互相关函数和自注意力机制是有效的方法，可以帮助优化编码矩阵的构建和特征提取，提高模型的准确性和性能。通过引入自注意力机制并结合残差网络，我们的方法可以提取光谱数据的重要特征，从而实现更准确的光谱重建。这一发现对于光谱重建技术的进一步研究和应用具有重要意义。

最后，需要指出的是，本研究的实验数据集仅限于仿真实验，因此需要进一步研究该算法在实际环境中的可行性和应用前景。为了验证算法的鲁棒性和实用性，未来研究可以将算法应用于更广泛的实际数据集，并且可以探索在不同的场景下算法的性能表现。此外，我们的研究也存在一些局限性，比如数据集的大小和多样性等方面可以进一步扩展和优化。因此，未来的研究可以进一步探索自注意力机制在光谱重建中的应用，包括如何优化网络结构、选择更有效的特征提取方法等方面，并尝试使用更多的数据集和方法来验证算法的有效性和泛化能力。通过不断的优化和改进，可以进一步提高光谱重建技术在实际应用中的效果，为相关领域的研究和应用提供更好的支持和帮助。

方法：

滤波函数设计和数据集的构建

本研究的数据构造分为两个部分。为了构建一个准确的光谱重建模型，需要采集合适的输入和输出数据。第一部分输入数据是由FDTD仿真出的一个完整的滤波函数库并通过期望设置为0.85为门限条件作为初筛，然后利用互相关函数最终筛选出100条滤波函数并排列成10X10的编码阵列。第二部分先用高级基函数模拟光谱曲线，并添加不同程度的高斯噪音，然后将其通过构建好的滤波函数编码矩阵进行转换成未编码的输入数据。该处一共仿真了20000条模拟光谱数据，并对此分别添加了5%，10%，15%，20%的高斯噪音，一共有着100000条模拟数据，然后对其以8:2的方式划分为训练集和验证集，其中验证集是以随机的方式在整体数据中进行抽取。最后拿现实测的1000条光谱数据转换后作为测试集，以测试模型的准确性。

神经网络模型的框架

用于光谱重建的神经网络的架构可以直接表示为“FC(100)->LR->Self-Attention(100)->FC(120)->LR->FC(150)->LR->FC(200)->LR”。每个数字表示对应图层中的单元数。LR表示ReLU函数。FC表示一个完全连接的层，输入单元数对应100个随机光谱滤波器，输出单元数200表示重建的光谱通道（8um-12um，0.02um步长）。且每层都有添加残差机制，在图2B中有表示。

训练和验证：

在训练时使用均方误差（MSE）作为损失函数，并采用批归一化技术来加速训练过程，Adam作为优化算法，并设置了适当的学习率。此外，为了避免过拟合问题，用L1和dropout进行正则化约束。

我们采用了实际由海洋光学光谱仪测量的光谱数据来验证模型的性能，以保证具有实际价值，在面对现实环境情况下能够准确的重建结果。

参考文献：

1. Z. Yang et al., Single-nanowire spectrometers. Science 365, 1017–1020 (2019). doi: 10.1126/science.aax8814; pmid: 31488686
2. L. P. Schuler, J. S. Milne, J. M. Dell, L. Faraone, MEMS-based microspectrometer technologies for NIR and MIR wavelengths. J. Phys. D 42, 133001 (2009). doi: 10.1088/ 0022-3727/42/13/133001
3. J. Malinen et al., Advances in miniature spectrometer and sensor development. Proc. SPIE 9101, 91010C (2014). doi: 10.1117/12.2053567
4. M. Ebermann et al., Tunable MEMS Fabry-Pérot filters for infrared microspectrometers: A review. Proc. SPIE 9760, 97600H (2016). doi: 10.1117/12.2209288
5. R. A. Crocombe, Portable Spectroscopy. Appl. Spectrosc. 72, 1701–1751 (2018). doi: .1177/0003702818809719; pmid: 30335465
6. R. F. Wolffenbuttel, MEMS-based optical mini- and microspectrometers for the visible and infrared spectral range. J. Micromech. Microeng. 15, S145–S152 (2005). doi: 10.1088/0960-1317/15/7/021.
7. Goodfellow, I., Bengio, Y. & Courville, A. Deep Learning 1 (MIT Press, 2016).
8. Jie Huang, Wengang Zhou, Qilin Zhang, Houqiang Li, and Weiping Li. Video-based sign language recognition without temporal segmentation. In Thirty-Second AAAI Conference on Artifificial Intelligence, 2018.
9. Gang Liu and Jiabao Guo. Bidirectional lstm with attention mechanism and convolutional layer for text classifification. Neurocomputing, 337:325–338, 2019.
10. Karim Ahmed, Nitish Shirish Keskar, and Richard Socher. Weighted transformer network for machine translation. arXiv preprint arXiv:1711.02132, 2017.
11. Matthias Sperber, Graham Neubig, Jan Niehues, and Alex Waibel. Attention-passing models for robust and data-effificient end-to-end speech translation. Transactions of the Association for Computational Linguistics, 7:313–325, 2019.
12. Kun Xu, Lingfei Wu, Zhiguo Wang, Yansong Feng, Michael Witbrock, and Vadim Sheinin. Graph2seq: Graph to sequence learning with attention-based neural networks. arXiv preprint arXiv:1804.00823, 2018.
13. Jianlong Fu, Heliang Zheng, and Tao Mei. Look closer to see better: Recurrent attention convolutional neural network for fifine-grained image recognition. In 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pages 4476–4484, Honolulu, HI, July 2017. IEEE.
14. Kai Han, Jianyuan Guo, Chao Zhang, and Mingjian Zhu. Attribute-aware attention model for fifine-grained representation learning. arXiv:1901.00392 [cs], January 2019. arXiv:1901.00392.
15. Dimitris Kastaniotis, Ioanna Ntinou, Dimitrios Tsourounis, George Economou, and Spiros Fotopoulos. Attention aware generative adversarial networks (ata-gans). In 2018 IEEE 13th Image, Video, and Multidimensional Signal Processing Workshop (IVMSP), pages 1–5. IEEE, 2018.
16. Jiahui Yu, Zhe Lin, Jimei Yang, Xiaohui Shen, Xin Lu, and Thomas S. Huang. Generative image inpainting with contextual attention. arXiv:1801.07892 [cs], January 2018. arXiv: 1801.07892.
17. Baraniuk, R. G. Compressive sensing. IEEE Signal Process. Mag. 24, 118–121 (2007).
18. Yang J , Cui K , Cai X , et al. Ultraspectral Imaging Based on Metasurfaces with Freeform Shaped Meta-Atoms[J]. Laser & Photonics Reviews, 2022(7):16.
19. Song, H. Y. et al. Deep-learned broadband encoding stochastic fifilters for computational spectroscopic instruments. Adv. Theory Simul. 4, 2000299 (2021).
20. Nie, S. J. et al. Deeply learned fifilter response functions for hyperspectral reconstruction. In Proc. IEEE/CVF Conference on Computer Vision and Pattern Recognition 18–23 June 2018 (IEEE, Salt Lake City, UT, USA, 2018).
21. Leibe B , Matas J , Sebe N , et al. [Lecture Notes in Computer Science] Computer Vision – ECCV 2016 Volume 9908 || Identity Mappings in Deep Residual Networks[J]. 2016, 10.1007/978-3-319-46493-0(Chapter 38):630-645.
22. He K , Zhang X , Ren S , et al. Deep Residual Learning for Image Recognition[C]// IEEE Conference on Computer Vision and Pattern Recognition. IEEE, 2016.
23. Kukaka J , Golkov V , Cremers D . Regularization for Deep Learning: A Taxonomy[J]. 2017.
24. Wen F , Chu L , Liu P , et al. A Survey on Nonconvex Regularization Based Sparse and Low-Rank Recovery in Signal Processing, Statistics, and Machine Learning[J]. 2018.
25. Engle R . Dynamic Conditional Correlation[J]. Journal of Business & Economic Statistics, 2002, 20(3):339-350.