

面向小麦种植面积反演的自监督学习光谱解混算法

胡金龙^{1,2}, 段金亮¹, 沙马阿各¹, 张瑞^{1,3}

1. 西南交通大学 地球科学与环境工程学院测绘遥感信息系, 成都 611756;

2. 西南石油大学 土木工程与测绘学院, 成都 610500;

3. 西南交通大学 高速铁路运营安全空间信息技术国家地方联合工程实验室, 成都 611756

摘要: 遥感作为大范围地表覆盖提取和时空动态监测的有力工具, 在小麦种植面积反演领域具有显著优势。然而, 混合像元现象限制了小麦种植面积估计的精度, 端元的光谱变异性现象也使得传统的混合像元分解方法表现不佳。针对小麦种植面积估计过程中存在的混合像元和端元光谱变异性等问题, 本文提出基于自监督学习的光谱解混算法SLSU (Self-supervised Learning-based Spectral Unmixing Algorithm)。首先, 使用变分自编码器实现无监督的端元光谱变异性解释和端元库生成; 然后, 使用交替最小二乘法和全约束最小二乘模型估计各类端元对应的丰度; 最后, 利用概率松弛标记法对解混结果进行空间邻域校正, 以进一步提高光谱解混和小麦面积估计的精度。以河南省新乡市3个典型的小麦种植区为实验区, 利用Sentinel-2影像获取小麦种植面积并利用实地测量的小麦分布现状数据计算提取精度。结果表明: 基于SLSU算法获取的小麦种植面积的相对提取误差中位数<1.3个像素, 显著优于全约束最小二乘、扩展线性混合模型等传统混合像元分解算法以及支持向量机、随机森林等监督学习分类方法所提结果。所提SLSU算法可以提高小麦提取的精度和稳定性, 为农作物分布提取和种植面积估算提供了有效的方法。

关键词: 遥感, 自监督学习, 变分自编码器, 混合像元分解, 地物提取, 作物面积估计

中图分类号: TP751/P2

引用格式: 胡金龙, 段金亮, 沙马阿各, 张瑞. 2024. 面向小麦种植面积反演的自监督学习光谱解混算法. 遥感学报, 28(9): 2405-2415

Hu J L, Duan J L, Shama A G and Zhang R. 2024. Wheat acreage retrieval based on self-supervised learning-based spectral unmixing algorithm. National Remote Sensing Bulletin, 28(9): 2405-2415 [DOI:10.11834/jrs.20222315]

1 引言

农业生产是国家社会经济的基础, 粮食产量对于确保国家粮食安全和社会稳定、指导国家和区域社会经济发展规划制定等方面具有重要意义(王娜等, 2017; 吴炳方等, 2019)。农作物种植面积是影响粮食产量的关键问题, 而小麦作为中国主要的粮食作物之一, 准确、及时的小麦种植面积, 对于各级政府管理生成和制定相关政策至关重要(刘扬等, 2015)。遥感是大范围监测地表覆盖及其时空动态的有力工具, 基于遥感的小麦种植面积提取及产量估计也取得了较大的进展(尹捷等, 2021; 尹瀚民等, 2022; 冯志立等, 2022)。

由于地物的复杂性、影像空间分辨率的限制和电磁辐射传输过程中的各种环境因素, 遥感影像中普遍存在光谱混合的现象(胡茂桂和王劲峰, 2010; 蓝金辉等, 2018), 存在这种现象的像元称为混合像元。混合像元内包含多种地物, 而传统影像分类方法将像元作为不可再分的最小单位, 势必造成分类结果与实际地物分布的不符, 进而导致分类精度下降(段金亮等, 2021)。基于此, 将混合像元分解方法引入遥感地物分类和信息提取对提高结果的精度具有重要意义(Zhu等, 2022a, 2022c)。

传统的混合像元分解方法首先为影像构建端元库, 即各类地物对应的纯净像元的光谱曲线, 依据所获取的端元库和光谱混合模型进行混合像

收稿日期: 2022-06-30; 预印本: 2022-09-14

基金项目: 高分辨率对地观测重大专项航空观测系统项目(编号:30-H30C01-9004-19/21)

第一作者简介: 胡金龙, 研究方向为定量遥感和点云信息提取。E-mail: ltpai91@hotmail.com

通信作者简介: 张瑞, 研究方向为SAR图像处理 and 定量遥感。E-mail: zhangrui@swjtu.edu.cn

元的光谱解混, 得到各端元在混合像元中所占的比例 (即丰度) (段金亮 等, 2018)。国内外学者进行了大量的相关研究, 在端元提取 (Boardman, 1993; Winter, 1999; Ren 和 Chang, 2003) 和光谱混合模型 (Heinz 和 Chang, 2001; Drumetz 等, 2016a; Imbiriba 等, 2018) 等方面取得了重要的进展。然而, 由于各种因素 (环境条件、光照、大气以及时间等) 的影响, 端元存在光谱变异性 (Zare 和 Ho, 2014; Borsoi 等, 2020a), 即某一端元的光谱曲线会随条件的改变和产生一定的变化。这对于混合像元分解的性能产生了巨大的影响。

近年来, 国内外学者已提出一些解释和克服端元光谱变异性的方法, 如构建端元集合、将端元假设为特定分布以及参数化端元的变化系数 (Imbiriba 等, 2014; Drumetz 等, 2016a; Borsoi 等, 2020b)。其中, 因为良好的解混效果, 且避免了对光谱先验信息的依赖, 参数化方法被广泛研究。但是该方法仍然难以实现精度和维度的平衡, 即模型往往非常复杂, 而实际上真实的光谱变异性通常是低维的 (Borsoi 等, 2020a)。另一方面, 已有研究利用支持向量机 SVM (Support Vector Machine) 等监督分类方法来克服端元光谱变异性 (Okujeni 等, 2013; Wang 等, 2013), 但这些方法都依赖于大量的训练数据, 而先验的端元光谱信息获取的时间和经济成本非常高昂, 且方法的物理意义难以解释。

基于此, 本文提出基于自监督学习的光谱解混算法 SLSU (Self-supervised Learning-based Spectral Unmixing Algorithm)。首先, 基于初始端元和丰度, 使用在无监督学习领域被广泛使用的生成方法——变分自编码器 VAE (Variational Auto Encoder) (Kingma 和 Welling, 2022) 对端元的光谱变异性

进行解释, 获取各类端元的端元库。由于 VAE 方法只需要输入从影像中直接提取的端元光谱信息, 因此该方法是无监督的, 降低了方法的使用成本和数据依赖性。其后, 使用交替最小二乘法 ALS (Alternating Least-squares Strategy), 基于 VAE 方法所得到的端元库来估计全约束最小二乘模型 FCLS (Fully Constrained Least Squares) (Heinz 和 Chang, 2001) 的参数, 在此基础上得到各类端元对应的丰度, 实现考虑端元变异性的光谱解混。同时, 考虑到地物分布具有空间相关性, 引入概率松弛标记法 PLR (Probabilistic Relaxation Labelling) 描述影像的空间邻域信息, 对解混结果进行空间邻域校正, 以进一步提高解混精度。

2 研究方法

本文所提出的 SLSU 方法可以分为光谱初解混、VAE 端元光谱变异性解释和邻域信息后校正 3 个主要步骤。首先, 结合像元纯度指数 PPI (Pixel Purity Index) (Boardman, 1993)、N-FINDR (Winter, 1999) 以及自动目标生成 ATPG (Automatic Target Generation Process) (Ren 和 Chang, 2003) 等方法提取结果与实地纯净像元信息确定初始端元光谱。获得初始端元后, 使用全约束最小二乘算法 FCLS (Fully Constrained Least Squares) (Heinz 和 Chang, 2001) 对影像进行光谱解混, 得到初始丰度图。其后, 使用 VAE 方法对端元光谱变异性进行解释, 得到考虑光谱变异性的端元库。最后, 结合交替最小二乘方法和全约束最小二乘模型估计丰度图, 利用概率松弛标记法对丰度图进行校正, 得到最终的丰度图。SLSU 方法的具体流程如图 1 所示。

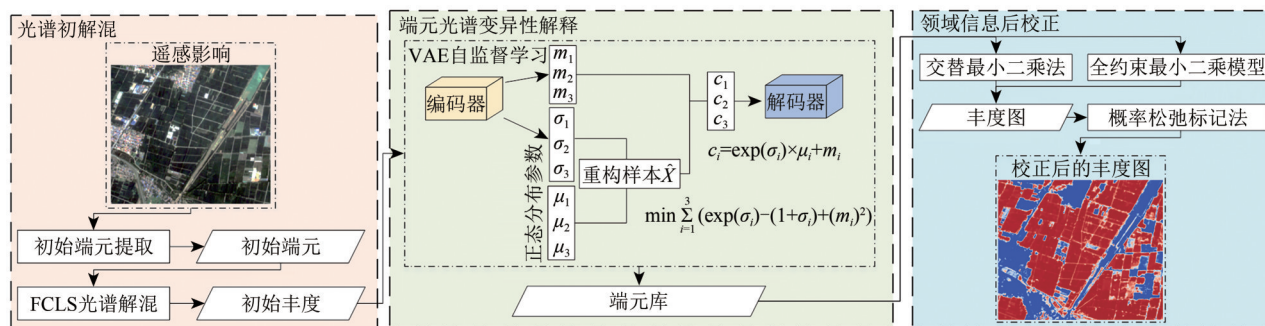


图1 SLSU 方法流程图

Fig. 1 Flow chart of the SLSU method

2.1 光谱解混

对于混合像元光谱解混问题，常采用的模型有线性混合模型、非线性混合模型和正态组分模型。其中，最常用的是线性混合模型 LMM (Linear Mixing Model)。LMM 假定对于一个有 L 个波段 N 个像元的影像，第 n 个混合像元：

$$\mathbf{y}_n = \mathbf{M}\mathbf{a}_n + \mathbf{e}_n \quad (1)$$

式中， $\mathbf{y}_n \in \mathbb{R}^L$ (即为有 L 个元素的实数向量)； $\mathbf{M} \in \mathbb{R}^{L \times P}$ 为端元矩阵，其中的每一列 $\mathbf{m}_k$ 为一个端元的光谱向量； P 为端元数目； $\mathbf{a}_n$ 是丰度向量； $\mathbf{e}_n$ 是满足白高斯分布 WGN (White Gaussian Noise) 的噪声向量。考虑到丰度的物理意义，丰度向量 $\mathbf{a}_n$ 应满足约束：

$$\begin{cases} \mathbf{1}^T \mathbf{a}_n = 1 \\ \mathbf{a}_n \geq 1 \end{cases} \quad (2)$$

即丰度各元素均为非负，且丰度之和为 1。

然而，由于各种因素的影响（环境条件、光照、大气或时间等），端元存在光谱变异性，即端元的光谱曲线会随条件产生一定的变化。这对于光谱解混的性能产生了巨大影响。为了削弱光谱变异性的影响，使用每种端元在不同情况下的光谱模型构建端元集，则待分解影像可表示为

$$\mathbf{Y} = \mathbf{M}\mathbf{A} + \mathbf{N} \quad (3)$$

式中， $\mathbf{Y} \in \mathbb{R}^{L \times N}$ 为影像矩阵， $\mathbf{A} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_N] \in \mathbb{R}^{P \times N}$ 为丰度矩阵， $\mathbf{N} \in \mathbb{R}^{L \times N}$ 为噪声矩阵。上述式 (3) 可以被视为一个矩阵分解问题，因此，可以使用 ALS 将该问题转换为一个最优化问题，其损失函数为

$$J(\mathbf{A}, \mathbf{B}) = \frac{1}{2} \sum_{n=1}^N \|\mathbf{y}_n - \mathbf{M}\mathbf{a}_n\|_F^2 + R(\mathbf{A}) + R(\mathbf{B}) \quad (4)$$

式中， $\mathbf{B} \in \mathbb{R}^{N \times L \times P}$ 为一个通过将所有像元的端元矩阵堆叠而形成的 3 波段张量， $R(\mathbf{A})$ 和 $R(\mathbf{B})$ 为正则化项。通过最小化损失函数：

$$(\hat{\mathbf{A}}, \hat{\mathbf{B}}) = \arg \min_{\mathbf{A}, \mathbf{B}} J(\mathbf{A}, \mathbf{B}) \quad (5)$$

即可求解得到所需的丰度矩阵 $\mathbf{A}$ 和对应的 $\mathbf{B}$ 。值得注意的是，当同时求解 $\mathbf{A}$ 和 $\mathbf{B}$ ，式 (5) 是非凸的。因此，通过对每个变量迭代最小化来近似求得最优解。

2.2 邻域信息后处理

传统解混方法只考虑单个像元的光谱属性，忽略了影像的邻域信息，如地物相关性和图像纹理等，制约了解混精度的提高。松弛标记法

(Rosenfeld 等, 1976; Wang 和 Hancock, 2008) 是利用符号来描述特性特征信息的方法，能将多种约束有效结合且不受实体数目的影响，利用松弛标记法能够实现解混结果的空间邻域信息校正。其中，处理对象被称为目标，描述目标的符号称为标记，先给定目标一组初始标记，通过迭代运算逐次更新标记，最后求得这组目标的确切标记集。目前，松弛标记法主要分为离散松弛标记法和概率松弛标记法（即连续松弛标记法）(Lee 和 Won, 2011; 张丁文 等, 2018)。离散松弛标记法在分配初始标记后，利用约束关系来剔除不一致的目标和标记所组成的分配组合。然而，固定的约束条件难以适用于光谱解混的应用。概率松弛标记法则通过引入相容性系数和支持函数，将固定的约束条件转换为概率分布，被广泛应用于遥感影像匹配和城市道路匹配等领域。因此，本文使用概率松弛标记法进行解混结果空间邻域校正。

首先，我们令中心像元 m 与其上方、下方、左侧、右侧的 4 个邻接像元组成 4 邻域，使其后验概率 $P_m(c_k) (k \in \mathbb{N}^+ \& k \in [1, R])$ 满足：

$$\sum_{k=1}^R P_m(c_k) = 1 \quad (6)$$

可见，中心像元的概率 P_m 将在一定程度上被相邻像元影响。利用概率松弛标记法，令相邻邻域的概率为 $Q_m(c_k)$ ，解混概率为 $P_m(c_k)$ ，则校正后的解混概率为

$$P'_m(c_k) = \frac{P_m(c_k)Q_m(c_k)}{\sum_{k'} P(c_{k'})Q(c_{k'})} \quad (7)$$

为了提高运算效率，对上式进行矢量化，得到：

$$P'_m(c_k) = \frac{P_m \otimes Q_R}{P_m^T Q_R} \quad (8)$$

根据贝叶斯概率理论，可推导出：

$$P'_m(c_k) = \frac{P_R \otimes (P_{mn} Q_R)}{P_m^T P_{mn} Q_R} \quad (9)$$

使用式 (9) 对解混结果进行校正，使最大后验概率近似于真实概率，即可使误分的像元 m 得以修正。

2.3 基于自监督学习光谱解混算法

在大量训练数据和适当的超参数辅助下，SVM 等监督学习方法能够取得较好的解混效果。然而，这些方法都依赖于大量的训练数据，而训

练数据的获取非常困难。一些工作试图采用混合软硬分类的方法降低计算成本,但是结果缺乏明确的物理解释。生成模型 GM (Generative Model) 是自监督学习中被广泛使用的神经网络。生成模型可以概括为通过学习和模仿输入样本 X 的特征,生成一些与 X 尽可能相似的新样本 $\hat{X}$ 的过程。自编码器 AE (Auto Encoder) 是经典的生成模型,它由编码器 (Encoder) 和解码器 (Decoder) 组成。编码器负责学习输入样本 X 的特征,而解码器则基于学习到的特征生成与 X 尽可能相似的 $\hat{X}$ 。然而,AE 只能依据输入的样本,生成一个新的样本。而我们需要依据输入的端元,建立对应地物类别的端元库,即多个样本,这是 AE 所无法处理的。因此,使用 AE 的改进算法——VAE 来建立端元库。VAE 方法不直接生成新的样本,而是通过估计输入样本 X 的分布,进而生成大量符合原始样本特征的新样本 $\hat{X}$ 。

VAE 的关键是基于变量 $\mathbb{R}^L \ni X \sim p(X)$ 的观测值 x_i 来估计 X 的概率密度分布 $p(X)$, 得到 $p(X)$ 的近似 $\hat{p}(X)$ 。通过对 $\hat{p}(X)$ 采样,便能得到与 X 具有相同特征的重构样本 $\hat{X}$ 。然而,由于 X 的维度 L 通常非常高,使得 $\hat{p}(X)$ 的获取变得非常困难,尤其是当 X 的样本不足时,因为这相当于对一个任意分布的高维概率密度函数进行估计和采样。针对这一问题,VAE 定义一个隐变量 $Z \in \mathbb{R}^K$ ($K \leq L$), 且假设 Z 服从某一高斯分布:

$$q_{\rho, \tau}(Z|X) = \mathcal{N}(\mu_\rho, \sigma_\tau^2) \quad (10)$$

式中, μ_ρ 和 σ_τ^2 分别为对应高斯分布的均值和标准差。与此同时,定义映射函数 $\mathcal{G}_\theta$ 使得:

$$\mathbb{R}^L \ni \hat{X} = \mathcal{G}_\theta(Z) \quad (11)$$

则只需要知道 Z 所符合的高斯分布的参数 μ_ρ 和 σ_τ^2 , 以及映射函数 $\mathcal{G}_\theta$, 就能得到原始数据的分布,进而可以基于此生成与 X 非常相似的重构样本 $\hat{X}$ 。

为了学习到 Z 对应高斯分布的参数 μ_ρ 和 σ_τ^2 , 需要足够多的 Z 的样本。因此,定义一个映射(条件概率分布)函数 $\mathcal{Q}_\varphi$, 使得:

$$Z = \mathcal{Q}_\varphi(Z|X) = f(X, \mathcal{N}(\mu_\rho, \sigma_\tau^2)) \quad (12)$$

即基于输入 X 和对应的正态分布参数生成隐变量 Z 的映射。使用两个神经网络分别拟合参数 μ_ρ 和 σ_τ^2 (称为编码器), 使用一个神经网络拟合映射函数 $\mathcal{G}_\theta$ (称为解码器)。将变量 X 的观测值输入模

型, 编码器将其编码为隐变量 Z , 解码器将 Z 解码为 $\hat{X}$ 。经过上次步骤, 即可实现从 X 的样本中创建大量具有 X 特征的新样本 $\hat{X}$ 。

这与 AE 方法中 $X \xrightarrow{\text{Encoder}} \text{Features} \xrightarrow{\text{Decoder}} \hat{X}$ 的处理流程非常相似。模型实现的关键在于训练两个映射函数 $\mathcal{Q}_\varphi$ 和 $\mathcal{G}_\theta$, 使得 X 和 $\hat{X}$ 尽可能接近。该模型的损失函数由两部分构成, 第一部分为 X 和 $\hat{X}$ 的重构误差:

$$L_E = \int_Z q_{\rho, \tau}(Z|X) \log p(X|Z) dZ \quad (13)$$

式中, $p(X|Z)$ 为原始数据的概率分布。此外, 由于假设 Z 服从高斯分布, 因此编码器对应的条件概率分布 $\mathcal{N}(\mu_\rho, \sigma_\tau^2)$ 应尽可能地与标准高斯分布 $\mathcal{N}(0, 1)$ 相似, 这两个分布的相似性可以用相对熵(又称为 KL 散度, Kullback-Leibler Divergence)来衡量, 计算公式为

$$\begin{aligned} L_{KL} &= \text{KL}(q_{\rho, \tau}(Z|X) \| p(X)) = \\ &= \text{KL}(\mathcal{N}(\mu, \sigma^2) \| \mathcal{N}(0, 1)) = \\ &= \frac{-\log \sigma^2 + \mu^2 + \sigma^2 - 1}{2} \end{aligned} \quad (14)$$

综上, VAE 模型的损失函数为

$$L = L_E - L_{KL} \quad (15)$$

使用上述方法建立基于 VAE 的端元生成模型, 并利用初始端元和 FLCs 方法估计出的初始丰度训练端元生成模型。其后, 将待分解影像 Y 的所有像元输入训练好的端元生成模型, 得到考虑每个像元光谱变异性的端元矩阵 $\hat{X}$ 。将 $\hat{X}$ 代入式 (3), 得到:

$$A = \hat{X}^{-1}Y - N \quad (16)$$

利用式 (5) 的最小化损失函数对上式进行求解, 得到待分解影像 Y 的丰度矩阵 A 。最后, 利用影像的邻域信息对 A 进行邻域信息后校正, 即可得到最终的丰度矩阵:

$$A' = AP'_m(c) \quad (17)$$

3 实验设计与结果分析

3.1 研究区概况

选择以小麦作为主要农作物的河南省新乡市为研究区。新乡市位于河南省北部, 地理坐标 $113^\circ 23'E - 115^\circ 01'E$, $34^\circ 53'N - 35^\circ 50'N$ 。新乡市东接油城濮阳, 与鲁西相连; 南邻黄河, 与郑州、开封隔河相望; 西接焦作, 并与山西接壤; 北依太行,

与鹤壁、安阳毗邻。新乡市总面积达到 8249 km²，境内地形以平原为主，除西北部为低山丘陵区，余境均为黄河、卫河冲击平原区，地势西北高、东南低（王学强，2007；刘战东 等，2008）。该市年平均气温 14 ℃，年降雨量为 573.4 mm，平均无霜期为 205 d，属于暖温带大陆性季风气候。从新乡市下辖的全国优质小麦产业化示范县——封丘县和延津县中选择 3 个面积为 9 km² 正方形区域（3 km×3 km）作为实验样区，具体样区位置及其概况分别如图 2 和表 1 所示。

3.2 实验数据及处理

Sentinel-2 是欧洲航天局 ESA（European Space Agency）哥白尼计划中的一项地球观测任务，使用多架卫星组成的卫星星座，拍摄地球表面的中分辨率影像（Pohl 和 Van Genderen，2016）。Sentinel-2 影像作为目前公众可免费获得的最高空间分辨率多光谱遥感数据，具有很高的应用价值，被广泛应用于森林监测、水质监测、土地覆盖变化侦测、自然灾害管理等领域，取得了显著的成果。因此，

本文以 Sentinel-2 影像作为实验数据，数据来源于 ESA 哥白尼开放获取中心（Copernicus Open Access Hub）。

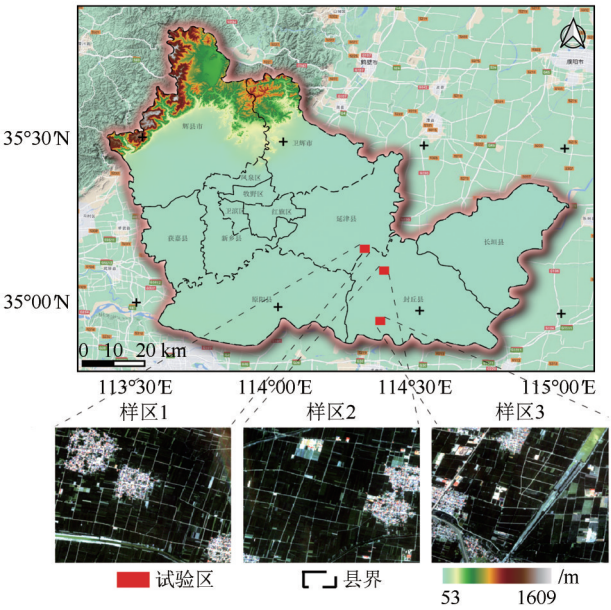


图2 试验区示意图
Fig. 2 Study area

表 1 实验样区概况

Table 1 Overview of the sample areas

样区编号	所跨行政区划	面积/km ²	地物类型
样区 1	延津县司寨乡封丘县城固乡	9	建设用地、小麦、河流、林地等
样区 2	封丘县王村乡	9	建设用地、小麦、蒜苗、金银花等
样区 3	封丘县荆隆宫乡	9	建设用地、小麦、河流、林地等

为了获取实验样本，本团队于 2021 年 4 月底至 5 月初对试验区进行了实地调查和测量，图 3 为试验区 2 的实地调研场景。经过整理后共获得 940 个图斑，各类别图斑数量如表 2 所示。可见，试验区地类以耕地和建筑用地为主，其中主要农作物为小麦、蒜苗和金银花，适合本文所提算法验证。



图3 试验区实地调研
Fig. 3 Field trip to the study area

表 2 实测图斑概况

Table 2 Overview of the measured patches

图斑类别	图斑数量/个	图斑总面积/hm ²
建筑用地	349	56.39
小麦	142	187.15
蒜苗	151	33.31
金银花	134	19.65
其他菜地	92	10.10
林地	72	27.04

与所测量的样本数据对应，选择 2021 年 5 月 2 日拍摄的 Sentinel-2 影像作为实验数据，该影像处理基线和相对轨道编号分别为 N0300 和 R075，影像云量均低于 15%，满足实验需求。由于 Sentinel-2 影像数据具有 10 m、20 m 以及 60 m 等多种空间分辨率，不便于后续使用，因此，使用超分辨率重建算法将 Sentinel-2 的 20 m 和 60 m 波段数据全部

重建为 10 m 空间分辨率的影像。同时, 鉴于获取的影像数据为 Level 1C 原始影像数据, 获得试验区的影像数据后, 分别对影像进行辐射定标和大气校正处理, 从而获得地面反射率数据。其后, 从校正后的影像中裁剪出试验区范围, 得到 3 个试验区的最终影像, 其真彩色合成效果见图 2。

另一方面, 为了对各种方法所提小麦面积精度进行定量评价, 将实地测量的各小麦矢量图斑使用“面转栅格”方式采样为 0.1 m 空间分辨率的栅格数据, 像素值分配依据为像元内面积最大的要素对应的值。其后, 计算各图斑所占面积, 作为小麦样本的面积真值 a_T 。依据栅格分辨率和小麦所占像元数计算监督分类方法小麦面积测量值, 使用小麦丰度图计算混合像元分解方法所得小麦面积测量值。设各样本由多种方法提取的面积测量值为 a_M , 则各方法所提小麦的相对误差 ERE (Extraction Relative Error) 可表示为

$$\text{ERE} = \text{abs}(a_M - a_T) / \text{size}_p^2 \quad (18)$$

式中, size_p 为像元的空间分辨率, 此处为 10 m。将各方法所提取小麦分布的误差由绝对尺度转化为相对于的像元空间分辨率的相对尺度, 便于不同空间分辨率影像之间的相互比较, 提高了精度指标的可比性。

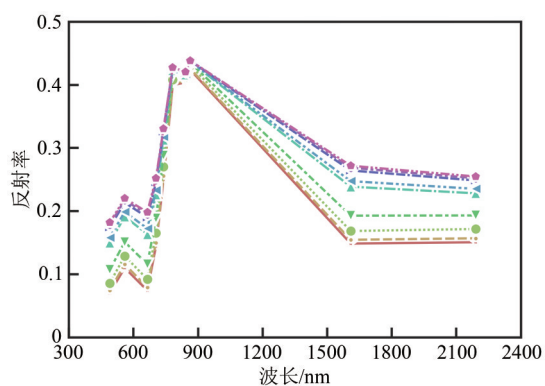
利用所提出的 SLSU 方法, 基于获取到的初始端元和初始丰度分别对 3 个试验区进行光谱解混与小麦提取。SLSU 方法使用 VAE 对端元的光谱变异性建模, 能够依据每个像素的特点生成对应的端元光谱。将试验区 1 生成的所有端元按照小数点后

1 位精度对其光谱反射率进行简化和去重, 得到样区典型地物的端元光谱曲线。为了验证所提方法解混的可靠性, 使用扩展线性混合模型 ELMM (Extended Linear Mixture Model) (Drumetz 等, 2016b)、扰动线性混合模型 PLMM (Perturbed Linear Mixture Model) (Thouvenin 等, 2016) 以及广义线性混合模型 GLMM (Generalized Linear Mixed Model) (Imbiriba 等, 2018) 等 3 种方法进行光谱解混和小麦提取。同时, 使用支持向量机 SVM (Support Vector Machine)、随机森林 RF (Random Forest) 和极大似然分类法 MLC (Maximum Likelihood Classification) 等 3 种监督学习方法提取试验区小麦分布。

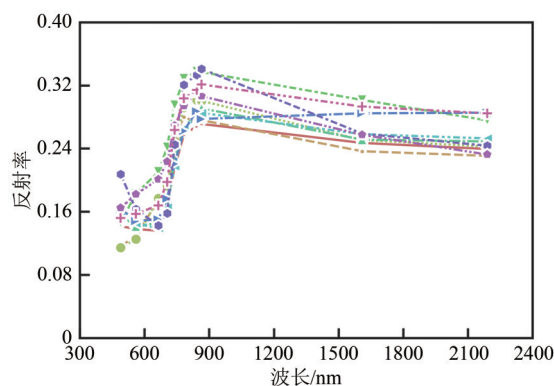
3.3 结果及分析

SLSU 方法所提取的样区典型地物的端元光谱曲线如图 4 所示。SLSU 方法所提取地物端元既满足地物的光谱特征, 又充分考虑了端元的光谱变异性, 方法结果符合预期。

5 种方法所提取的小麦丰度图如图 5 所示。FCLS 方法所提结果丰度较为粗糙, 丰度值主要分布于 0 和 1 附近; PLMM 方法所提结果破碎性较强, 且整体丰度偏低; ELMM、GELMM 以及本文所提的 SLSU 等 3 种方法所提结果丰度过渡平滑, 能够体现小麦丰度在不同地区的变化。小麦分布提取结果如图 6 所示。SVM 等监督分类方法能够较好地研究区中提取植被, 但难以实现植被类型的区分, 将绝大部分植被覆盖区识别为小麦; 而本文所提出的 SLSU 方法所提结果则能够较好地体现不同区域的小麦丰度。



(a) 小麦
(a) Wheat



(b) 建设用地
(b) Imperviousness

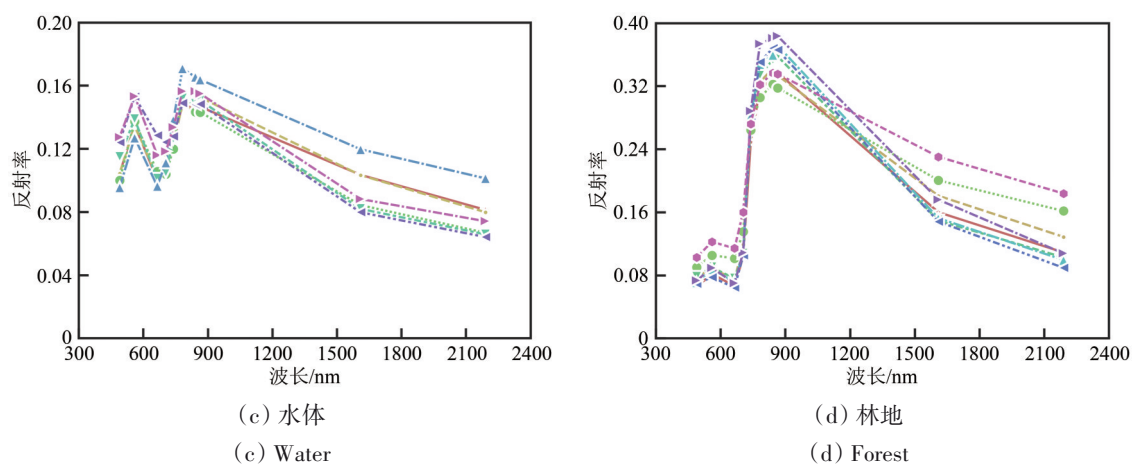


图4 SLSU所提取的试验区典型地物的光谱特征

Fig. 4 Spectral reflective curve of typical objects in the sample areas extracted by the SLSU method

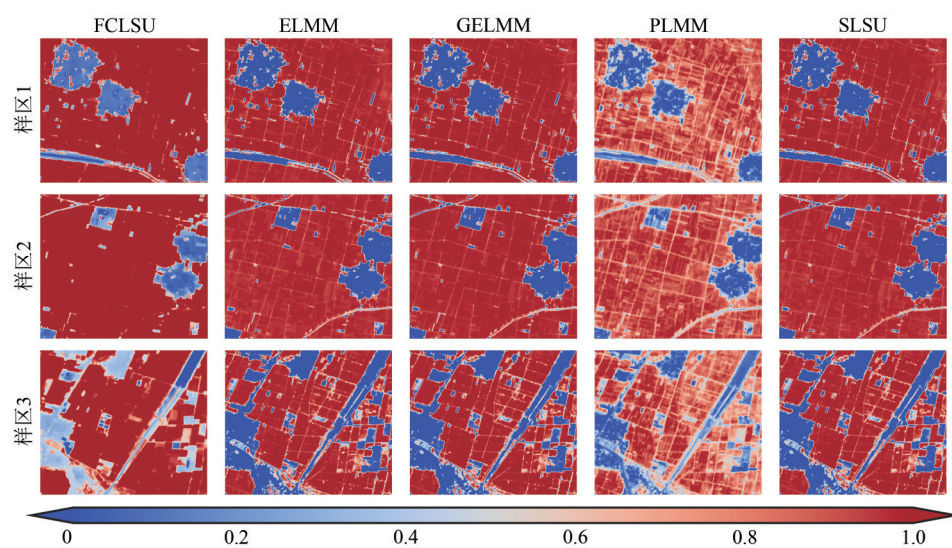


图5 3个试验区所提取小麦丰度图

Fig. 5 Abundance of wheat extracted from three sample areas

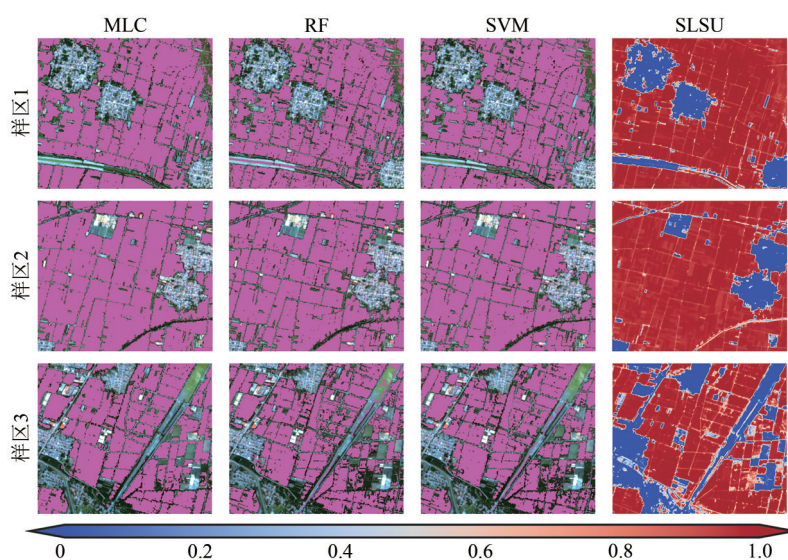


图6 试验区监督分类方法所提取小麦分布

Fig. 6 Distribution of wheat extracted by the supervised classification method in the sample areas

分别计算3个试验区小麦样本多种提取方法的误差ERE，结果见表3。基于混合像元分解的方案相比基于像素的监督分类方法普遍具有更高的精度；SLSU方法在所有方法中具有最高的精度，其所提取小麦的ERE在3个实验区的均方根误差分别为1.61、1.25和12.91个像素；SVM等监督分类方法ERE最低均值和中位数分别为12.49和7.30个像素，接近SLSU方法的6倍。为了更加直观的比较各方法提取精度，使用箱形图展示小麦提取误差分布，结果如图7所示。监督分类方法的结果相比光谱解混方法的结果普遍存在ERE均值和中位数较高，且四分位数区间较大的特征；与监督分类方法相比，光谱解混方法的结果普遍具有较高

的精度。其中PLMM方法对小麦丰度存在低估，导致其所提小麦面积精度相对较低；SLSU方法所提结果的ERE最大中位数为1.23个像素，ERE最大上四分位接近其他方法的ERE下四分位数。进一步分析各方法的时间成本，各方法的训练时间如表4所示。可知SLSU等混合像元分解方法普遍比以MLC为代表的监督分类方法具有更高的时间成本，其中尤以PLMM方法为著，其时间成本远超其他方法。受限于深度学习的复杂性，本文所提SLSU方法时间成本也较高，如何降低方法的时间复杂度仍有待进一步研究。综上，本文所提出的SLSU方法无论是精度还是稳定性相比其他方法都具有较大的优势。

表3 各方法所提小麦面积的相对误差
Table 3 Extraction relative errors of wheat area

方法	样区1				样区2				样区3			
	均值	中值	标准差	RMSE	均值	中值	标准差	RMSE	均值	中值	标准差	RMSE
MLC	14.46	7.48	17.21	22.21	12.49	7.30	14.48	18.84	29.17	14.41	49.11	56.43
RF	17.62	11.78	20.21	26.51	15.54	10.37	16.86	22.61	39.63	22.04	68.36	78.05
SVM	17.38	7.62	21.42	27.25	14.79	12.12	15.91	21.43	37.56	18.04	73.66	81.62
FCLSU	3.38	2.18	3.86	5.08	6.58	4.68	5.52	8.50	11.05	4.28	17.47	20.43
ELMM	2.44	1.39	2.81	3.68	4.53	3.37	3.86	5.89	9.99	3.72	17.24	19.69
GELMM	3.33	2.13	3.81	5.00	6.47	4.62	5.43	8.36	11.01	4.23	17.46	20.40
PLMM	14.82	8.01	17.66	22.78	7.79	6.62	7.95	10.99	24.87	10.99	33.18	41.04
SLSU	1.28	1.12	1.00	1.61	0.90	0.56	0.88	1.25	4.03	1.23	12.46	12.91

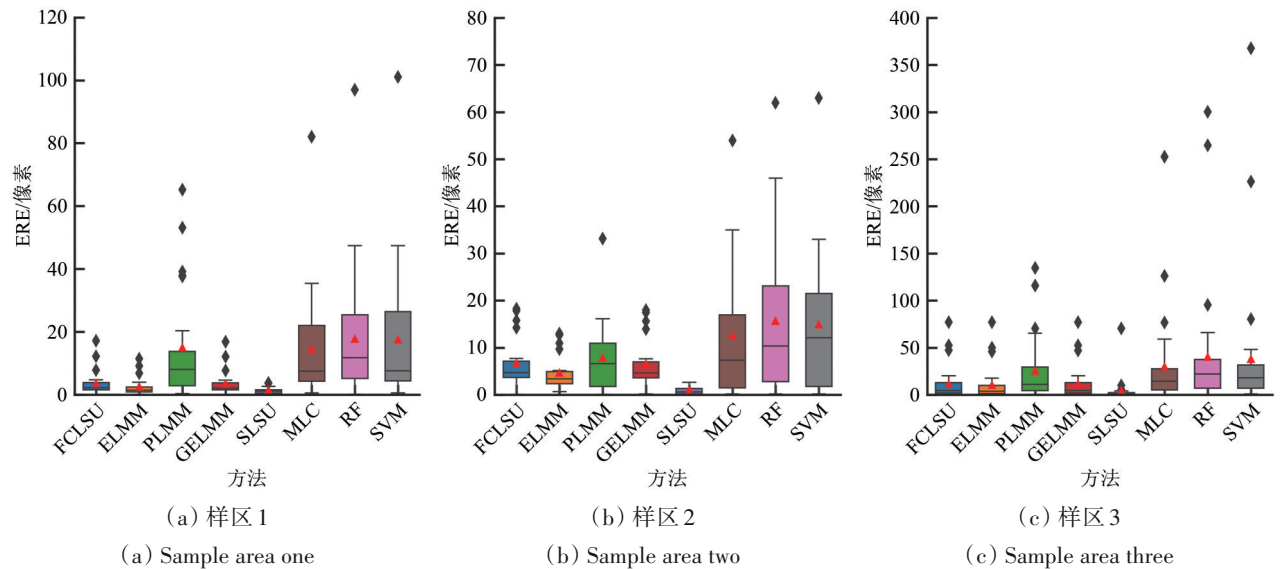


图7 各方法所提小麦误差分布
Fig. 7 Distribution of wheat extraction relative errors

表4 各方法训练时间
Table 4 Training time for each method

方法	训练时间			/s
	样区 1	样区 2	样区 3	
MLC	0.86	0.15	0.14	
RF	0.55	0.43	0.42	
SVM	123.61	60.12	50.34	
FCLSU	5.75	5.55	5.97	
ELMM	73.58	126.17	79.48	
GELMM	31.97	67.99	32.82	
PLMM	1999.91	1576.51	1423.17	
SLSU	1238.91	1160.73	1311.26	

4 讨 论

由于遥感影像空间分辨率的限制，影像像元内不可避免的存在光谱混合现象。基于像素的遥感地物分类和提取方法，如SVM等监督分类方法和NDVI等指数方法，都不可避免的受到光谱混合的影响，限制了提取结果的精度。基于混合像元分解的方法则考虑了像元内不同地物的光谱混合，运用光谱混合的物理模型得到像元内不同地物（端元）所占的面积比例（丰度），从而获得更加精确的地物分布。然而，由于各种因素的影响，端元存在光谱变异性，限制了光谱解混的性能。SLSU方法将VAE算法引入模型，使用自监督学习算法实现端元库的无监督生成，通过该方式将端元的光谱变异性纳入解混模型，一方面提高了解混的精度，另一方面增强了方法在不同场景下的稳定性和鲁棒性。同时，概率松弛标记法的引入实现了基于邻域信息的解混结果校正，在DAEN等模型基础上进一步提高了所提方法的精度（Su等，2019）。然而，相比S³TRM方法在解混的同时考虑光谱空间特征，SLSU方法在解混完成后再进行空间邻域后校正，模型更为复杂且后校正缺乏对数据的自适应（Zhu等，2022b），如何协同光谱解混与空间邻域信息仍是值得研究的问题。

5 结 论

本文针对遥感小麦面积反演中存在的混合像元和光谱解混时的端元光谱变异性现象，将小麦提取转换为小麦丰度估计问题，并提出了一种基于自监督学习的光谱解混算法——SLSU算法。通

过实验验证主要结论如下：

（1）SLSU算法在影像光谱解混方面较ELMM、PLMM和GLMM等传统解混模型更有优势，其通过变分自编码器实现对端元光谱变异性的无监督建模，使得SLSU算法解混精度显著提高。

（2）在面向小麦面积反演应用，SLSU算法较传统解混模型以及SVM、RF和MLC等机器学习方法更加准确，不仅其反演面积更加接近样本实测面积，而且更为稳定，体现出较好的应用潜力。

总体来看，SLSU算法对于多光谱影像的谱解混有较好的效果，同时在遥感小麦面积反演领域也表现出较高的精度和稳定性。然而，SLSU算法基于自监督深度学习实现端元光谱变异性建模，所需算力相对较高，后续的研究中将重点考虑对变分自编码器模型进行模型压缩和模型修剪，以提升算法运行效率（Han等，2015；He等，2019）。

参考文献(References)

- Boardman J W. 1993. Automating spectral unmixing of AVIRIS data using convex geometry concepts//Summaries of the 4th Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop. California: JPL: 11-14
- Borsoi R A, Imbiriba T and Bermudez J C M. 2020a. Deep generative endmember modeling: an application to unsupervised spectral unmixing. *IEEE Transactions on Computational Imaging*, 6: 374-384 [DOI: 10.1109/TCI.2019.2948726]
- Borsoi R A, Imbiriba T and Bermudez J C M. 2020b. A data dependent multiscale model for hyperspectral unmixing with spectral variability. *IEEE Transactions on Image Processing*, 29: 3638-3651 [DOI: 10.1109/TIP.2020.2963959]
- Drumetz L, Chanussot J and Jutten C. 2016a. Variability of the endmembers in spectral unmixing: recent advances//2016 8th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). Los Angeles: IEEE: 1-5 [DOI: 10.1109/WHISPERS.2016.8071675]
- Drumetz L, Veganzones M A, Henrot S, Phlypo R, Chanussot J and Jutten C. 2016b. Blind hyperspectral unmixing using an extended linear mixing model to address spectral variability. *IEEE Transactions on Image Processing*, 25(8): 3890-3905 [DOI: 10.1109/TIP.2016.2579259]
- Duan J L, Wang J and Zhang T. 2018. A kind of vegetation cover fraction retrieval algorithm based on spectral normalization frame. *Remote Sensing Technology and Application*, 33(2): 252-258 (段金亮, 王杰, 张婷. 2018. 一种基于光谱归一化下的植被覆盖度反演算法. *遥感技术与应用*, 33(2): 252-258) [DOI: 10.11873/j.issn.1004-0323.2018.2.0252]
- Duan J L, Zhang R, Li K and Pang J T. 2021. A kind of extended linear

- spectral unmixing algorithm based on noise level estimation. *Remote Sensing Technology and Application*, 36(4): 820-826 (段金亮, 张瑞, 李奎, 庞家泰. 2021. 一种基于噪声水平估计的扩展线性光谱分解算法. *遥感技术与应用*, 36(4): 820-826 [DOI: 10.11873/j.issn.1004-0323.2021.4.0820])
- Feng Z L, Xiao F, Lu X P, Hao B, Wang R Y and Zhu R. 2022. Winter wheat classification method based on feature optimization of random forest. *Bulletin of Surveying and Mapping*, (3): 70-75 (冯志立, 肖锋, 卢小平, 郝波, 王如意, 朱睿. 2022. 基于随机森林特征优选的冬小麦分类方法. *测绘通报*, (3): 70-75 [DOI: 10.13474/j.cnki.11-2246.2022.0080])
- Han S, Pool J, Tran J and Dally W J. 2015. Learning both weights and connections for efficient neural networks//*Proceedings of the 28th International Conference on Neural Information Processing Systems*. Montreal: MIT Press: 1135-1143
- He Y, Liu P, Wang Z W, Hu Z L and Yang Y. 2019. Filter pruning via geometric median for deep convolutional neural networks acceleration//*Proceedings of the 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Long Beach: IEEE: 4335-4344 [DOI: 10.1109/CVPR.2019.00447]
- Heinz D C and Chang C I. 2001. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Transactions on Geoscience and Remote Sensing*, 39(3): 529-545 [DOI: 10.1109/36.911111]
- Hu M G and Wang J F. 2010. Mixed-pixel decomposition and super-resolution reconstruction of RS image. *Progress in Geography*, 29(6): 747-756 (胡茂桂, 王劲峰. 2010. 遥感影像混合像元分解及超分辨率重建研究进展. *地理科学进展*, 29(6): 747-756 [DOI: 10.11820/dlkxjz.2010.06.015])
- Imbiriba T, Bermudez J C, Tournet J Y and Richard C. 2014. Detection of nonlinear mixtures using Gaussian processes: application to hyperspectral imaging//*2014 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Florence: IEEE: 7949-7953 [DOI: 10.1109/ICASSP.2014.6855148]
- Imbiriba T, Borsoi R A and Bermudez J C M. 2018. Generalized linear mixing model accounting for endmember variability//*2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. Calgary: IEEE: 1862-1866 [DOI: 10.1109/ICASSP.2018.8462214]
- Kingma D P and Welling M. 2022. Auto-encoding variational bayes. *arXiv preprint arXiv: 1312.6114* [DOI: 10.48550/arXiv.1312.6114]
- Lan J H, Zou J L, Hao Y S, Zeng Y L, Zhang Y Z and Dong M W. 2018. Research progress on unmixing of hyperspectral remote sensing imagery. *Journal of Remote Sensing (in Chinese)*, 22(1): 13-27 (蓝金辉, 邹金霖, 郝彦爽, 曾溢良, 张玉珍, 董铭巍. 2018. 高光谱遥感影像混合像元分解研究进展. *遥感学报*, 22(1): 13-27 [DOI: 10.11834/jrs.20186502])
- Lee J H and Won C H. 2011. Topology preserving relaxation labeling for nonrigid point matching. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 33(2): 427-432 [DOI: 10.1109/TPAMI.2010.179]
- Liu Y, Fu Z Y and Zheng F B. 2015. Review on high resolution remote sensing image classification and recognition. *Journal of Geo-information Science*, 17(9): 1080-1091 (刘扬, 付征叶, 郑逢斌. 2015. 高分辨率遥感影像目标分类与识别研究进展. *地球信息科学学报*, 17(9): 1080-1091 [DOI: 10.3724/SP.J.1047.2015.01080])
- Liu Z D, Duan A W, Gao Y and Liu H. 2008. Study on dynamic model of leaf area index (LAI) for winter wheat in Xinxiang area. *Journal of Triticeae Crops*, 28(4): 680-685 (刘战东, 段爱旺, 高阳, 刘浩. 2008. 河南新乡地区冬小麦叶面积指数的动态模型研究. *麦类作物学报*, 28(4): 680-685)
- Okujeni A, van der Linden S, Tits L, Somers B and Hostert P. 2013. Support vector regression and synthetically mixed training data for quantifying urban land cover. *Remote Sensing of Environment*, 137: 184-197 [DOI: 10.1016/j.rse.2013.06.007]
- Pohl C and Van Genderen J. 2016. *Remote Sensing Image Fusion: A Practical Guide*. Boca Raton: CRC Press [DOI: 10.1201/9781315370101]
- Ren H and Chang C I. 2003. Automatic spectral target recognition in hyperspectral imagery. *IEEE Transactions on Aerospace and Electronic Systems*, 39(4): 1232-1249 [DOI: 10.1109/TAES.2003.1261124]
- Rosenfeld A, Hummel R A and Zucker S W. 1976. Scene labeling by relaxation operations. *IEEE Transactions on Systems, Man, and Cybernetics*, SMC-6(6): 420-433 [DOI: 10.1109/TSMC.1976.4309519]
- Su Y C, Li J, Plaza A, Marinoni A, Gamba P and Chakravorty S. 2019. DAEN: deep autoencoder networks for hyperspectral unmixing. *IEEE Transactions on Geoscience and Remote Sensing*, 57(7): 4309-4321 [DOI: 10.1109/TGRS.2018.2890633]
- Thouvenin P A, Dobigeon N and Tournet J Y. 2016. Hyperspectral unmixing with spectral variability using a perturbed linear mixing model. *IEEE Transactions on Signal Processing*, 64(2): 525-538 [DOI: 10.1109/TSP.2015.2486746]
- Wang H F and Hancock E R. 2008. Probabilistic relaxation labelling using the Fokker-Planck equation. *Pattern Recognition*, 41(11): 3393-3411 [DOI: 10.1016/j.patcog.2008.03.030]
- Wang L G, Liu D F, Wang Q M and Wang Y. 2013. Spectral unmixing model based on least squares support vector machine with unmixing residue constraints. *IEEE Geoscience and Remote Sensing Letters*, 10(6): 1592-1596 [DOI: 10.1109/LGRS.2013.2262371]
- Wang N, Li Q Z, Du X, Zhang Y, Zhao L C and Wang H Y. 2017. Identification of main crops based on the univariate feature selection in Subei. *Journal of Remote Sensing (in Chinese)*, 21(4): 519-530 (王娜, 李强子, 杜鑫, 张源, 赵龙才, 王红岩. 2017. 单变量特征选择的苏北地区主要农作物遥感识别. *遥感学报*, 21(4): 519-530 [DOI: 10.11834/jrs.20176373])
- Wang X Q. 2007. *Study on Wheat's Productive Potential and Developmental Strategy in Henan Province*. Yangling: Northwest A&F University: 129 (王学强. 2007. 河南小麦生产潜力及发展战略研究. 杨凌: 西北农林科技大学: 129)
- Winter M E. 1999. N-FINDR: an algorithm for fast autonomous spectral end-member determination in hyperspectral data//*Proceedings Volume 3753, Imaging Spectrometry V*. Denver: SPIE: 266-275 [DOI: 10.1117/12.366289]

- Wu B F, Zhang M, Zeng H W, Yan N N, Zhang X, Xing Q and Chang S. 2019. Twenty years of CropWatch: progress and prospect. *Journal of Remote Sensing (in Chinese)*, 23(6): 1053-1063 (吴炳方, 张森, 曾红伟, 闫娜娜, 张鑫, 邢强, 常胜. 2019. 全球农情遥感速报系统 20 年. *遥感学报*, 23(6): 1053-1063) [DOI: 10.11834/jrs.20198156]
- Yin H M, Jiapaer G, Yu T, Umuhoza J and Li X. 2022. Wheat yield estimation with remote sensing in northern Kazakhstan. *Arid Land Geography*, 45(2): 488-498 (尹瀚民, 古丽·加帕尔, 于涛, Umuhoza J, 李旭. 2022. 哈萨克斯坦北部小麦遥感估产方法研究. *干旱区地理*, 45(2): 488-498) [DOI: 10.12118/j.issn.1000-6060.2021.236]
- Yin J, Zhou L L, Li L W, Zhang Y Q, Huang W J, Zhang H L, Wang Y, Zheng S J, Fan H S, Ji C, Chen J J and Peng D L. 2021. A comparative study on wheat identification and growth monitoring based on multi-source remote sensing data. *Remote Sensing Technology and Application*, 36(2): 332-341 (尹捷, 周雷雷, 李利伟, 张雅琼, 黄文江, 张赫林, 王岩, 郑诗军, 范海生, 纪婵, 陈俊杰, 彭代亮. 2021. 多源遥感数据小麦识别及长势监测比较研究. *遥感技术与应用*, 36(2): 332-341) [DOI: 10.11873/j.issn.1004-0323.2021.2.0332]
- Zare A and Ho K C. 2014. Endmember variability in hyperspectral analysis: addressing spectral variability during spectral unmixing. *IEEE Signal Processing Magazine*, 31(1): 95-104 [DOI: 10.1109/MSP.2013.2279177]
- Zhang D W, Chen Z L and Xie Z. 2018. Spatial scenes matching with on relaxation labeling approach. *Geomatics and Information Science of Wuhan University*, 43(5): 752-758 (张丁文, 陈占龙, 谢忠. 2018. 利用松弛标记法进行空间场景匹配. *武汉大学学报 (信息科学版)*, 43(5): 752-758) [DOI: 10.13203/j.whugis.20150719]
- Zhu Q Q, Deng W H, Zheng Z, Zhong Y F, Guan Q F, Lin W H, Zhang L P and Li D R. 2022c. A spectral-spatial-dependent global learning framework for insufficient and imbalanced hyperspectral image classification. *IEEE Transactions on Cybernetics*, 52(11): 11709-11723 [DOI: 10.1109/TCYB.2021.3070577]
- Zhu Q Q, Lei Y, Sun X L, Guan Q F, Zhong Y F, Zhang L P and Li D R. 2022a. Knowledge-guided land pattern depiction for urban land use mapping: a case study of Chinese cities. *Remote Sensing of Environment*, 272: 112916 [DOI: 10.1016/j.rse.2022.112916]
- Zhu Q Q, Wang L L, Chen J L, Zeng W, Zhong Y F, Guan Q F and Yang Z J. 2022b. S³TRM: spectral-spatial unmixing of hyperspectral imagery based on sparse topic relaxation-clustering model. *IEEE Transactions on Geoscience and Remote Sensing*, 60: 5515613 [DOI: 10.1109/TGRS.2021.3117250]

Wheat acreage retrieval based on self-supervised learning-based spectral unmixing algorithm

HU Jinlong^{1,2}, DUAN Jinliang¹, SHAMA Age¹, ZHANG Rui^{1,3}

1. Faculty of Geoscience and Environmental Engineering, Southwest Jiaotong University, Chengdu 611756, China;

2. School of Civil Engineering and Geomatics, Southwest Petroleum University, Chengdu 610500, China;

3. State-Province Joint Engineering Laboratory of Spatial Information Technology of High-Speed Rail Safety, Southwest Jiaotong University, Chengdu 611756, China

Abstract: Wheat is an important food crop in China, and its distribution acquisition and planting area estimation are crucial to ensuring national food security and development planning. As a powerful tool for large-scale land cover extraction and temporal and spatial dynamic monitoring, remote sensing has remarkable advantages in the field of wheat planting area estimation. However, wheat is often mixed with other ground objects and forms mixed pixels limited by the spatial resolution of remote sensing images and the influence of various factors in the transmission of electromagnetic radiation. The spectral mixing limits the accuracy of wheat planting area estimation. Meanwhile, endmember spectral variability also makes traditional spectral unmixing methods perform poorly. Thus, this study proposes a Self-supervised Learning-based Spectral Unmixing algorithm (SLSU) to alleviate the influence of spectral mixing and endmember spectral variability wheat planting area estimation. First, the variational autoencoder is used to achieve unsupervised interpretation of the endmember spectral variability and endmember library generation. Then, abundances corresponding to various endmembers are estimated by using the alternating least-squares strategy and fully constrained least squares. Finally, the unmixed results are corrected on the basis of the spatial neighborhood by using the probabilistic relaxation labeling algorithm to improve the accuracy of spectral unmixing and wheat extraction further. Three typical wheat planting areas in Xinxiang, Henan Province were selected as experimental areas, and the wheat planting area was obtained by Sentinel 2 images. The extraction accuracy was evaluated by wheat distribution data measured in the field, and the results showed that the median value and the R^2 score of the wheat extraction's relative extraction error are close to 1.3 pixels and 1.00, respectively, which are remarkably better than the results extracted by traditional spectral unmixing algorithms (including fully constrained least squares, extended linear mixed model and so on) and traditional supervised learning classification methods (such as support vector machines and random forests). The proposed SLSU algorithm can improve the accuracy and stability of wheat planting area estimation and provide an effective method for crop distribution extraction and planting area estimation.

Key words: remote sensing, self-supervised learning, variational auto encoder, spectral mixing, feature extraction, crop area estimation

Supported by High Resolution Earth Observation Major Special Project (No. 30-H30C01-9004-19/21)