SpectralFormer: Rethinking Hyperspectral Image Classification With Transformers

符雅雯

摘要

高光谱 (HS) 图像的特征是近似连续的光谱信息，能够通过捕捉细微的光谱差异来精细识别物质。卷积神经网络 (convolutional neural networks, cnn) 由于其出色的局部上下文建模能力，在 HS 图像分类中已被证明是一种强大的特征提取器。然而，cnn 由于其固有的网络骨干网的限制，不能很好地挖掘和表示光谱特征的序列属性。为了解决这个问题，我们从 Transformer 的顺序角度重新思考 HS 图像分类，并提出了一种新的骨干网 SpectralFormer。除了经典 Transformer 的逐波段表示，SpectralFormer 能够从 HS 图像的邻近波段学习光谱局部序列信息，从而产生分组光谱嵌入。更重要的是，为了减少在分层传播过程中丢失有价值信息的可能性，我们设计了一种跨层跳过连接，通过自适应学习融合跨层的“软”残差，将类似内存的组件从浅层传递到深层。值得注意的是，提出的 SpectralFormer 是一个高度灵活的骨干网，可以适用于像素级和补丁级输入。通过广泛的实验，我们评估了所提出的 SpectralFormer 在三个 HS 数据集上的分类性能，显示了其优于经典变压器的优势，并与最先进的骨干网相比取得了显著的改进。这项工作的原代码将在https://github.com/danfenghong/IEEE_TGRS_SpectralFormer 上提供。

关键词：卷积神经网络；深度学习；高光谱 (HS) 图像分类；局部上下文信息；遥感；序列数据；跳跃融合；变压器；

1 引言

超高光谱 (HS) 成像，在整个电磁波谱的每个像素处收集数百个 (窄) 波长波段，从而能够在细粒度水平上识别或检测物体，特别是对于那些在视觉线索中具有极其相似光谱特征的材料 (例如 RGB)[1]。这为各种高水平的地球观测 (EO) 任务提供了巨大的潜力，如精确的土地覆盖测绘、精准农业、目标/目标检测、城市规划、树种分类和矿产勘探。在 HS 图像分类系统中，一般的顺序过程包括图像恢复 (如去噪和缺失数据恢复)[2-5]，降维[6],[7]，光谱解混[8-12]，特征提取[13-16]。其中，特征提取是 HS 图像分类的关键步骤，越来越受到研究者的关注。在过去的十年中，针对 HS 图像分类[17]，人们提出了大量先进的手工和基于子空间学习的特征提取方法。这些方法在小样本分类问题中表现良好。但是，当训练规模逐渐增加，训练集变得更加复杂时，它们往往会遇到性能瓶颈。可能的原因是这些传统方法的数据拟合和表示能力有限。深度学习 (DL) 能够从大量的多元数据[18]中发现有内涵的、内在的和潜在有价值的知识，受到其巨大成功的启发，人们在网络中设计和添加先进的模块，以从遥感数据中提取更多的诊断特征。例如，Zhao 等[19]利用 HS 和光探测与测距 (LiDAR) 数据开发了一个联合分类框架，该框架已被证明在从多源 RS 数据中提取特征方面表现出色。Zhang 等[20]设计了一种非凡的 patch-topatch 卷积神经网络 (CNN)，其结果明显优于其他技术。

2 相关工作

近年来，许多公认的骨干网络，如自动编码器 (AEs)、cnn、循环神经网络 (RNNs)、生成对抗网络 (GANs)、胶囊网络 (CapsNet)、图卷积网络 (GCNs) 等，已广泛并成功应用于 HS 图像分类任务[21]。Chen 等人[22]堆叠了多个自编码器网络，从主成分分析 (PCA)[23]生成的降维 HS 图像中提取深度特征表示，并应用于 HS 图像分类。Chen 等[24]使用 cnn 代替堆叠 AEs，考虑 HS 图像的局部上下文信息，从语义上提取空间光谱特征，获得了更高的分类精度。Hang 等[25]利用 RNN 可以对序列性进行建模，有效地表示相邻光谱波段的关系，设计了一种用于 HS 图像分类的级联 RNN。在[26]中，对 gan 进行了改进，使其适用于 HS 图像分类任务，输入三个 PCA 分量和随机噪声。Paoletti 等[27]通过定义一种新的空间光谱胶囊单元，扩展了基于 cnn 的模型，得到了一个高性能的 HS 图像分类框架，同时降低了网络设计的复杂性。Hong 等[28]对 cnn 和 GCNs 在 HS 图像分类上进行了定性和定量的全面比较，提出了一种 minibatch GCN (miniGCNs)，为解决 GCNs 中的大图问题提供了一种可行的解决方案，用于最先进的 HS 图像分类。虽然这些骨干网络及其变体已经能够获得有希望的分类结果，但它们在表征光谱序列信息 (特别是在捕捉光谱维度上的细微光谱差异) 方面的能力仍然不足。图 1 给出了在 HS 图像分类任务中这些最先进的骨干网的概述说明。具体限制可以大致概括如下。1) cnn 作为主流的骨干架构，在从 HS 图像中提取空间结构信息和局部上下文信息方面表现出了强大的能力。然而，一方面，cnn 很难很好地捕获序列属性，特别是中长期依赖关系。这就不可避免地遇到了 HS 图像分类任务的性能瓶颈，特别是当待分类的类别种类繁多，光谱特征极其相似时。另一方面，cnn 过度关注空间内容信息，导致学习特征中的顺序信息发生了频谱畸变。这在很大程度上增加了挖掘诊断光谱属性的难度。2) 与 cnn 不同的是，rnn 是为序列数据设计的，它从 HS 图像中逐带有序地累积学习光谱特征。这种模式极度依赖于光谱波段的顺序，容易产生梯度消失，因此很难学习长期依赖关系[29]。这可能进一步导致难以捕捉时间序列中的光谱显著变化。更重要的是，在真实的 HS 图像场景中，通常有大量的 HS 样本 (或像素)，而 rnn 无法并行训练模型，限制了实际应用中的分类性能。3) 对于其他骨干网，如 GANs、CapsNet 和 GCNs，尽管它们在学习光谱表示 (例如，鲁棒性、等效性和样本之间的长程相关性) 方面都有各自的优势，但有一个共同点是，几乎所有的骨干网都无法有效地对序列信息建模。即光谱信息利用不足 (这是利用 HS 数据进行精细土地覆盖分类或制图的关键瓶颈)。针对上述局限性，我们利用目前最先进的变压器[30]，从序列数据的角度重新思考 HS 图像分类过程。与 cnn 和 rnn 完全不同的是，由于使用了自注意技术，变压器是 (目前) 最先进的骨干网之一，它的设计很好，可以更有效地处理和分析顺序 (或时间序列) 数据。这将为 HS 数据的处理和分析，如 HS 图像分类提供一个很好的适应。众所周知，变压器中的自注意块能够通过位置编码[31]的方式捕获全局顺序信息。然而，变压器也存在一些缺陷，阻碍了其性能的进一步提高。例如, 1）尽管变换器在解决频谱特征的长期依赖性问题上表现突出，但它们失去了捕捉局部语境或语义特征的能力（参考 CNN 或 RNN）。2）正如 [32][32]中提到的，跳过连接（SC）在变换器中起着关键作用。这可以通过以下方式来解释这可以通过使用” 残差” 来使梯度更好地传播，或者加强” 记忆” 以减少遗忘或丢失的关键信息。然而。不幸的是，简单的加法 SC 操作只发生在每个 transformer 块内，削弱了不同层或块之间的连接。基于这些原因，我们的目标是开发一种新的基于变压器的网络架构，简称 SpectralFormer，实现高性能的 HS 图像分类任务。针对上述两个缺点，SpectralFormer 提供了点对点解决方案。更具体地说，SpectralFormer 能够在每个编

码位置从多个相邻波段而不是单个波段 (在原始变压器中) 学习局部光谱表示，例如，分组嵌入和带嵌入。此外，在 SpectralFormer 中设计了跨层 SC，通过自适应学习融合“软”残差，逐步将类似内存的组件从浅层传递到深层。本文的主要贡献可以概括为以下几点。1) 我们从顺序的角度重新审视 HS 图像分类问题，并提出了一种新的基于变压器的骨干网 SpectralFormer，以替代基于 CNN 或 rnn 的架构。据我们所知，这是第一次将变压器 (没有任何预处理操作，例如使用卷积和循环单元或其他转换技术进行特征提取) 纯粹应用于 HS 图像分类任务。2) 我们在 SpectralFormer 中设计了两个简单但有效的模块，即分组光谱嵌入 (GSE) 和跨层自适应融合 (CAF)，分别用于学习局部详细的光谱表示和从浅层到深层传递类似内存的组件。3) 我们定性和定量地评估了 SpectralFormer 在三个具有代表性的 HS 数据集上的分类性能，即印第安松树、帕维亚大学和休斯顿大学，并进行了广泛的消融研究。实验结果表明，与经典变压器 (OA 提高约 10%) 和其他最先进的骨干网 (OA 提高至少 2%) 相比，该方法具有显著优势。本文的其余部分组织如下。第二节首先回顾了经典变压器相关文献，然后详细介绍了所提出的 SpectralFormer 与两个精心设计的 HS 图像分类模块。在第三节中进行了广泛的消融研究和讨论实验。第四部分综合总结，并对未来可能的研究方向进行了简要展望。

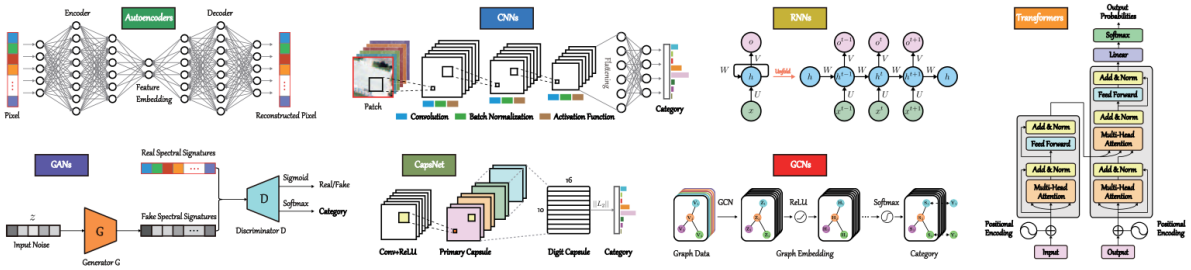

图 1: 目前公认的用于 HS 图像分类任务的骨干网络概述，如 AEs[22]，CNNs[24]，RNNs[25]，GANs[26]，CapsNet[27]，GCNs[28]，transformer[30]。

3 SPECTRAL FORMER

在这一节中，我们开始回顾经典变换器的一些前言。在此基础上，我们提出了带有两个精心设计的模块，即 GSE 和 CAF 的 SpectralFormer，使其更适用于 HS 图像分类任务。最后，我们还研究了所提出的 SpectralFormer 对输入图像 patch 的空间上下文信息建模的能力。

3.1 Transformer 简介

众所周知,[30]变压器在处理自然语言处理 (NLP) 中的序列到序列问题 (如机器翻译) 中占有重要地位。由于它们摒弃了 rnn 中的序列依赖特性，而引入了一种全新的自注意机制，从而能够对任意位置的单元进行全局信息 (长期依赖) 捕获，这极大地促进了时间序列数据处理模型的发展。不仅限于 NLP，图像处理和计算机视觉领域也开始探索变压器架构。最近，视觉转换器 (ViT)[33]似乎已经在各种视觉领域任务中达到或接近基于 cnn 的最先进效果，为视觉相关任务提供了新的见解、灵感和创意空间。

transfoermer 的成功在很大程度上取决于多注意的使用，其中多个自注意 (SA)[34]层被堆叠和集成。顾名思义，SA 机制更善于捕捉数据或特征的内部相关性，从而减少对外部信息的依赖。图 2(a) 为 transformer 中 SA 模块的工作过程。更具体地说，SA 层可以按照六个步骤执行. 此部分对本文将要复现的工作进行概述，图的插入如图 2 所示。

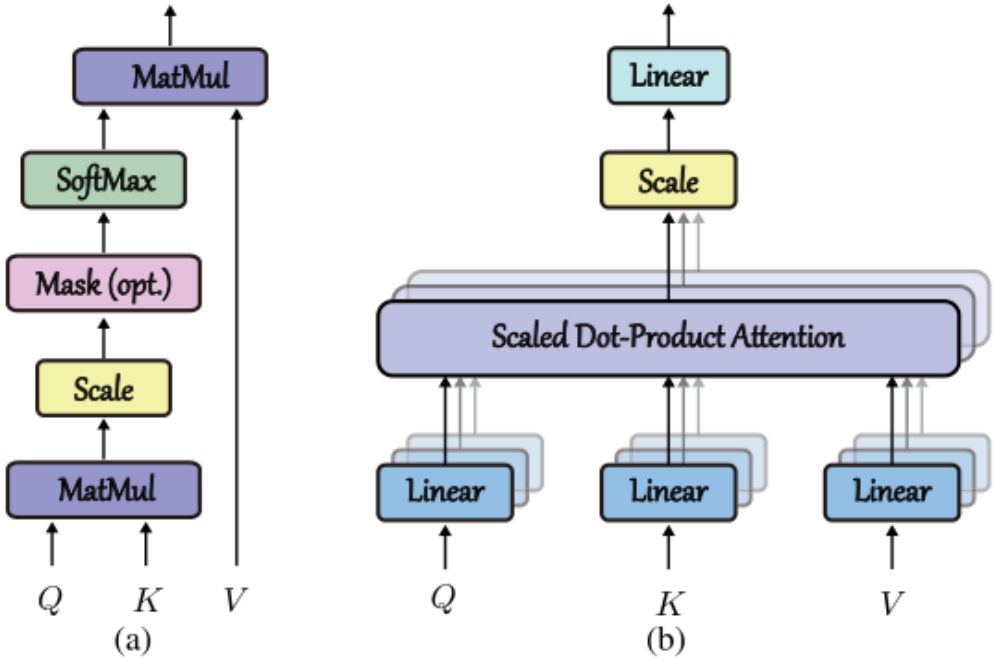

图 2: transformer 中注意机制的说明。(a) 自我注意模块。(b) 多注意。

3.2 SpectralFormer 概述

我们的目标是开发一种新颖通用的基于 vv 的基线网络 (即 SpectralFormer)，重点关注光谱特性，使其很好地适用于 HS 图像的高精度精细分类。为此，我们设计了两个关键模块，即 GSE 和 CAF，并将其集成到变压器框架中，分别提高了细微光谱差异的细节捕捉能力和增强层与层之间的信息传递性 (或连通性)(即减少随着层的逐渐加深而造成的信息损失)。此外，所提出的 SpectralFormer 不仅应用于按像素的 HS 图像分类，而且还可扩展到批量输入的空间-光谱分类，从而得到空间-光谱 SpectralFormer 版本。图 3 展示了所提出的 SpectralFormer 在 HS 图像分类任务中的概述，而表 I 详细描述了所提出的 SpectralFormer 中使用的符号的定义。

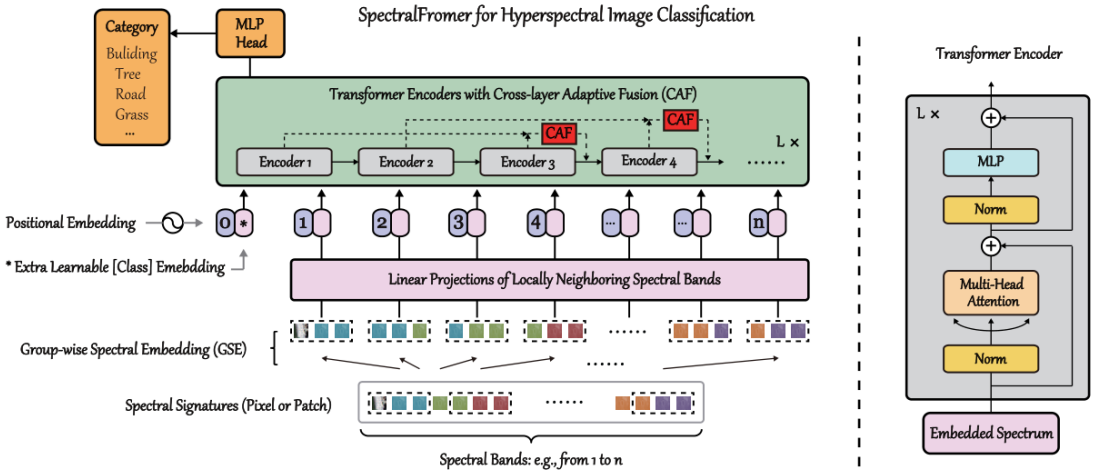

图 3: 用于 HS 图像分类任务的 SpectralFormer 网络概述。SpectralFormer 由两个设计良好的模块组成，即 GSE 和 CAF，使其 (基于变压器的骨干) 更好地适用于 HS 图像。

3.3 分组光谱嵌入

与经典 transformer 或 ViT 中的离散序列性 (例如，图块) 不同，HS 图像中的数百或数千个光谱通道以微妙的间隔 (例如，10 nm) 密集地从电磁频谱中采样，产生近似连续的光谱特征。不同位置的光谱信息反映了不同波长对应的不同吸收特性。这在很大程度上显示了当前材料的物理性质。捕捉这种

光谱特征的局部详细吸收 (或变化) 是准确和精细地对 HS 场景中的物质进行分类的关键因素。为此，我们建议学习分组光谱嵌入，即 GSE，而不是按频带输入和表示。

3.4 跨层自适应融合

在深度网络中，SC 机制已被证明是一种有效的策略，它可以增强层间的信息交换，减少网络学习过程中的信息丢失。近年来，SC 的应用在图像识别和分割方面取得了巨大的成功，例如 ResNet[35]的短 SC 和 U-Net[36]的长 SC。但需要注意的是，短序列的信息“记忆”能力仍然有限，而长序列由于高低特征之间的差距较大，往往融合不足。这也是变压器存在的一个关键问题，对变压器的结构设计提出了新的挑战。为此，我们在 SpectralFormer 中设计了一个中等范围的 SC，自适应学习跨层特征融合。

4 复现细节

高光谱图像的数据多，光谱波段冗余，如果将所有波段都输入进网络中进行训练的话，有些噪声或者不重要的数据可能会影响分类结果，而且造成网络训练时间较长的缺点。引入主成分分析（PCA）技术，在数据预处理阶段对高光谱图像数据进行降维处理，提取图像数据的主要特征分量，去除噪声数据，使数据更易被网络处理。

4.1 与已有开源代码对比

本文复现了作者所提出的基于 transformer 的 SpectralFormer 模型，并且引入主成分分析（PCA）技术，在数据预处理阶段对高光谱图像数据进行降维处理。本实验使用了 Pavia University、Indian Pines、Houston2013 等数据集进行实验，验证了作者所提出模型的有效性。本文参考了他人的相关代码，链接如下：https://github.com/danfenghong/IEEE_TGRS_SpectralFormer

4.2 实验环境搭建

操作系统	Windows10 专业版 21H2
CPU	11th Gen Intel(R) Core(TM) i7-11700 @ 2.50GHz
机带 RAM	40GB
存储器容量	1000GB

CPU	12 × Xeon Gold 6271
GPU	NVIDIA Tesla P100-16GB
显存	16GB
内存	48GB
硬盘	200GB

5 实验结果分析

在引入主成分分析后，将结果与以 SpectralFormer 网络下像素级和图块级输入以及 ViT 网络的结果进行对比，通过评估指标 OA、AA 和 Kappa 系数可以看出，加入主成分分析的分类精度明显提升。在将高光谱图像数据输入进网络前，对数据进行 PCA 降维处理不但可以减少网络的训练时间，还能增强网络的分类效果。且在四个高光谱遥感数据集上都证明我们的改进是有效的。以下是其中两个数据集的实验结果示意图。

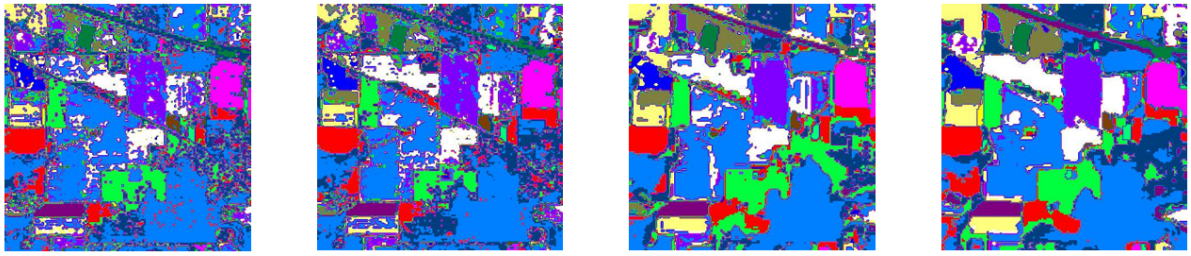

图 4: Indian Pines 实验结果示意

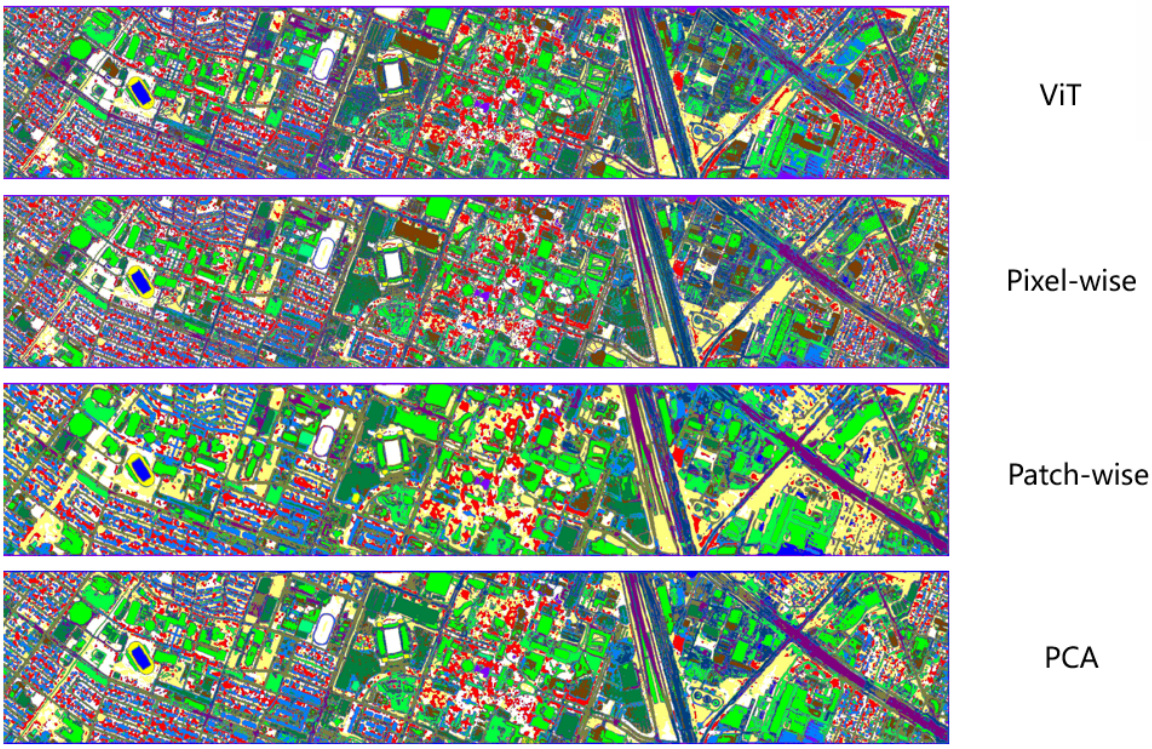

图 5: Houston2013 实验结果示意

6 总结与展望

HS 图像通常被收集 (或表示) 为一个具有空间光谱信息的数据立方体，一般可以将其视为沿光谱维度的数据序列。与 cnn 主要关注上下文信息建模不同，transformer 已被证明是一种强大的体系结构，可以在全局范围内描述顺序属性。然而，经典的基于 transformer 的视觉网络，如 ViT，在处理高光谱图像类数据时，不可避免地会出现性能下降的问题。这可以很好地解释为，ViT 未能对局部详细的光谱差异进行建模，并有效地传递“记忆”样的成分 (从浅层到深层)。为此，在本文中，我们提出了一种新的基于 transformer 的骨干网，称为 SpectralFormer，它更专注于提取光谱信息。在不使用任何卷积或循环单元的情况下，所提出的 SpectralFormer 可以获得最先进的高光谱图像分类结果。在未来，我们将研究进一步改进基于变压器的架构的策略，利用更先进的技术，如注意力，自监督学习，使其更适用于 HS 图像分类任务，并尝试建立一个轻量级的基于变压器的网络，以降低网络的复杂性，同时保持其性能。此外，我们还希望在所提出的框架中嵌入更多光谱波段的物理特征和 HS 图像的先验知识，从而产生更多可解释的深度模型。此外，CAF 模块中跳过和连接编码器的数量可能是提高 SpectralFormer 分类性能的一个重要因素，在未来的工作中应该更多地关注这一点。

参考文献

[1] HONG D, HE W, YOKOYA N, et al. Interpretable Hyperspectral Artificial Intelligence: When nonconvex modeling meets hyperspectral remote sensing[Z]. 2021.

[2] WANG Y, PENG J, ZHAO Q, et al. Hyperspectral Image Restoration via Total Variation Regularized Low-rank Tensor Decomposition[Z]. 2017.

[3] CAO W, WANG K, HAN G, et al. A Robust PCA Approach With Noise Structure Learning and Spatial–Spectral Low-Rank Modeling for Hyperspectral Image Restoration[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2018.

[4] WANG M, WANG Q, CHANUSSOT J, et al. L-l Hybrid Total Variation Regularization and Its Applications on Hyperspectral Image Mixed Noise Removal and Compressed Sensing[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, PP(99): 1-16.

[5] PENG J, SUN W, LI H C, et al. Low-Rank and Sparse Representation for Hyperspectral Image Processing: A review[J]. IEEE Geoscience and Remote Sensing Magazine, 2022, 10(1): 10-43. DOI: 10.1109/MGRS.2021.3075491.

[6] HONG D, YOKOYA N, CHANUSSOT J, et al. Joint and Progressive Subspace Analysis (JPSA) with Spatial-Spectral Manifold Alignment for Semi-Supervised Hyperspectral Dimensionality Reduction[Z]. 2020.

[7] LUO F, GUO T, LIN Z, et al. Semisupervised Hypergraph Discriminant Learning for Dimensionality Reduction of Hyperspectral Image[J/OL]. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens., 2020, 13: 4242-4256. https://doi.org/10.1109/JSTARS.2020.3011431. DOI: 10.1109/JSTARS.2020.3011431.

[8] YAO J, MENG D, ZHAO Q, et al. Nonconvex-Sparsity and Nonlocal-Smoothness-Based Blind Hyperspectral Unmixing[J/OL]. IEEE Trans. Image Process., 2019, 28(6): 2991-3006. https://doi.org/10.1109/TIP.2019.2893068. DOI: 10.1109/TIP.2019.2893068.

[9] HONG D, YOKOYA N, CHANUSSOT J, et al. An Augmented Linear Mixing Model to Address Spectral Variability for Hyperspectral Unmixing[J]. IEEE Transactions on Image Processing, 2019, 28(4): 1923-1938. DOI: 10.1109/TIP.2018.2878958.

[10] YUAN Y, ZHANG Z, WANG Q. Improved Collaborative Non-Negative Matrix Factorization and Total Variation for Hyperspectral Unmixing[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2020, 13: 998-1010. DOI: 10.1109/JSTARS.2020.2977399.

[11] GAO L, HAN Z, HONG D, et al. CyCU-Net: Cycle-Consistency Unmixing Network by Learning Cascaded Autoencoders[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, PP(99): 1-14.

[12] HONG D, GAO L, YAO J, et al. Endmember-Guided Unmixing Network (EGU-Net): A General Deep Learning Framework for Self-Supervised Hyperspectral Unmixing[J]. ArXiv e-prints, 2021.

[13] HONG D, YOKOYA N, CHANUSSOT J, et al. Learning to propagate labels on graphs: An iterative multitask regression framework for semi-supervised hyperspectral dimensionality reduction[J]. Isprs Journal of Photogrammetry and Remote Sensing, 2019, 158: 35-49.

[14] PENG J, SUN W, DU Q. Self-Paced Joint Sparse Representation for the Classification of Hyperspectral Images[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(2): 1183-1194. DOI: 10.1109/TGRS.2018.2865102.

[15] HONG D, WU X, GHAMISI P, et al. Invariant Attribute Profiles: A Spatial-Frequency Joint Feature Extractor for Hyperspectral Image Classification[Z]. 2019.

[16] LI Q, ZHENG B, TU B, et al. Ensemble EMD-Based Spectral-Spatial Feature Extraction for Hyperspectral Image Classification[J/OL]. IEEE J. Sel. Top. Appl. Earth Obs. Remote. Sens., 2020, 13: 5134-5148. https://doi.org/10.1109/JSTARS.2020.3018710. DOI: 10.1109/JSTARS.2020.3018710.

[17] RASTI B, HONG D, HANG R, et al. Feature Extraction for Hyperspectral Imagery: The Evolution From Shallow to Deep: Overview and Toolbox[J]. IEEE Geoscience and Remote Sensing Magazine, 2020, 8(4): 60-88. DOI: 10.1109/MGRS.2020.2979764.

[18] LECUN Y, BENGIO Y, HINTON G. Deep learning[J]. Nature, 2015, 521(7553): 436-444.

[19] ZHAO X, TAO R, LI W, et al. Joint classification of hyperspectral and LiDAR data using hierarchical random walk and deep CNN architecture[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(10): 7355-7370.

[20] ZHANG M, LI W, DU Q, et al. Feature extraction for classification of hyperspectral and LiDAR data using patch-to-patch CNN[J]. IEEE transactions on cybernetics, 2018, 50(1): 100-111.

[21] LI S, SONG W, FANG L, et al. Deep learning for hyperspectral image classification: An overview[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(9): 6690-6709.

[22] CHEN Y, LIN Z, ZHAO X, et al. Deep learning-based classification of hyperspectral data[J]. IEEE Journal of Selected topics in applied earth observations and remote sensing, 2014, 7(6): 2094-2107.

[23] ABDI H, WILLIAMS L J. Principal component analysis[J]. Wiley interdisciplinary reviews: computational statistics, 2010, 2(4): 433-459.

[24] CHEN Y, JIANG H, LI C, et al. Deep feature extraction and classification of hyperspectral images based on convolutional neural networks[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(10): 6232-6251.

[25] HANG R, LIU Q, HONG D, et al. Cascaded recurrent neural networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(8): 5384-5394.

[26] ZHU L, CHEN Y, GHAMISI P, et al. Generative adversarial networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(9): 5046-5063.

[27] PAOLETTI M E, HAUT J M, FERNANDEZ-BELTRAN R, et al. Capsule networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 57(4): 2145-2160.

[28] HONG D, GAO L, YAO J, et al. Graph convolutional networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 59(7): 5966-5978.

[29] BENGIO Y, SIMARD P, FRASCONI P. Learning long-term dependencies with gradient descent is difficult[J]. IEEE transactions on neural networks, 1994, 5(2): 157-166.

[30] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need[J]. Advances in neural information processing systems, 2017, 30.

[31] KE G, HE D, LIU T Y. Rethinking positional encoding in language pre-training[J]. ArXiv preprint arXiv:2006.15595, 2020.

[32] DONG Y, CORDONNIER J B, LOUKAS A. Attention is not all you need: Pure attention loses rank doubly exponentially with depth[C]//International Conference on Machine Learning. 2021: 2793-2803.

[33] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: Transformers for image recognition at scale[J]. ArXiv preprint arXiv:2010.11929, 2020.

[34] WANG X, GIRSHICK R, GUPTA A, et al. Non-local neural networks[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2018: 7794-7803.

[35] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition[C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 770-778.

[36] RONNEBERGER O, FISCHER P, BROX T. U-net: Convolutional networks for biomedical image segmentation[C]//Medical Image Computing and Computer-Assisted Intervention–MICCAI 2015: 18th International Conference, Munich, Germany, October 5-9, 2015, Proceedings, Part III 18. 2015: 234-241.