

PCTU-Net: 用于高光谱图像的多尺度池化-Transformer协同解混网络

陈建锋, 王继红*

贵州大学大数据与信息工程学院, 贵州 贵阳

收稿日期: 2023年10月9日; 录用日期: 2023年11月18日; 发布日期: 2023年11月24日

摘要

近年来, 基于深度学习的高光谱解混技术越来越受到关注, 并取得了重大进展。然而, 仅依靠Transformer方法不足以有效捕获全局和细粒度信息, 从而影响解混任务的准确性。为了充分利用高光谱图像中包含的信息, 本研究探索了一个通过池化操作来加深网络并提取图像细节特征进而与Transformer协同作用于高光谱图像解混的网络, 称为PCTU-Net。它端到端地充分学习了全局和局部信息, 以实现更有效地解混。该网络包括两个核心模块: 一个是多尺度池化模块, 该模块由最大池化操作、条纹池化操作和平均池化操作组成; 另一个是Transformer编码器, 它包括了嵌入层、自注意力模块、线性层以及多层感知器。本研究在三个数据集(Samson、Apex和合成数据集)上广泛评估了PCTU-Net和其他六种高光谱解混方法。实验结果有力地表明, 所提出的方法在精度方面优于其他方法, 具有有效解决高光谱解混任务的潜力。

关键词

深度学习, 高光谱图像解混, 多尺度池化, Transformer

PCTU-Net: Multi-Scale Pooling and Transformer Collaborative Unmixing Network for Hyperspectral Images

Jianfeng Chen, Jihong Wang*

College of Big Data and Information Engineering, Guizhou University, Guiyang Guizhou

Received: Oct. 9th, 2023; accepted: Nov. 18th, 2023; published: Nov. 24th, 2023

*通讯作者。

文章引用: 陈建锋, 王继红. PCTU-Net: 用于高光谱图像的多尺度池化-Transformer 协同解混网络[J]. 建模与仿真, 2023, 12(6): 5572-5584. DOI: 10.12677/mos.2023.126506

Abstract

In recent years, deep learning-based hyperspectral unmixing techniques have gained increasing attention and made significant strides. However, relying solely on the Transformer approach falls short in effectively capturing both global and fine-grained information, thereby impacting the accuracy of unmixing tasks. To fully harness the information embedded in hyperspectral images, this study explores a network called PCTU-Net, which deepens the network and extracts image detail features through pooling operations, synergizing with the Transformer for hyperspectral image unmixing. PCTU-Net end-to-end learns both global and local information to achieve more effective unmixing. The network comprises two core modules: a multi-scale pooling module, consisting of max-pooling, stripe pooling, and average pooling operations, and a Transformer encoder, which includes embedding layers, self-attention modules, linear layers, and multi-layer perceptron. This research extensively evaluates PCTU-Net and six other hyperspectral unmixing methods on three datasets (Samson, Apex, and Synthetic datasets). The experimental results robustly demonstrate that the proposed approach outperforms others in terms of accuracy, showcasing its potential for effectively addressing hyperspectral unmixing tasks.

Keywords

Deep Learning, Hyperspectral Image Unmixing, Multi-Scale Pooling, Transformer

Copyright © 2023 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

高光谱成像是一项备受关注的遥感技术。高光谱图像(HSI) [1]将反映物质辐射的光谱信息与地形的空间信息相结合。由于 HSI 中丰富的光谱和空间信息,它们在食品安全[2]、环境监测[3]、矿产勘探[4]等领域有着广泛的应用。然而,由于高光谱成像仪器[5]的空间分辨率限制,HSI 中的像素通常由混合光谱[6]、[7]组成,代表各种材料的组合,称为混合像素。在实际应用中,混合像素的丰富性会显著影响基于像素的材料分类和面积测量方法的准确性,使高光谱图像的开发和应用更具挑战性。为了解决这个问题,通常有两种方法需要考虑:第一种方法涉及提高光谱仪的空间分辨率,这不可避免地会增加人力和财力成本。第二种方法,高光谱解混[8],通常用于降低成本。光谱解混的主要目标是仅基于观测到的高光谱图像提取/估计每个像素中的端元及其丰度分数[9]。

在众多的高光谱混合方法中,线性光谱混合模型(LSMM) [10]以其简单、高效和能够很好地描述真实的光谱混合过程而脱颖而出。在 LSMM 框架的基础上,研究人员引入了几种有效的解混算法,包括一些最具代表性的算法,如几何、统计或稀疏方法。在几何方法领域,顶点分量分析(VCA) [11]和完全约束最小二乘解混合(FCLSU) [12]是最常用的技术。在稀疏解混方法领域,如最小绝对收缩和选择算子(LASSO) [13]和基于增强拉格朗日量的稀疏解混(SUnSAL) [14]由于其卓越的性能,引起了人们的极大关注。基于统计的方法中,由于在学习基于分量的表示方面的独特优势,非负矩阵因子分解(NMF) [15]和 L1/2-NMF [16]是统计方法类别中同时估计端元和丰度的两种最常用的算法。

近年来,随着深度学习的兴起,基于卷积神经网络(CNN)的各种方法在高光谱分解领域得到了快速发展。其中, Danfeng Hong 等人提出的 EGU-Net [17]针对端元引导的解混,引入了使用端元来引导解混

网络的概念。它代表了在解混研究中首次利用此类技术,为解混研究的未来发展提供了新的见解。Behnood Rasti 等人提出了最小单纯形卷积网络(MiSiC-Net) [18], 该网络结合了相邻像素之间的空间相关性和线性单纯形的几何特性。

近年来, Transformer [19] [20]的开发在 NLP 中取得了巨大成功, 而 Vision Transformer [21]将这种架构扩展到了计算机视觉领域。它展示了其独立于卷积运算的独特能力, 并在各类图像处理任务中取得了非凡的成果。例如最近 Preetam Ghosh 及其同事引入了 Deep Trans [22]网络, 该网络首次尝试将 Transformer 架构应用于高光谱解混任务, 并取得了显著的研究成果。这令人信服地证明了 Transformer 在图像解混任务中的可行性。

2. 方法

本文提出了一种基于多尺度池化和 Transformer 的双协同高光谱解混编解码器网络。所提出的网络的编码器部分分为两个模块: 多尺度池化模块和 Transformer 编码器模块, 如图 1 所示。

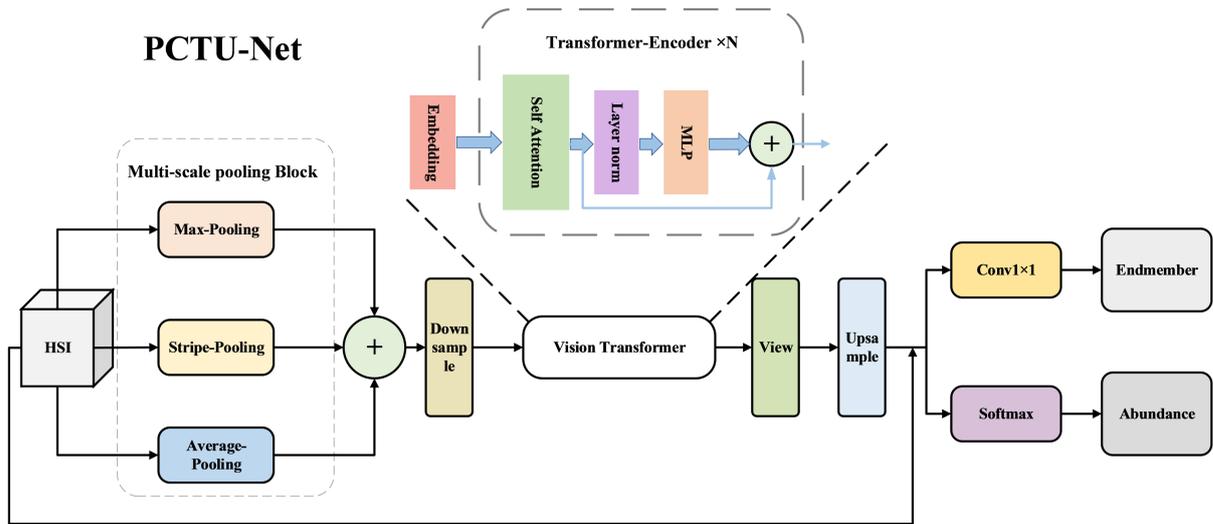


Figure 1. PCTU-Net structure diagram

图 1. PCTU-Net 结构图

2.1. 相关问题

2.1.1. 公式和有关符号表示

本文中涉及的符号由以下表示: 设高光谱图像 $I \in \mathbb{R}^{B \times H \times W}$, 其中空间维度用 $H \times W$ 表示, 光谱通道用 B 表示。HSI 可以被重塑以产生矩阵 $Y = [y_1, y_2, y_3, \dots, y_n] \in \mathbb{R}^{B \times n}$, 表示 B 波段 $n = H \cdot W$ 像素的输入 HSI。 y_i 代表第 i 个观测光谱。端元矩阵将由 $E = [e_1, e_2, e_3, \dots, e_R] \in \mathbb{R}^{B \times R}$ 表示, e_i 为第 i 个端元向量, R 表示 HSI 中存在的端元的数量。相应的丰度立方体由 $M \in \mathbb{R}^{R \times H \times W}$ 表示。丰度立方体可以被重塑以产生矩阵 $A = [a_1, a_2, a_3, \dots, a_n] \in \mathbb{R}^{R \times n}$, a_i 表示第 i 个观测像素对应的分数丰度。LMM 已经被广泛用于解混, 在 LMM 中观测到的反射率公为:

$$Y = EA + N \tag{1}$$

其中 $N \in \mathbb{R}^{B \times n}$ 是 Y 中存在的加性噪声。另外, 在解混任务中通常应满足三个物理约束条件: 首先端元矩阵应该是非负的 $E \geq 0$, 其次需要满足丰度非负约束和丰度合一约束, 即 $ANC(A \geq 0)$, $ASC(1_R^T A = 1_n^T)$, 其中 1_n 表示 1 的 n 维列向量。

2.1.2. 编码器

由于 AE 具有强大的表示和重构能力, 非常有能力从给定的输入中重建和提取信息, 因此已成为无监督深度学习模型在高光谱解混领域的代表。一般来说, AE 由编码器与解码器构成。本文中编码器部分输入像素 $y_i \in \mathbb{R}^n$ 将其降维为 $v_i \in \mathbb{R}^R$, 用公式可以表示为:

$$v_i = F_D(y_i) = F(W^{(d)T} y_i + b^{(d)}) \quad (2)$$

其中 $F(\cdot)$ 为非线性激活函数, 如 sigmoid 和 ReLU 等, $W^{(d)}$ 和 $b^{(d)}$ 表示第 d 个编码器部分的权重和偏置。

2.1.3. 解码器

解码器的主要任务是基于 LMM 将提取的隐藏特征转化为原始输入像素, 重构像素 $\hat{y}_i \in \mathbb{R}^n$ 表示为:

$$\hat{y}_i = f_E(v_i) = W^{(e)T} v_i \quad (3)$$

其中, $W^{(e)}$ 为解码器部分的权重矩阵, 所提取的端元矩阵的结果 \hat{e}_i 和估计的丰度向量 \hat{a}_i 分别对应 $W^{(e)}$ 和 v_i

2.1.4. 损失函数

AE 解混网络的目标函数是通过在均方误差(MSE)和光谱角距离(SAD)等不同测量形式下 y_i 和 \hat{y}_i 之间的重构误差来实现的

$$L_{RE}(y_i, \hat{y}_i) = \frac{1}{H \cdot W} \sum_{i=1}^H \sum_{j=1}^W (\hat{y}_i - y_i)^2 \quad (4)$$

$$L_{SAD}(y_i, \hat{y}_i) = \frac{1}{R} \sum_{i=1}^R \cos^{-1} \left(\frac{\hat{y}_i^T y_i}{\|\hat{y}_i\|_2 \|y_i\|_2} \right) \quad (5)$$

RE 损失由均方误差(MSE)目标函数计算, 并帮助编码器部分只学习输入 HSI 的基本特征, 而丢弃非必要的细节。SAD 损失是一个尺度不变目标函数。MSE 区分端元, 基于它们的绝对大小, 这在 HSI 解混的情况下是不可取的。加入 SAD 损失有助于抵消 MSE 目标函数的这一缺点, 使整个模型收敛得更快。总损失按这两个损失的加权和计算:

$$L = \beta L_{RE} + \gamma L_{SAD} \quad (6)$$

正则化参数为 β 和 γ , 我们会在后续实验部分介绍这两个参数的影响。

2.2. 多尺度池化模块

本文设计了多尺度池化模块对于输入图像 $I \in \mathbb{R}^{L \times H \times W}$ 进行三个维度的池化操作以期待更细节的学习图像特征, 同时加深网络使 Transformer 编码器在长程网络上可以更好地处理图像信息以此提高解混性能, 其池化过程如图 2 所示。

在最大池化中, 输入的三维图像 I 经过最大池化操作后与条纹池化和平均池化进行融合, 可表示为:

$$y_m = \text{Maxpool}(I) \quad (7)$$

在条纹池化中, 数学上, 输入的三维图像 $I \in \mathbb{R}^{L \times H \times W}$ 经过最大池化与卷积操作后得到输出 y_i^h 和 y_j^w 。其中水平池化输出 y_i^h 可以定义为:

$$y_i^h = \frac{1}{W} \sum_{0 \leq j \leq W} I'_{i,j} \quad (8)$$

垂直池化输出 y_j^w 可以定义为:

$$y_j^w = \frac{1}{H} \sum_{0 \leq i \leq H} I'_{i,j} \quad (9)$$

式中, i, j 表示图像像素的位置, h, w 表示池化的空间范围。

Multi-scale-Pool Block

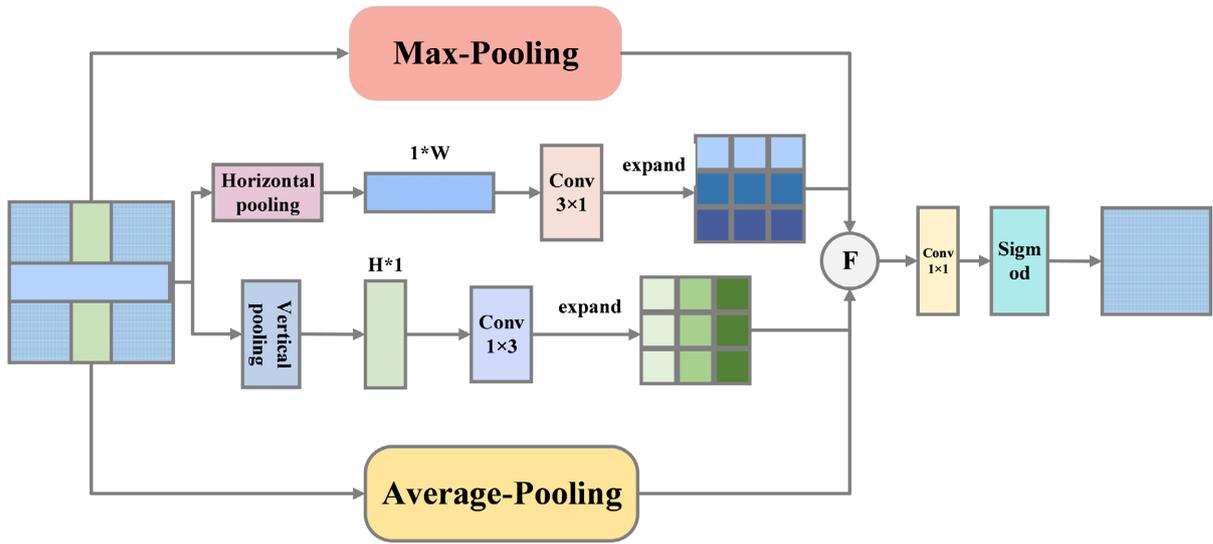


Figure 2. Multi scale pooling process
图 2. 多尺度池化过程

在平均池化中, 输入的三维图像 I 经过平均池化操作后与条纹池化和最大池化进行融合, 可表示为:

$$y_a = Averagepool(I) \quad (10)$$

然后我们把 y_i^h, y_j^w, y_m 和 y_a 进行融合, 如下所示:

$$Y = y_i^h + y_j^w + y_m + y_a \quad (11)$$

最终, 将 Y 经过一个 1×1 的卷积和 Sigmoid 函数得到输出 z 可以被计算为:

$$z = \sigma f(Y) \quad (12)$$

其中, σ 表示 sigmoid 函数, $f()$ 表示 1×1 卷积。

2.3. Transformer 模块

本文中 Transformer 由四个模块组成, 接下来将详细讲述各模块功能及步骤。

首先, 对于输入图像的嵌入(embedding), 将输入图像分成固定大小的图像块, 每个图像块被视为一个 “patch”。这些图像块首先被线性变换(通常是一个卷积层), 以产生初始的嵌入表示。这些表示构成了模型的输入序列。为了考虑到图像中各个位置的信息, 通常会添加位置嵌入向量。位置嵌入向量与每个输入图像块的嵌入相加, 以表示图像块在图像中的位置。

其次, 对于自注意力机制(self-attention), Vision Transformer 使用了多头自注意力机制(Multi-Head Self-Attention), 其中包括了多个注意力头。自注意力机制用于捕捉图像块之间的关系, 以便模型可以理解全局和局部特征之间的相互作用。

注意力权重计算的公式如下:

$$Attention(Q, K, V) = softmax\left(\frac{QK^T}{\sqrt{d_k}}\right)V \quad (13)$$

其中, Q 表示查询(Query)矩阵, K 表示键(Key)矩阵, V 表示值(Value)矩阵, 而 d_k 是注意力头的维度。

然后, 对于层归一化(Layer Normalization), 在每个子层(如多头自注意力和前馈神经网络)之后, 通常会应用层归一化来稳定训练。

最后, MLP 通常由两个全连接层组成, 每个全连接层后面跟着一个激活函数。这两个全连接层分别称为前馈神经网络(Feed-Forward Network)的一部分。MLP 用于在 Transformer 模型中对每个位置的特征进行映射和变换。具体来说, MLP 接收来自注意力子层的输出作为输入。

2.4. 解码器

将协同网络输出进行重塑并与输入特征图进行融合后, 我们将自动编码器的输出给到解码器模块进行解码, 其结构图如图 1 所示。其具体操作步骤如下: 首先, 本文解码器通过 4 层卷积核为 1×1 的卷积层将通道压缩至相应数据集端元的数目。其次, 通过一个核为 3×3 的卷积及 softmax 函数后得到估计丰度。最终, 将所得到的丰度经过一个核为 1×1 的卷积层后获得重构端元。

3. 仿真实验结果与分析

本文将所提出的网络的结果与六种不同的解混技术(即 FCLSU、NMF、Collab、EGU-Net、MiSiC-Net 和 Trans-Net)分别应用于两个真实数据集(即 Samson 和 Apex)和一个合成数据集进行了比较。

3.1. 实验结果及与其他网络的比较

本文评估了所提出的模型的性能, 并在上述三个数据集上将其与六种不同的方法进行了比较。其中包括三种传统的解混方法: 利用 VCA 进行端元提取的 FCLSU, $L_{1/2}$ -NMF 采用丰度稀疏性进行解混, 以及考虑空间光谱信息的 collab 进行联合解混。此外, 还考虑了三种最先进的深度神经网络方法: EGU-Net, 一种端元引导的解混合网络; 频谱空间协作网络 MiSiC-Net; 以及 Trans-Net, 这是用于解混合的变压器网络的第一个应用。下列图 3~8 中, 1 表示 FCLSU 方法, 2 表示 NMF 方法, 3 表示 Collab 方法, 4 表示 EGU-Net, 5 表示 MiSiC-Net, 6 表示 Trans-Net, 7 表示本文拟议方法 PTCU-Net。

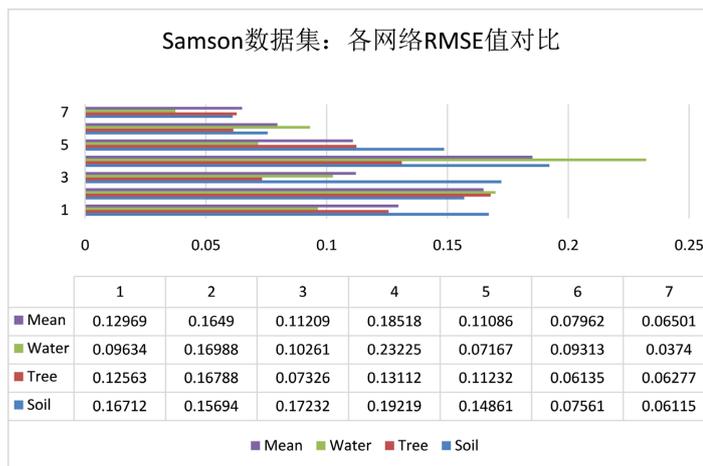


Figure 3. The comparison chart of RMSE values in the Samson dataset
图 3. Samson 数据集 RMSE 值结果对比图

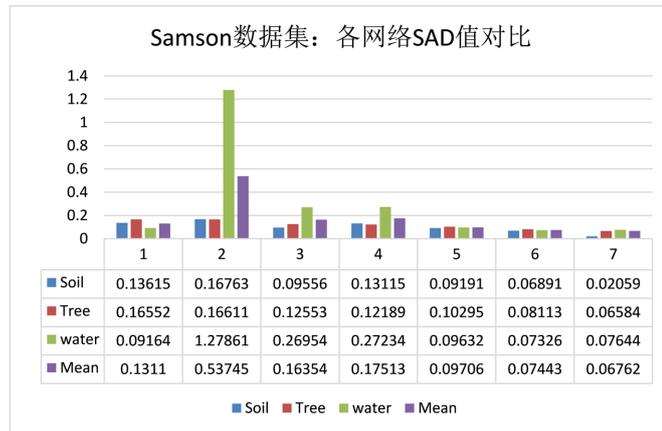


Figure 4. The comparison chart of SAD values in the Samson dataset
图 4. Samson 数据集 SAD 值结果对比图

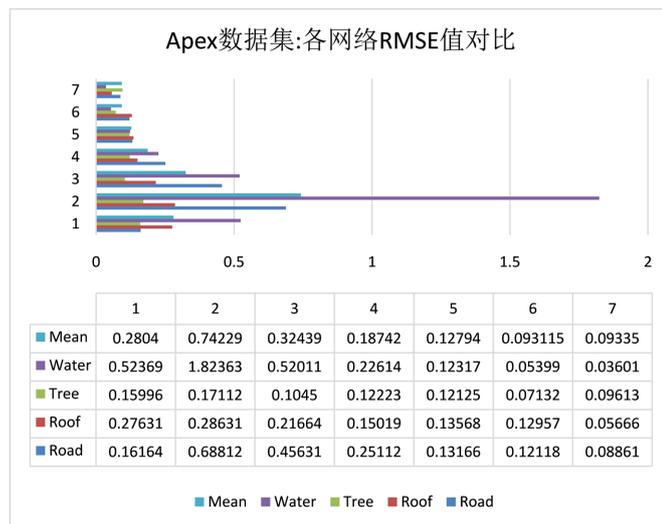


Figure 5. The comparison chart of RMSE values in the Apex dataset
图 5. Apex 数据集 RMSE 值结果对比图

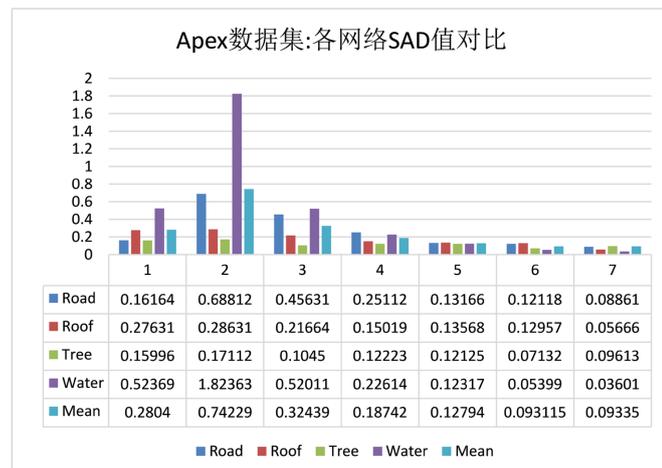


Figure 6. The comparison chart of SAD values in the Apex dataset
图 6. Apex 数据集 SAD 值结果对比图

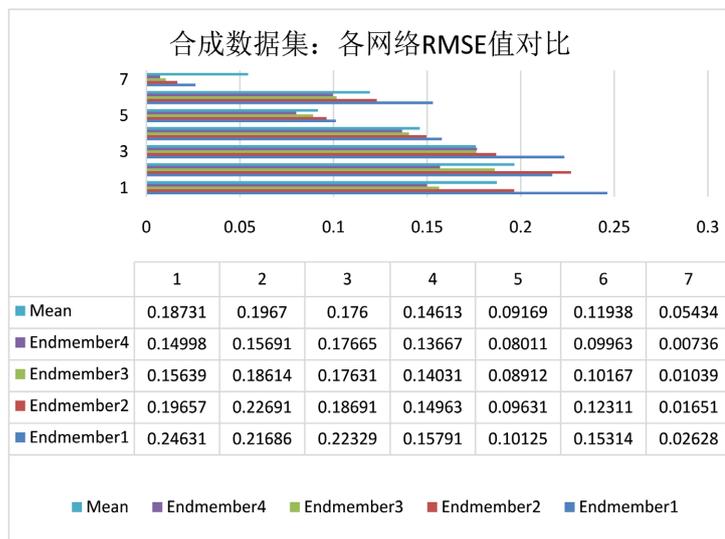


Figure 7. The comparison chart of RMSE values in the synthesized dataset
图 7. 合成数据集 RMSE 值结果对比图

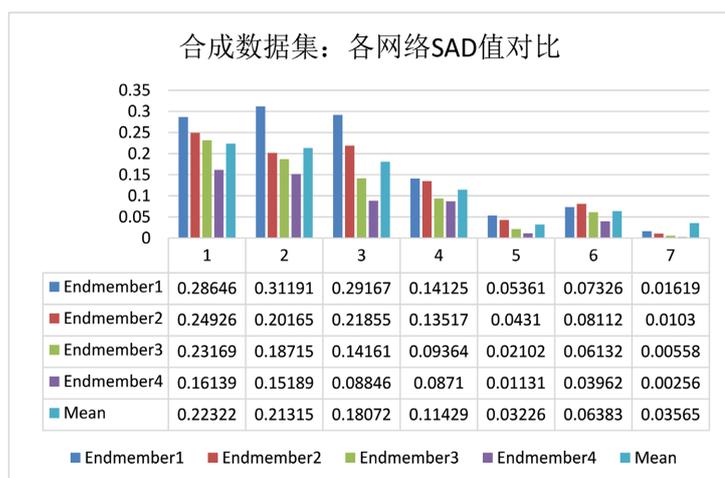


Figure 8. The comparison chart of SAD values in the synthesized dataset
图 8. 合成数据集 SAD 值结果对比图

Samson 数据集：Samson 数据集中的定量结果图 3 和图 4。结果表明，在大多数情况下，所提出的模型在丰度和端元估计方面优于其他技术。其中，我们网络的平均均方根误差为比次优方法高 30.90%；其平均 SAD 比次优方法高 22.58%。实验结果表明了所提出的网络在 Samson 数据集上的竞争力，并证明了多尺度池化和 Transformer 之间的联合应用的网络在解混任务中的可行性和优越性。

Apex 数据集：Apex 数据集中的定量结果如图 5 和图 6 所示。从表中可以看出，与大多数其他方法相比，Apex 数据中的终点“路”和“水”构成了相当大的挑战。然而，所提出的方法极大地提高了这两个端元的估计，这主要是由于所提出的网络对光谱空间信息的协同学习策略，该策略考虑了端元的几何信息，并充分利用了有效的谱带。该方法的平均 RMSE 值比次优方法高 34.26%，平均 SAD 比次优法高 71.10%。此外，基于 SAD 值，所提出的方法为所有端元提供了最佳端元估计。

合成数据集：为了评估所提出的方法对噪声的鲁棒性，我们在模拟数据集中添加了具有不同噪声功率的高斯白噪声，并获得了信噪比分别为 20、30、40 和 50 dB 的数据。不同噪声条件下的端元和丰度估

计结果如图 7 和图 8 所示, 包括所有备选解混方法的结果。总体而言, 一方面, 随着信噪比的增加, 每个网络的丰度和端元的估计结果都有所改善。另一方面, 与替代网络相比, 本文提出的网络产生了优越的解混结果, 在所有信噪比的 RMSE 和 SAD 方面表现出显著优势。这主要归功于所提出的方法融合了多尺度池化和 Transformer 的优势, 对全局与局部信息进行了全面细致的学习。

3.2. 可视化分析

在对不同的数据集和比较网络进行实验研究之后, 我们决定可视化地分析每个数据集生成的丰度图和端元图。从丰度图和端元图可以看出, 与其他比较网络相比, 本文方法得到的丰度图和端元图在视觉上最接近地面真值图(GTs)。接下来, 将逐一讨论每个数据集上的结果。

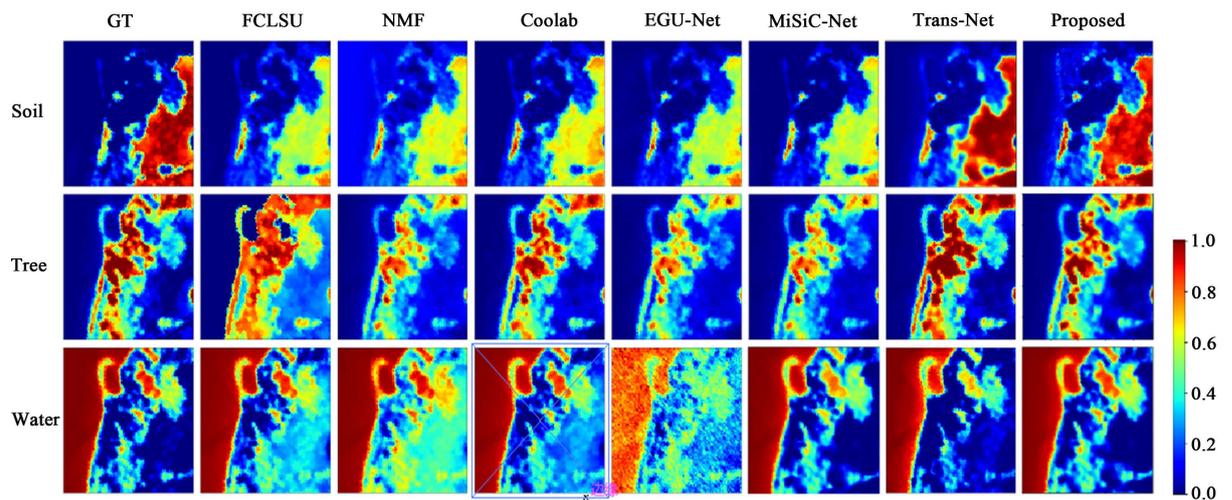


Figure 9. Visual comparison chart of Samson dataset obtained through different abundance unmixing techniques

图 9. 通过不同丰度分解技术获得的 Samson 数据集的可视化对比图

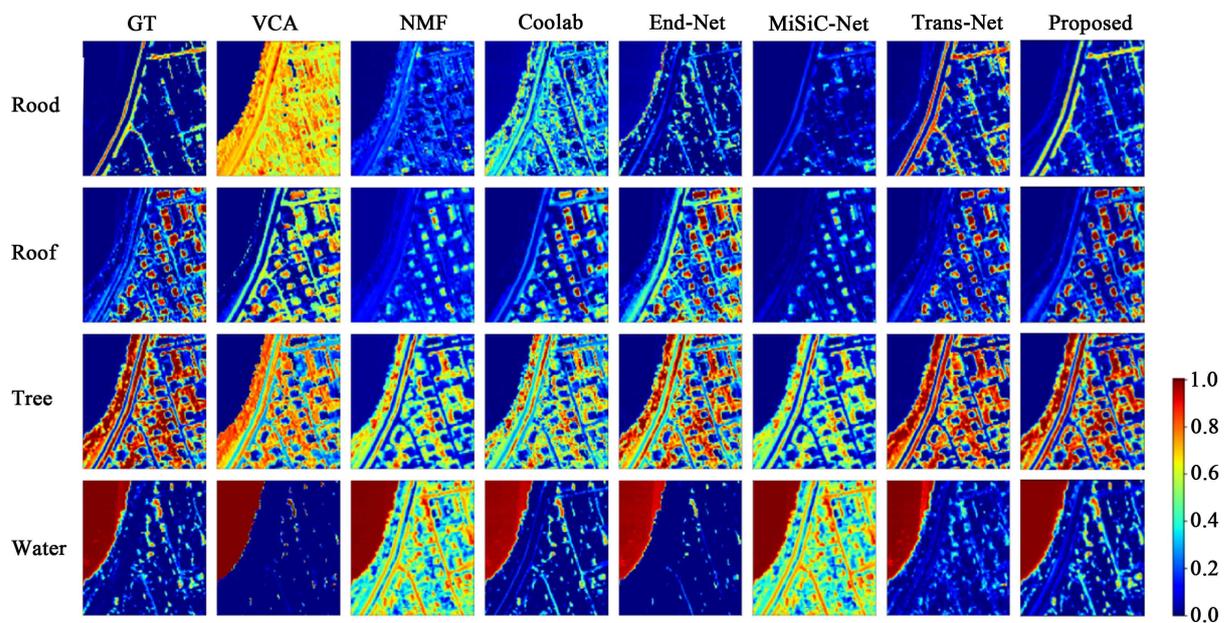


Figure 10. Visual comparison chart of Apex datasets obtained through different abundance unmixing techniques

图 10. 通过不同丰度分解技术获得的 Apex 数据集的可视化对比图

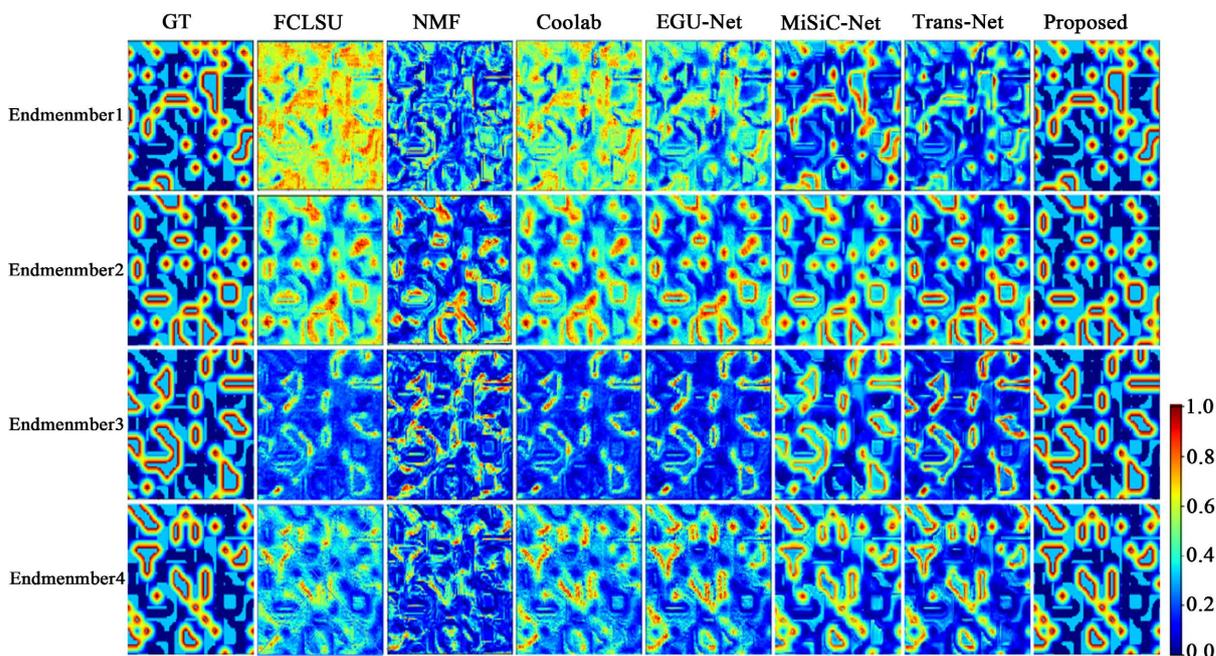


Figure 11. Visual comparison chart of synthesized datasets obtained through different abundance unmixing techniques

图 11. 通过不同丰度分解技术获得的合成数据集的可视化对比图

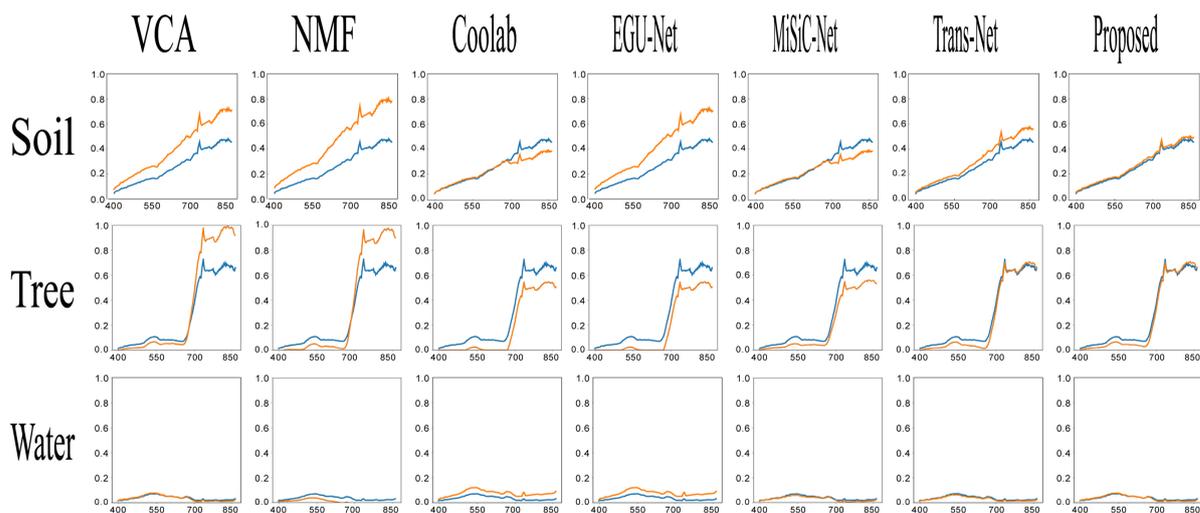


Figure 12. Visual comparison chart of Samson dataset obtained through different end member unmixing techniques

图 12. 通过不同端元分解技术获得的 Samson 数据集的可视化对比图

对于图 9 和图 12 所示的 Samson 数据集：传统的分解方法(如 FCLSU、NMF 和 Coolab)在丰度和端元图上显示的结果不充分。这是因为它们的性能受到终端元素之间差异的影响，这限制了它们在分解任务中的准确性，并导致这些模型的总体性能损失。EGU 网法总体效果较好，但对端元素“水”的处理效果较差。这可能是由于网络主要针对纯端元进行分解，而这类任务在混合像元分解中的性能并不突出。MiSiC 网络作为近年来应用于高光谱分解任务的代表性神经网络之一，在 Samson 上也显示出竞争性的分解结果。然而，由于 CNN 本身的限制，它不能有效地捕获全局信息，从而在最终重建的丰度图和端元图中留下了增强的空间。Trans-Net 作为第一个基于 Transformer 处理分解任务的网络，提高了分解任务的精度。然而，由于变压器主要关注全局信息，其解混性能通常不如本文拟议的方法。

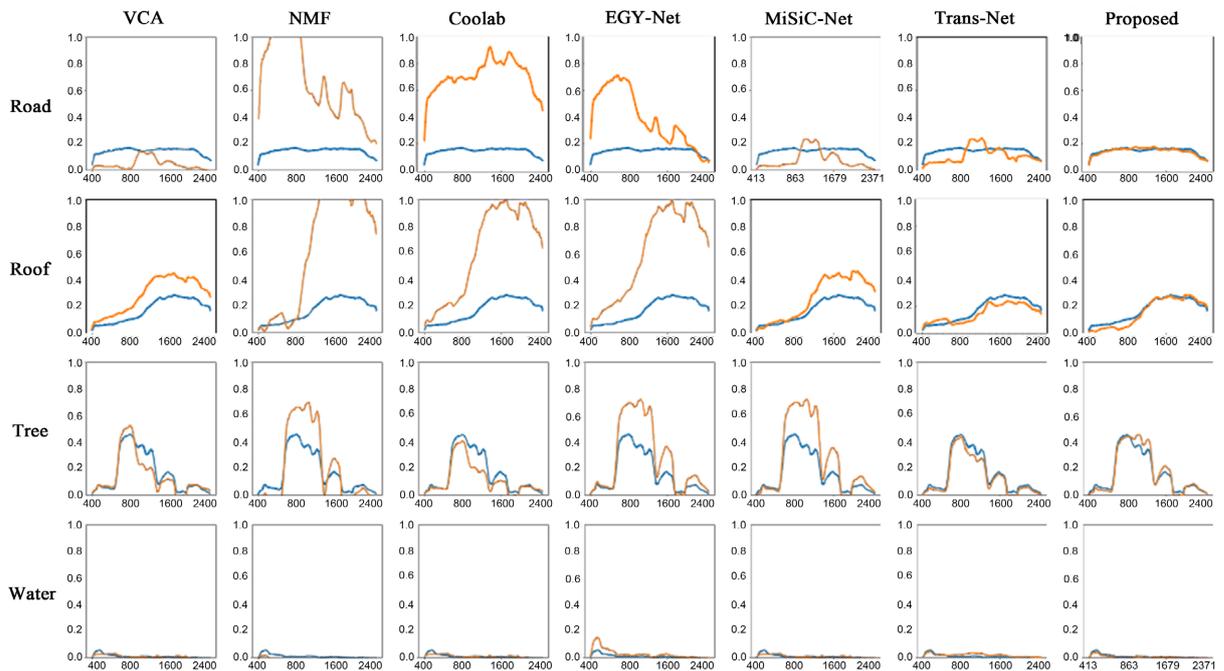


Figure 13. Visual comparison chart of Apex datasets obtained through different end member unmixing techniques
图 13. 通过不同端元分解技术获得的 Apex 数据集的可视化对比图

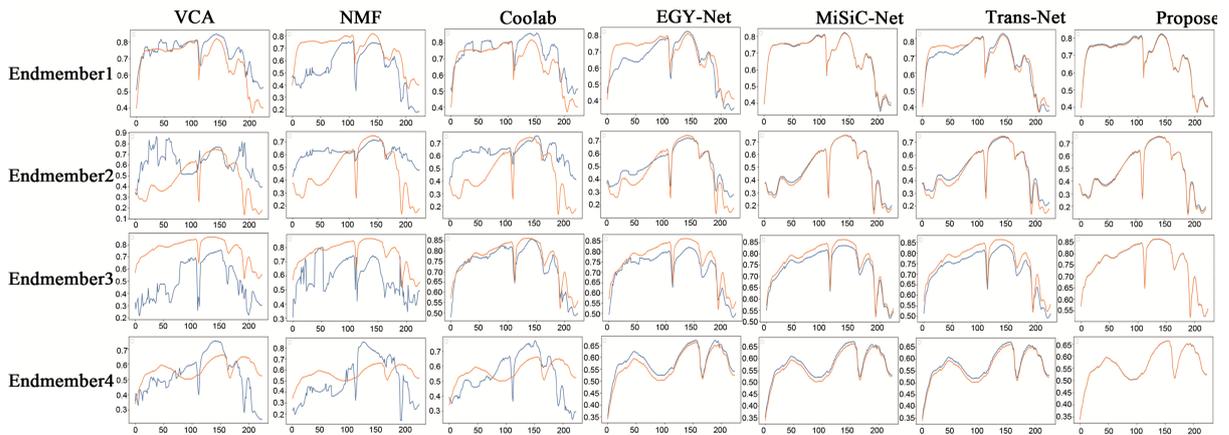


Figure 14. Visual comparison chart of synthesized datasets obtained through different end member unmixing techniques
图 14. 通过不同端元分解技术获得的合成数据集的可视化对比图

对于如图 10 和图 13 所示的 Apex 数据集：我们观察到，在“道路”端元上提出的方法最接近原始丰度图，相对而言，其他方法在估算端元和丰度方面的精度低于在屋顶和树端元上提出的方法，Trans-Net 和提出的方法明显优于以往的方法；在水端元上，该方法也更具竞争力。整体网络性能的显著增强主要源于本文中多尺度池化和 Transformer 架构的协作培训。这种协作方法使网络能够获得全局和复杂的细节，从而大大提高网络性能。

对于图 11 和图 14 所示的合成数据集：可以看出，近年来提出的深度学习方法与经典分解方法相比取得了更好的结果。MiSiC 网络、Trans 网络以及本文提出的方法生成的丰度图和端元图显示出更接近于地面真值图(GTs)的视觉相似性。这主要是因为深度学习方法的学习能力较强，覆盖面更广，适应性更强。

4. 结论

本文中, 该研究引入了一个基于多尺度池化和 Transformer 的网络结构以用于高光谱图像进行解混、降噪、超分辨率等任务。在任务性能上, 本文以解混任务为例, 与基准网络相比该网络在平均解混精度方面均取得了最佳实验结果。在网络结构上, 本研究在 PTCU 网络的编码器中设计多尺度池化操作的使用保证了 Transformer 编码器不但关注了局部信息的细粒度细节, 还加深了网络结构以充分发挥 Transformer 在长程网络上的优势, 从而进一步学习了图像的全局信息, 提高了网络性能。在实验研究上, 为了验证所提出的 PTCU 网络的有效性, 本文在 Samson、Apex 和合成数据集中进行了广泛的研究。实验结果表明, PTCU 网络是一种有效的解混方法, 在三种数据集上均展现出了优异的解混性能和富有竞争力的结果。最后, 本文研究者期待 PTCU 网络在相关领域的进一步应用, 为高光谱解混领域的发展做出贡献。

参考文献

- [1] Bioucas-Dias, J., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N.M. and Chanussot, J. (2013) Hyperspectral Remote Sensing Data Analysis and Future Challenges. *IEEE Transactions on Geoscience and Remote Sensing*, **1**, 6-36. <https://doi.org/10.1109/MGRS.2013.2244672>
- [2] Smith, M.O., Johnson, P.E. and Adams, J.B. (1985) Quantitative Determination of Mineral Types and Abundances from Reflectance Spectra Using Principal Component Analysis. *Proceedings Lunar and Planetary Science Conference*, Vol. 90, 797-904. <https://doi.org/10.1029/JB090iS02p0C797>
- [3] Lv, Z., Huang, H., Li, X., Zhao, M., Benediktsson, J.A., Sun, W. and Falco, N. (2022) Land Cover Change Detection with Heterogeneous Remote Sensing Images: Review, Progress, and Perspective. *Proceedings of the IEEE*, **110**, 1976-1991. <https://doi.org/10.1109/JPROC.2022.3219376>
- [4] Lv, Z., Zhong, P., Wang, W., You, Z. and Falco, N. (2023) Multiscale Attention Network Guided with Change Gradient Image for Land Cover Change Detection Using Remote Sensing Images. *IEEE Geoscience and Remote Sensing Letters*, **20**, 1-5. <https://doi.org/10.1109/LGRS.2023.3267879>
- [5] Keshava, N. and Mustard, J.F. (2002) Spectral Unmixing. *IEEE Signal Processing Magazine*, **19**, 44-57.
- [6] Bioucas-Dias, J.M., Plaza, A., Camps-Valls, G., Scheunders, P., Nasrabadi, N.M. and Chanussot, J. (2013) Hyperspectral Remote Sensing Data Analysis and Future Challenges. *IEEE Geoscience and Remote Sensing Magazine*, **1**, 6-36. <https://doi.org/10.1109/MGRS.2013.2244672>
- [7] Ghamisi, P., et al. (2017) Advances in Hyperspectral Image and Signal Processing: A Comprehensive Overview of the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, **5**, 37-78. <https://doi.org/10.1109/MGRS.2017.2762087>
- [8] Green, R.O., Eastwood, M.L., Sarture, C.M., Chrien, T.G., et al. (1998) Imaging Spectroscopy and the Airborne Visible/Infrared Imaging Spectrometer (AVIRIS). *Remote Sensing of Environment*, **65**, 227-248. [https://doi.org/10.1016/S0034-4257\(98\)00064-9](https://doi.org/10.1016/S0034-4257(98)00064-9)
- [9] Shaw, G. and Manolakis, D. (2002) Signal Processing for Hyperspectral Image Exploitation. *IEEE Signal Processing Magazine*, **19**, 12-16. <https://doi.org/10.1109/79.974715>
- [10] Keshava, N. and Mustard, J.F. (2002) Spectral Unmixing. *IEEE Signal Processing Magazine*, **19**, 44-57. <https://doi.org/10.1109/79.974727>
- [11] Manolakis, D. and Shaw, G. (2002) Detection Algorithms for Hyperspectral Imaging Applications. *IEEE Signal Processing Magazine*, **19**, 29-43. <https://doi.org/10.1109/79.974724>
- [12] Heinz, D.C. and Chang, C.-I. (2001) Fully Constrained Least Squares Linear Spectral Mixture Analysis Method for Material Quantification in Hyper-Spectral Imagery. *IEEE Transactions on Geoscience and Remote Sensing*, **39**, 529-545. <https://doi.org/10.1109/36.911111>
- [13] Drumetz, L., Meyer, T.R., Chanussot, J., Bertozzi, A.L. and Jutten, C. (2019) Hyperspectral Image Unmixing with Endmember Bundles and Group Sparsity Inducing Mixed Norms. *The IEEE Transactions on Image Processing*, **28**, 3435-3450. <https://doi.org/10.1109/TIP.2019.2897254>
- [14] Iordache, M.-D., Bioucas-Dias, J.M. and Plaza, A. (2011) Sparse Unmixing of Hyperspectral Data. *IEEE Transactions on Geoscience and Remote Sensing*, **49**, 2014-2039. <https://doi.org/10.1109/TGRS.2010.2098413>
- [15] Li, J., Bioucas-Dias, J.M., Plaza, A. and Liu, L. (2016) Robust Collaborative Nonnegative Matrix Factorization for

- Hyperspectral Unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, **54**, 6076-6090.
<https://doi.org/10.1109/TGRS.2016.2580702>
- [16] Qian, Y, Jia, S., Zhou, J. and Robles-Kelly, A. (2011) Hyperspectral Unmixing via $L_{1/2}$ Sparsity-Constrained Nonnegative Matrix Factorization. *IEEE Transactions on Geoscience and Remote Sensing*, **49**, 4282-4297.
<https://doi.org/10.1109/TGRS.2011.2144605>
- [17] Hong, D., *et al.* (2022) Endmember-Guided Unmixing Network (EGU-Net): A General Deep Learning Framework for Self-Supervised Hyperspectral Unmixing. *IEEE Transactions on Neural Networks and Learning Systems*, **33**, 6518-6531. <https://doi.org/10.1109/TNNLS.2021.3082289>
- [18] Rasti, B., Koirala, B., Scheunders, P. and Chanussot, J. (2022) MiSiCNet: Minimum Simplex Convolutional Network for Deep Hyperspectral Unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, Article ID: 5522815, <https://doi.org/10.1109/TGRS.2022.3146904>
- [19] Vaswani, A., *et al.* (2017) Attention Is All You Need.
- [20] Devlin, J., Chang, M.-W., Lee, K. and Toutanova, K. (2018) BERT: Pre-Training of Deep Bidirectional Transformers for Language Understanding.
- [21] Dosovitskiy, A., *et al.* (2020) An Image Is Worth 16×16 Words: Transformers for Image Recognition at Scale.
- [22] Ghosh, P., Roy, S.K., Koirala, B., Rasti, B. and Scheunders, P. (2022) Hyperspectral Unmixing Using Transformer Network. *IEEE Transactions on Geoscience and Remote Sensing*, **60**, Article ID: 5535116.
<https://doi.org/10.1109/TGRS.2022.3196057>