

• 计算机科学与技术 •

DOI:10.12454/j.jsuese.202400048



本刊网刊

## 基于光谱特征的高光谱图像分类的优化方法

李洵枫, 李晓华\*

(四川大学 计算机学院, 四川 成都 610065)

**摘要:** 高光谱图像分类(HSIC)是遥感图像处理领域的重要研究方向之一,其目标是将高光谱图像中的不同地物进行像素级分类。近年来,随着深度学习技术的持续进步,其在高光谱图像分类领域的应用越发突出,已成为主导技术。特别是基于空间-光谱联合特征的HSIC方法的研究进展迅猛,受到了广泛关注。因其能够联合利用空间信息与光谱信息,通常能获得较高的分类精度。然而,基于空间-光谱联合特征的深度学习方法忽略了高光谱图像数据集与自然图像数据集的区别,在提出及实验阶段沿袭了自然图像分类中采用的样本随机划分方式。在现有的公共高光谱图像数据集上采用样本随机划分方式,基于空间-光谱的方法很难避免测试样本与训练样本重叠的问题,这在一定程度上影响了分类精度的真实性并削弱了模型的泛化能力。为了解决这一问题,提出了一种特别的思路,通过缩小基于空间-光谱联合特征方法中的输入图像块大小,实现训练集与测试集的有效分离。当图像块缩小至一定程度时,其特征将逐渐转变为纯光谱特征。考虑到这一事实,重新聚焦基于光谱特征的HSIC方法,提出了一种改进的基于深度像素对特征的方法,用于更快、更稳健地进行高光谱图像分类。所提出的方法从两方面改进了原始的像素对特征方法。首先,提出了基于距离约束的像素对生成方法。该方法不仅可以高效构建大规模的像素对训练集,还可以消除冗余和不合理的像素对。其次,提出了一种基于结果复用的投票策略联合分类方法,可以避免不必要的模型调用,从而显著减少分类时间。在公共数据集的实验结果表明,与现有基于空间-光谱联合特征的方法相比,所提出的方法在实际用例下能够实现更高的分类精度。与标准的基于像素对的方法相比,所提出的方法可以显著降低时间消耗。

**关键词:** 高光谱图像分类;深度学习;像素对特征;样本采取策略

中图分类号: TP391

文献标志码: A

文章编号: 2096-3246(2025)06-0323-12

高光谱图像(HSI)由数百个相邻波段组成,携带丰富的光谱信息,在遥感等诸多应用领域中极为有用。高光谱图像分类(HSIC)旨在基于其光谱或空间-光谱联合特征将每个高光谱像素分配到特定的地物覆盖类别,是高光谱“社区”中最具活力的研究领域之一。HSIC的研究涉及遥感技术、计算机视觉、图像处理等多个学科领域,其应用领域包括但不限于农业、环境监测、城市规划、地质勘探等<sup>[1-3]</sup>。研究者通过开发不同的分类算法和模型,致力于提高HSIC的准确性、效率和适用性。

### 1 国内外研究现状

在过去几十年中,HSIC方法取得了显著进展。依

据所利用的信息类型,现有方法主要可分为两类:基于光谱特征的方法和基于空间-光谱联合特征的方法。

早期的HSIC方法主要依赖高光谱数据丰富的光谱波段信息<sup>[4]</sup>,直接以单个像素的光谱向量作为输入进行分类。这类方法通常将每个像素视为一个高维向量,结合传统机器学习模型完成分类任务。例如,Bandos等<sup>[5]</sup>采用正则化线性判别分析(RLDA)提升分类稳定性;Melgani等<sup>[6]</sup>将支持向量机(SVM)应用于高光谱数据,验证了其在小样本下的有效性。此外,基于独立成分分析<sup>[7]</sup>、多层感知机<sup>[8]</sup>的方法也被提出,均以光谱信息为核心特征。这些方法在传感器空间分辨率较低的时代具有代表性,但难以有效应对光谱变异性和

收稿日期:2024-01-19 修回日期:2024-06-06 网络出版日期:2024-06-12

作者简介:李洵枫(1999-),男,硕士生。研究方向:高光谱图像分类。E-mail:lmom\_xa@163.com

\*通信作者:李晓华,副教授, E-mail:lxhw@scu.edu.cn

类别混淆问题。

随着成像技术的发展, 高光谱传感器的空间分辨率显著提高, 使得邻域像素间的空间结构信息逐渐被重视。研究表明, 融合空间上下文信息可显著提升分类精度<sup>[9-10]</sup>。因此, 研究重点逐步转向结合空间与光谱信息的联合建模范式。例如, Li 等<sup>[11]</sup>通过主动学习框架结合多变量逻辑回归进行半监督分割; Licciardi 等<sup>[12]</sup>利用扩展形态学剖面提取空间特征, 并与非线性主成分分析结合进行分类。这些早期探索为后续空间-光谱联合建模奠定了基础。

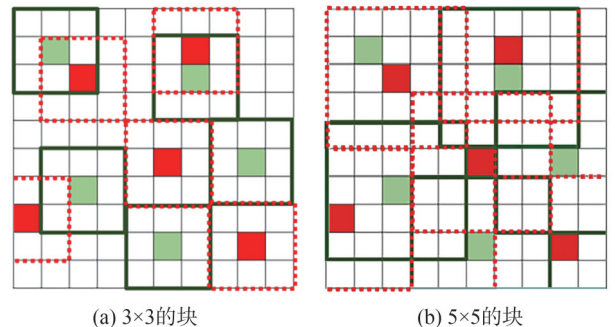
近年来, 深度学习在图像识别、目标检测等计算机视觉任务中取得突破性进展<sup>[13-15]</sup>, 其优势在于能够将特征提取与分类决策集成于统一框架, 实现端到端学习。受此启发, 大量基于深度神经网络的空间-光谱分类方法被提出, 并在 HSIC 任务中展现出卓越性能。

这类方法通常以 3 维图像块(3D patch)为输入, 通过设计专用网络结构从中自动学习判别性空间-光谱联合特征, 进而对中心像素进行分类。例如, Ben 等<sup>[16]</sup>提出了一种纯 3 维卷积神经网络(3D CNN), 首次实现了光谱与空间维度的联合卷积操作; Roy 等<sup>[17]</sup>设计了 HybridSN 架构, 在 3D CNN 提取光谱-空间特征后接 2 维卷积神经网络(2D CNN)进一步挖掘空间细节; Zhong 等<sup>[18]</sup>构建了空间-光谱残差网络(SSRN), 利用残差块缓解深层网络训练难题, 有效学习光谱响应与空间背景特征。

为进一步提升特征表达能力, 后续研究引入了更复杂的结构设计。Paoletti 等<sup>[19]</sup>提出的 PresNet 采用金字塔式残差模块捕获多尺度空间-光谱联合特征。Roy 等<sup>[20]</sup>设计的 A2S2K 网络通过注意力机制增强 3D 残差块的特征选择能力。Zhu 等<sup>[21]</sup>提出的残差空间-光谱注意力网络(RSSAN)可自适应强调同类像素, 抑制干扰信息。Fang 等<sup>[22]</sup>提出利用超像素分割提取光谱-空间信息, 并结合多核学习实现有效分类。Li 等<sup>[23]</sup>进一步探索了像素级与超像素级分类结果的概率融合策略, 提升了边界区域的判别能力。在模型架构创新方面, Transformer 结构因其强大的长程依赖建模能力被引入 HSI 领域。Zhong 等<sup>[24]</sup>提出的光谱-空间变换网络(SSTN)结合空间注意力与光谱关联模块, 克服了传统卷积感受野受限的问题。Song 等<sup>[25]</sup>进一步提出瓶颈空间-光谱变换器(BS2T), 通过压缩-恢复机制高效捕捉全局上下文依赖。Sun 等<sup>[26]</sup>在编码器中融合了特征表示的空间和光谱特征化, 不仅有助于提取局部空间信息, 还能在相邻序列之间建立远程关系。Roy 等<sup>[27]</sup>开发了一种多模态融合变压器(MFT),

用于从 HSI 中提取特征, 并将其与来自光探测和测距(LiDAR)数据的 CLS 令牌融合, 以增强联合分类性能。SpectralFormer<sup>[28]</sup>、HSI-BERT<sup>[29]</sup>、MST-Net<sup>[30]</sup>等模型通过自注意力机制捕捉长距离依赖关系, 在多个基准数据集上取得了优异表现。Gao 等<sup>[31]</sup>提出的 STransFuse 则融合 Swin Transformer 与 CNN, 在遥感图像语义分割任务中展现出了强大潜力, 也为 HSIC 提供了新的架构思路。

几乎所有上述基于空间-光谱联合特征的方法都取得了近乎完美的分类准确率(超过 95%)。然而, 这种高性能是通过使用随机采样策略将数据集随机分割为训练集和测试集来实现的。随机采样策略不符合基于空间-光谱联合特征的 HSIC 的实际用例, 且泛化能力较差<sup>[32]</sup>。在随机抽样中, 由于基于空间-光谱联合特征的方法使用 3 维空间-光谱块作为输入样本进行中心像素分类的特征提取, 存在无意中将测试样本包含在训练集中的可能性。换句话说, 在随机采样策略下, 一些测试样本的部分或全部信息将在训练阶段被“看到”。这种现象被称为训练样本和测试样本之间的重叠问题, 如图 1 所示。当使用 5×5 的块并随机选择 5% 的样本作为训练集时, 测试样本与训练样本的重叠率超过 85%。显然, 重叠问题可能导致虚高的分类精度和较差的泛化能力。原则上, 如果将训练样本的比例保持在 10% 以下, 随着块的缩小, 重叠率将减小。当块缩小到 1×1 时, 变为仅提取光谱特征, 从而形成一种基于光谱特征的方法。显然, 基于光谱特征的方法不受重叠问题的困扰。



注: 绿色中心、绿色实线框表示训练样本; 红色中心、红色虚线框表示测试样本。

图 1 训练样本与测试样本的重叠问题

Fig. 1 Overlap between train samples and test samples

事实上, 在早期就有学者提出了一些基于深度光谱特征的图像分类方法, 这些方法使用了更简单的网络架构, 如自动编码器<sup>[33-34]</sup>、1 维卷积神经网络(1D CNN)<sup>[35-36]</sup>、循环神经网络(RNN)<sup>[37]</sup>等, 将 1 维光谱向量(用于表示 HSI 中的单个像素)映射到地物覆盖

类别。这些方法直截了当且高效,但它们的分类准确率通常较低。

与传统方法不同,Li等<sup>[38]</sup>引入了一种新颖的光谱特征方法,该方法提取深度像素对特征(DPPF),然后通过基于联合分类的投票策略为每个像素分配类别。与传统的基于光谱特征的方法相比,DPPF可以实现更高的分类准确率,而且不像流行的基于空间-光谱联合特征的方法那样受到重叠问题的困扰。尽管基于DPPF的HISC可以取得可靠且更高的分类准确率和良好的泛化能力,但它仍然是一种朴素的方法,面临以下限制:

1)像素对生成方法未考虑所选训练像素之间的差异,而是直接对所有像素进行配对。这种朴素的方法导致了大量像素对的生成,因此在训练过程中的时间和空间消耗显著。此外,重叠问题干扰了模型在训练期间的收敛。

2)在推理/测试过程中,每个像素都被独立处理,忽略了邻接关系的对称性。这将导致大量冗余计算,从而增加了时间消耗。

为了解决上述问题,本文提出了一种更快且更强大的基于DPPF的HISC方法,称为Faster-DPPF。与当前流行的方法专注于设计复杂的空间-光谱联合特征提取模块不同,本文方法从实际用例出发,仅使用光谱信息以实现更快、更实际的HISC。本文的主要贡献可以总结如下:

1)提出了一种基于距离约束的像素对生成方法,不仅能高效构建大规模的像素对训练集,还能消除冗余和不合格的像素对;

2)提出了一种基于投票策略的联合分类重用策略,以避免不必要的模型调用,从而显著减少推理时间;

3)通过对比实验证实基于空间-光谱联合特征的方法在面临训练样本与测试样本的重叠问题时,会导致虚高的分类精度且泛化能力较差;

4)强分离样本选取下的实验结果表明了本文方法的有效性和高效性。

## 2 本文方法

### 2.1 DPPF

DPPF的初衷是通过扩充训练集来解决HSIC中的小样本问题。与一般基于光谱特征的方法直接从单个像素中提取光谱特征不同,DPPF从像素对中提取光谱特征。DPPF的框架主要包括3个步骤:1)使用选定的训练样本生成像素对;2)构建深度卷积神经网络(DCNN)架构提取像素对特征;3)通过

基于联合分类的投票(voting)策略确定测试像素的标签。

#### 2.1.1 像素对生成

假设一个可用的训练集,其中有 $M$ 个带标记的样本,如式(1)所示:

$$T^0 = \left\{ \langle \mathbf{x}_i, y_i \rangle \right\}_{i=1}^M \quad (1)$$

式中: $\mathbf{x}_i \in \mathbf{R}^d$ 为第 $i$ 个单像素样本, $\mathbf{R}^d$ 为 $d$ 维的实数集, $d$ 为光谱维度数目; $y_i \in \{1, 2, \dots, C\}$ 为 $\mathbf{x}_i$ 所对应的类别, $C$ 为类别的数目。

利用像素对生成方法,构造一个新的增强训练集 $T^1$ ,其中每个元素是来自 $T^0$ 的两个训练样本的组合集。

$$T^1 = \left\{ \langle \mathbf{p}_i^k, l_i^k \rangle \right\}_{i,k=1 \text{ 且 } i \neq k}^M \quad (2)$$

$$\mathbf{p}_i^k = [\mathbf{x}_i, \mathbf{x}_k] \quad (3)$$

式(2)、(3)中, $\mathbf{p}_i^k$ 为从 $T^0$ 中选择两个训练样本 $\mathbf{x}_i$ 和 $\mathbf{x}_k$ 的组合像素对, $l_i^k$ 为 $\mathbf{p}_i^k$ 对应的类别。新类别 $l_i^k$ 的定义如下:如果两个样本来自同一类别,则类别不变;如果两个样本来自不同的类别,则类别变为新类别0。用式(4)表示:

$$l_i^k = \begin{cases} y_i, y_i = y_k; \\ 0, \text{其他} \end{cases} \quad (4)$$

需要注意的是,新的增强训练集 $T^1$ 有 $C+1$ 个类别,其中 $\mathbf{p}_i^k \in \mathbf{R}^{d \times 2}$ 。

#### 2.1.2 基于投票策略的联合分类

由于相邻像素往往大概率属于同一类,因此在测试阶段构建基于投票策略的联合分类,用于获取测试像素 $\mathbf{x}_{ij}$ 的最终标签 $\hat{y}_{ij}$ ,可表示为式(5):

$$\hat{y}_{ij} = D(\hat{l}_{ij}^k) = D\left(N(\mathbf{p}_{ij}^k)\right) = D\left(N(\mathbf{x}_{ij}, \mathbf{x}_{ij}^k)\right), \quad (5)$$

$$k = 1, 2, \dots, n^2 - 1$$

$$\mathbf{p}_{ij}^k = [\mathbf{x}_{ij}, \mathbf{x}_{ij}^k] \quad (6)$$

式(5)、(6)中: $\mathbf{x}_{ij}$ 为位于HSI第 $i$ 行、第 $j$ 列的测试像素; $\mathbf{x}_{ij}^k$ 为在 $n \times n$ 邻域中的第 $k$ 个相邻像素; $\mathbf{p}_{ij}^k$ 为 $\mathbf{x}_{ij}$ 和 $\mathbf{x}_{ij}^k$ 生成的像素对; $N(\cdot)$ 为训练后的DCNN,它可以为像素对 $\mathbf{p}_{ij}^k$ 分配一个预测类别 $\hat{l}_{ij}^k$ ; $D(\cdot)$ 为投票策略,它决定了 $\hat{l}_{ij}^k (k=1, 2, \dots, n^2-1)$ 中最频繁出现的标签是 $\mathbf{x}_{ij}$ 最终标签 $\hat{y}_{ij}$ 。

图2为DPPF框架。在训练阶段生成像素对,并对DCNN进行训练,以提取像素对特征并对每个像素对进行分类。在测试阶段使用周围样本生成像素对,将其输入到已训练的DCNN中以获取每个像素对的标签,通过投票策略获取中心像素的标签。

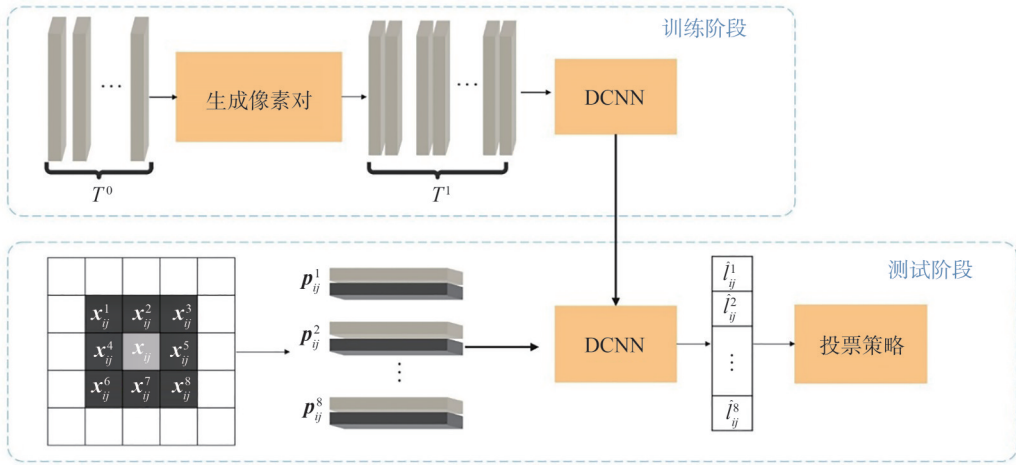


图2 DPPF框架

Fig. 2 Framework of DPPF

## 2.2 Faster-DPPF

与基于空间频谱特征的HSIC方法相比,DPPF在测试阶段具有更可靠的分类结果,在实际用例中不存在重叠问题,具有更强的泛化能力。然而,DPPF的实现虽然很简单,但需要大量的时间与空间消耗,这限制了DPPF的可用性。本节通过优化像素对生成方法和基于投票的联合分类模块,重点改进训练阶段的时间与空间消耗和测试阶段的时间消耗,并保持分类精度基本不变。对于DCNN模块,Faster-DPPF使用与DPPF相同的架构。

### 2.2.1 基于距离约束的像素配对方法

从DPPF的算法可以看出,普通像素对生成方法采用穷举方案。虽然它可以极大地扩充训练集,但也带来了大量的冗余。假设原始训练集 $T^0$ 有 $M$ 个单像素样本,则会生成 $[M \times (M-1)]$ 个像素对。再假设 $m_c$ 表示第 $c(c=1,2,\dots,C)$ 类像素对中可用的训练样本数量,则 $M = \sum_{c=1}^C m_c$ 。对于第 $c$ 类像素对的数量(记为 $n_c$ ),有 $n_c =$

$P_2^{m_c} = m_c \times (m_c - 1)$ 。即使为了保持数据平衡,对于第0类样本只保留一个近似相等的像素对数量(记为 $n_0$ ),在 $T^1$ 中总像素对的数量仍然非常庞大,达到 $n_0 + \sum_{c=1}^C n_c$ 。庞大的训练样本规模将导致训练过程中大量的时间与空间消耗。通常认为生成如此多的像素对是没有必要的,特别是使用两个空间距离非常远的像素来生成像素对。实际上,HSI通常具有较低的分辨率,一个像素往往覆盖几平方米甚至几百平方米的地面。即使是属于同一类的土地,当两个像素点的空间距离很远时,传感器捕捉到的光谱曲线也会有很大的差异。使用这种训练集中生成的像素对对网络模型进行训练,不仅增加了时间与空间的消耗,还会误导网络,使其提取到不相关

的特征,从而导致测试阶段的分类准确率降低。

考虑到上述情况,本文提出了一种基于距离约束的像素对生成方法。对于在 $T^0$ 中带标记的单像素点样本 $\langle x_i, y_i \rangle$ ,假设 $y_i = c$ ,用 $A_i$ 表示属于第 $c$ 类的其他 $m_c - 1$ 个单像素样本,并根据其在HSI中与 $x_i$ 的空间距离排序:

$$A_i = \left\{ \left\langle x_i^k, c \right\rangle \right\}_{k=1}^{m_c-1} \quad (7)$$

式中, $x_i^k$ 是第 $k$ 个最接近 $x_i$ 的像素点。在 $A_i$ 中,只有最接近 $x_i$ 的 $K(K < m_c)$ 个单像素样本才能与 $x_i$ 组合以生成像素对。因此,优化后的训练集表示为:

$$T^{10} = \left\{ \left\langle p_i^k, l_i^k \right\rangle \right\}_{i=1, k=1}^{i=M, k=K} \quad (8)$$

$$p_i^k = (x_i, x_i^k), i = 1, 2, \dots, M; k = 1, 2, \dots, K \quad (9)$$

$$l_i^k = y_i \quad (10)$$

除了式(8)~(10)中描述的同类别的像素对外,对于第0类样本,从不同的类中随机选择适量的像素对,生成近似相等数量的像素对(记为 $n'_0$ ),并将其加入到 $T^{10}$ 中。显然, $T^{10}$ 中的像素对总数将是 $n'_0 + \sum_{c=1}^C K m_c$ 。当 $K \ll m_c$ 时, $T^{10}$ 中的元素数量远少于 $T^1$ ,因此可以显著减少训练过程中的时间与空间消耗。

### 2.2.2 基于结果复用的分类方法

如第2.1节所述,为了在HSI中对像素进行分类,训练好的DCNN模型必须运算 $n^2 - 1$ 次,之后再使用投票策略来确定像素的类别标签。给定大小为 $W \times H$ 的HSI,训练好的模型必须运算 $W \times H \times (n^2 - 1)$ 次,因为DPPF使用逐像素处理方式处理真实的HSI,这意味着从上到下、从左到右逐个分类所有像素。很明显,整个过程的预测时间会随着相邻窗口的增大而迅速增加。虽然在DCNN中主要使用1维卷积,但在预测阶段的时间消耗可能超过基于3D CNN的方法。考虑到在分

类相邻像素时存在大量重复调用 DCNN 模型的情况,本文提出了一个显著的改进。例如,若邻域窗口的大小为 3,则像素  $x_{ij}$  对应的像素对  $(x_{ij}, x_{ij}^2)$  通过网络输出的结果与像素  $x_{i-1,j}$  对应的像素对  $(x_{i-1,j}, x_{i-1,j}^2)$  应当相同。假设当前像素为  $x_{ij}$ ,则不需要将所有  $n^2-1$  个像素对输入到 DCNN 模型中,只需将其中的一半输入到 DCNN 模型中即可。另一半像素对的分类结果可以直接重用前一像素对应的像素对的分类结果。其核心思想描述如下。

以  $3 \times 3$  邻域为例,对于当前像素  $x_{ij}$ ,训练好的 DCNN 只需要运算 4 次,就可以获得 4 个像素对  $\{p_{ij}^5, p_{ij}^6, p_{ij}^7, p_{ij}^8\}$  的预测标签  $\{\hat{l}_{ij}^5, \hat{l}_{ij}^6, \hat{l}_{ij}^7, \hat{l}_{ij}^8\}$ 。而其与前 4 个像素对  $\{p_{ij}^1, p_{ij}^2, p_{ij}^3, p_{ij}^4\}$  对应的预测标签  $\{\hat{l}_{ij}^1, \hat{l}_{ij}^2, \hat{l}_{ij}^3, \hat{l}_{ij}^4\}$ ,则

直接重用像素对  $\{p_{i-1,j-1}^8, p_{i-1,j}^7, p_{i-1,j+1}^6, p_{i,j-1}^5\}$  的预测标签  $\{\hat{l}_{i-1,j-1}^8, \hat{l}_{i-1,j}^7, \hat{l}_{i-1,j+1}^6, \hat{l}_{i,j-1}^5\}$ 。它们分别从对之前像素  $x_{i-1,j-1}, x_{i-1,j}, x_{i-1,j+1}, x_{i,j-1}$  的调用中获取,因为这些像素对与像素对  $\{p_{ij}^1, p_{ij}^2, p_{ij}^3, p_{ij}^4\}$  相同。

显然,使用重用策略可以减少 50% 的模型调用次数,因此整个 HSI 的预测时间将减少约 50%。

### 2.2 网络结构

像素对方法的核心在于深入剖析样本配对带来的深远效应。为探究此影响,本文巧妙地设计了一个简洁高效的深度卷积神经网络模型,其网络结构如图 3 所示。这一模型汇聚了卷积网络、跳变连接以及全连接神经网络的精髓,形成了一个功能卓越的综合体系。

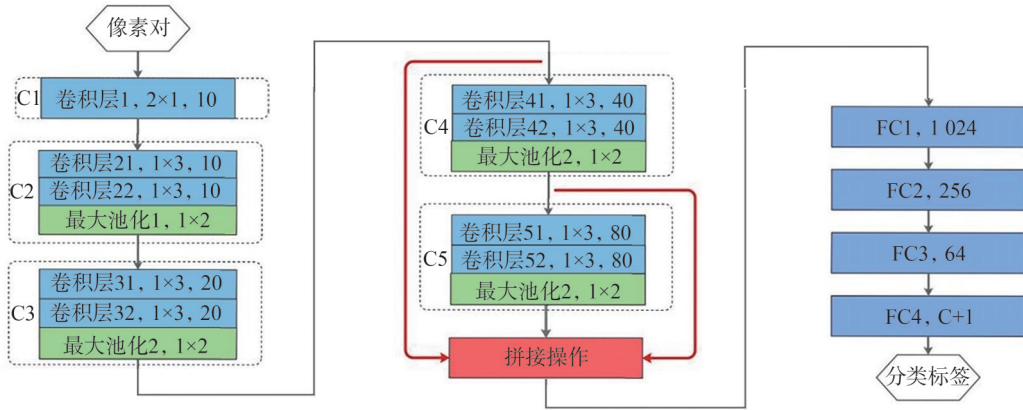


图 3 网络结构示意图

Fig. 3 Schematic diagram of network structure

在模型的架构搭建过程中,独具匠心地在每个卷积层(conv)之后融入了批归一化与ReLU操作,这一策略显著增强了模型的稳定性并提升了其整体性能。

随着输入数据历经多层卷积处理的流转,从 C1~C5 的每一层级都精准地捕捉了不同维度的特征信息。值得注意的是,C3、C4 与 C5 的输出并非孤立存在,而是经过 1 维向量化转换后,通过精细的拼接操作(flatten+concat)实现了高效整合。这一设计策略极大地增强了模型对多层次特征信息的利用效率,进而显著提升了模型的表达与泛化能力。

最终,经过精心整合的特征信息被传递至全连接层 FC1~FC4 进行进一步处理。通过这一系列对全连接层的精细调控,模型得以输出精确可靠的分类结果。

## 3 实验结果

### 3.1 数据集描述

本文方法采用 3 个公共的高光谱图像分类数据集来评估提出的 Faster-DPPF 方法的有效性和效率,分

别是印第安纳松木数据集 (Indiana Pines dataset, IP 数据集)、萨利纳斯谷数据集 (Salinas Valley dataset, SV 数据集) 和肯尼迪航天中心数据集 (Kennedy Space Centre dataset, KSC 数据集)。

IP 数据集:该数据集是由 AVIRIS 传感器在美国印第安纳州西北部的印第安纳松木试验地上采集的,包括一张  $145 \times 145$  像素的图片,含 224 个光谱波段,波长范围为  $400 \sim 2\,500$  nm。需要注意的是,尽管该数据集有 16 个类别,但在本文中仅使用了 9 个类别,以避免一些类别训练样本非常有限的问题。同时,通过去除覆盖水吸收区域的波段,将波段数量减少到 200。

SV 数据集:该数据集由 224 波段的 AVIRIS 传感器在美国加利福尼亚州的萨利纳斯谷上采集,具有高空间分辨率。SV 数据集包括一张  $512 \times 217$  像素的图片,有 204 个光谱波段,地面实况包含 16 个类别的地物覆盖。

KSC 数据集:肯尼迪航天中心数据集是由 AVIRIS 传感器在 1996 年采集的,波长范围为  $400 \sim 2\,500$  nm。

这些图像的空间维度为  $512 \times 614$  像素,在去除一些信噪比(SNR)较低的光谱波段后,共有 176 个光谱波段。KSC 数据集包含了 13 个类别的总共 5 202 个样本。

### 3.2 实验参数

1)在生成训练样本的过程中,本实验考虑了 3 种抽样策略:随机抽样策略、弱分离抽样策略和强分离抽样策略。对于使用随机抽样策略的实验,从整个高光谱图像中随机选择每个类别的 200 个样本作为训练集,其余的样本按照 9:1 的比例划分为测试集和验证集。对于使用弱分离采样策略的实验,首先将地面实况(GT)中的每个连通分量分为两个空间不相交的分量:一个用于训练,另一个用于测试;随后从训练子块中随机选择每个类别的 200 个样本作为训练集,而测试子块中的样本按照 9:1 的比例划分为测试集和验证集。对于使用强分离采样策略的实验,首先通过从左到右和从上到下的方式扫描每个类别,将每个类别中的样本均等划分为两个空间不相交的部分,一个用于训练,另一个用于测试;然后,从训练子块中随机选择每个类别的 200 个样本作为训练集,而测试子块中的样本按照 9:1 的比例划分为测试集和验证集。IP 数据集上的弱分离与强分离抽样策略的样本划分如图 4 所示。值得注意的是,由于 KSC 数据集中标记样本有限,每个类别仅选择了 20 个训练样本。

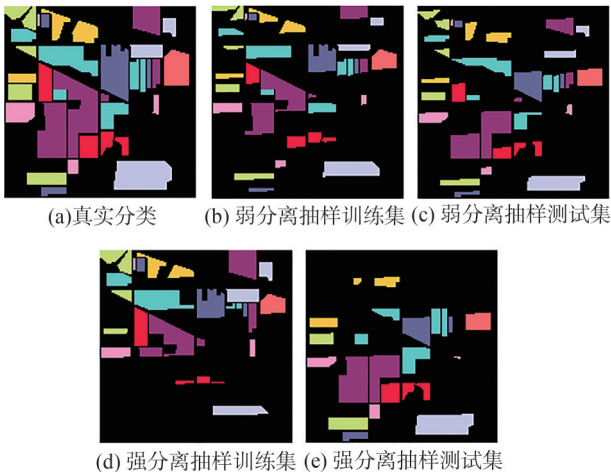


图 4 IP 数据集上训练集与测试集的划分

Fig. 4 Train and test splits on the IP dataset

2)除了 DPPF 之外,本实验还将 Faster-DPPF 与 4 种目前最先进的基于空间-光谱联合特征的方法(3D CNN<sup>[16]</sup>、SSRN<sup>[18]</sup>、SSTN<sup>[24]</sup>、BS2T<sup>[25]</sup>),以及两种典型的基于光谱的方法(1D CNN<sup>[16]</sup>、RNN<sup>[17]</sup>)进行了比较。对于基于空间-光谱联合特征的方法,使用  $5 \times 5$  大小的块作为模型输入;而对于基于光谱的方法,则使用单个像素的光谱向量作为模型输入。对于 DPPF 和 Faster-DPPF,在基于投票的联合分类阶段使用了  $5 \times 5$  大小的

邻域。

3)选取整体准确率(overall accuracy, OA)、平均准确率(average accuracy, AA)、Kappa 系数、训练时间和整个 HSI 的测试时间作为定量评估指标。所有涉及的模型都独立实验 10 次,实验结果取 10 次实验所得的平均值。

4)所有实验在一台搭载 AMD Ryzen™ 5 5600X CPU、16 GB 内存和 NVIDIA GeForce RTX™ 3060 GPU 的计算机上运行,平台为 Pytorch 框架的 Python 3.7,采用 Adam 优化器进行反向传播。批量(batch)的大小为 128,学习率为 0.001。训练周期设置为 200,每 5 个周期进行验证。

### 3.3 强分离抽样策略下的实验结果

正如第 1 节中讨论的,采用随机抽样策略时,对于基于空间-光谱联合特征的方法,测试样本与训练样本存在严重重叠问题,因此通常会得到虚高的分类准确率。为确保公正性,本节使用强分离抽样策略进行实验。对于基于空间-光谱联合特征的方法,这种策略可以确保测试集中只有极少数像素会无意中被包含在训练集中。

表 1 为 IP 数据集上不同方法的实验结果。由表 1 可知,本文提出的 Faster-DPPF 方法在所有方法中实现了最佳准确率,OA 为 87.76%,AA 为 88.21%,Kappa 系数为 85.51。与 BS2T(基于空间-光谱联合特征的方法中最好的)相比,本文方法在 OA 上实现了 7.28 个百分点的显著提升,Kappa 系数提高了 8.79;与 RNN(传统基于光谱特征的方法中表现较好的)相比,本文方法 OA 提高了 9.5 个百分点,Kappa 系数提高了 11.15。这都证明了基于深度像素对特征方法的有效性。与 DPPF 相比,本文方法的训练时间约为其 25.3%,测试时间约为其 54.2%,且分类精度略有提高。

此外,值得注意的是,本文提出的 Faster-DPPF 方法在大多数类别中均实现了最高准确率;而使用 DPPF 方法时每个类别的准确率也相对较高,不同类别之间没有明显差异。这表明在 DPPF 框架内,整体分类性能更加稳定。

图 5 为 IP 数据集每种方法的可视化分类图。由图 5(c)~(f)可见,在训练集与测试集重叠很少的情况下,基于空间-光谱联合特征的方法倾向于错误分类测试集中的像素(用绿色边界标记的区域)。而由图 5(g)~(j)可见,基于光谱特征的方法还是能取得相对一致的分类结果。此外,与传统的基于光谱特征的方法 1D CNN 和 RNN(图 5(g)~(h))相比,DPPF 和 Faster-DPPF 的分类图(图 5(i)~(j)),具有较少的椒盐噪声。

表 1 IP数据集上不同方法的实验结果  
Tab. 1 Quantitative results of different methods on the IP dataset

类别编号	各类别整体准确率/%							
	3D CNN	SSRN	SSTN	BS2T	1D CNN	RNN	DPPF	Faster-DPPF
1	39.57±5.71	44.38±6.62	63.66±7.27	61.18±5.13	63.26±6.97	66.77±5.47	78.26±3.29	80.68±2.73*
2	31.47±3.22	27.16±2.37	40.46±4.26	38.90±2.93	52.68±4.92	55.72±2.77	63.75±2.49	64.27±1.12*
3	83.70±1.82	88.15±5.71	91.56±5.21	91.36±2.97	88.19±2.95	94.42±1.42	96.48±0.75*	94.71±1.73
4	96.86±1.05	98.31±0.69	98.55±1.13	99.26±0.44*	96.71±0.78	95.95±1.12	99.18±0.15	99.14±0.39
5	100.00±0.00*	99.92±0.17	100.00±0.00*	100.00±0.00*	99.58±0.71	100.00±0.00*	100.00±0.00*	100.00±0.00*
6	57.37±6.03	71.81±2.57	72.77±5.07	76.30±3.40	66.65±7.10	69.43±4.41	85.65±2.13	86.28±2.27*
7	78.27±5.86	91.12±3.06*	87.88±3.32	92.80±3.18	74.19±7.42	74.15±3.37	88.81±2.86	88.96±2.54
8	31.04±9.07	46.78±5.19	57.35±10.28	51.75±1.83	66.48±5.66	71.99±4.32	81.17±5.50*	80.09±4.10
9	100.00±0.00*	99.98±0.05	99.97±0.06	100.00±0.00*	98.60±0.47	99.18±0.58	99.72±0.14	99.78±0.09
评价指标	3D CNN	SSRN	SSTN	BS2T	1D CNN	RNN	DPPF	Faster-DPPF
整体准确率/%	68.70±1.77	75.34±0.94	79.64±0.61	80.48±0.69	76.44±2.99	78.26±0.73	87.39±0.71	87.76±0.66*
平均准确率/%	68.70±1.64	74.18±1.27	79.13±1.17	79.06±0.40	78.48±2.35	80.84±0.47	88.11±0.45	88.21±0.44*
Kappa 系数	62.52±2.10	70.49±1.12	75.78±0.78	76.72±0.78	72.17±3.48	74.36±0.84	85.06±0.82	85.51±0.76*
训练时间/s	419.05	86.00	118.83	452.77	53.50*	81.63	1854.26	469.55
测试时间/s	49.23	20.19	3.71	73.07	1.04*	2.32	31.37	17.00

注: \*为最佳结果。

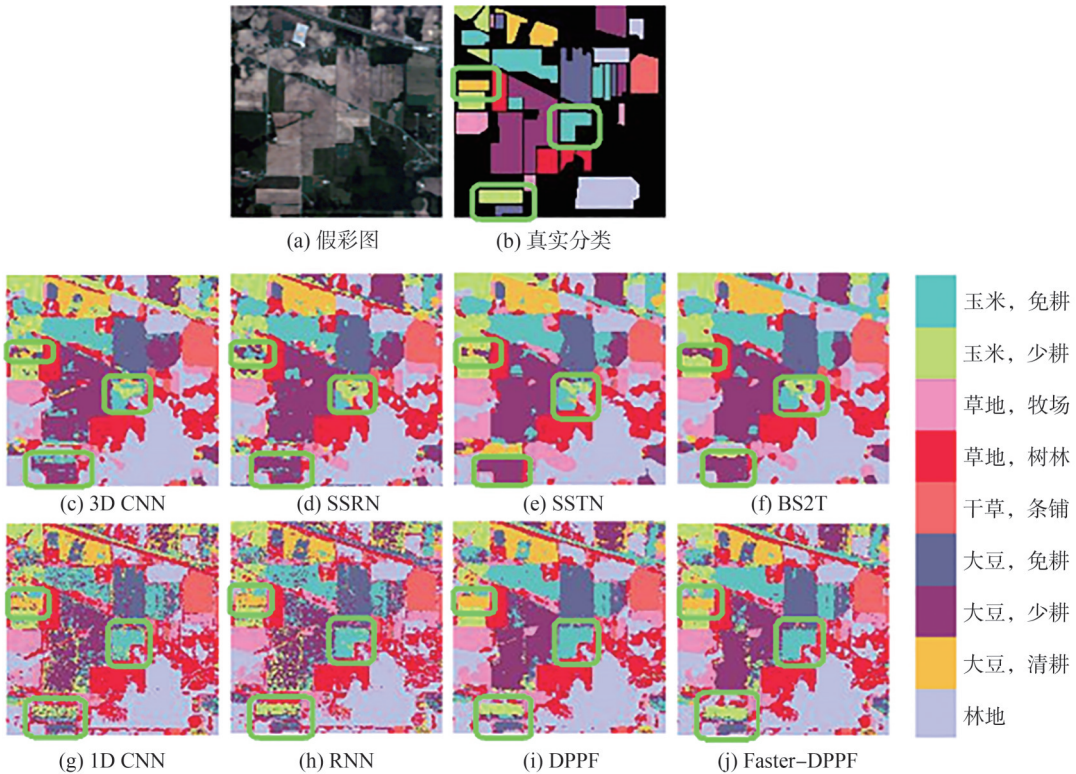


图 5 IP数据集上不同方法的可视化分类图展示

Fig. 5 Classification maps of the different approaches on IP dataset

表 2 和 3 分别为在 SV 和 KSC 数据集上的不同方法的实验结果。类似于 IP 数据集上的实验结果, 本文提出的 Faster-DPPF 方法在 SV 和 KSC 数据集上都有最高的 OA、AA、Kappa 系数和相对较少的训练时间与测试时

间。由表 2 和 3 可见, 尽管只有少量类别获得了最佳准确率(SV 数据集中有 3 个, KSC 数据集中有 1 个), 但本文提出的 Faster-DPPF 对每个类别的识别准确率均相对较高。这表明 DPPF 框架的分类性能较其他方法更加稳定。

表 2 在 SV 数据集上的不同方法的实验结果

Tab. 2 Quantitative results of different methods on the SV dataset

类别编号	各类别整体准确率/%							
	3D CNN	SSRN	SSTN	BS2T	1D CNN	RNN	DPPF	Faster-DPPF
1	96.71±1.08	99.94±0.13	99.90±0.30	100.00±0.00*	99.98±0.04	99.72±0.48	99.66±0.21	99.58±0.28
2	99.94±0.14	100.00±0.00*	100.00±0.00*	100.00±0.00*	99.84±0.04	99.68±0.41	99.75±0.19	99.87±0.07
3	68.42±2.33	62.72±3.27	59.29±4.08	59.39±3.37	59.59±4.94	48.63±5.77	71.86±2.23	72.80±2.83*
4	98.73±0.75	99.34±0.31	98.78±0.89	99.57±0.26*	99.06±0.64	99.23±0.27	98.90±1.13	98.84±0.87
5	72.90±1.84	72.22±3.73	83.72±3.48	70.47±3.49	67.55±2.05	78.06±5.25	83.05±0.96	83.29±0.81*
6	100.00±0.00*	100.00±0.00*	100.00±0.00*	100.00±0.00*	99.09±0.27	99.69±0.15	99.81±0.16	99.72±0.16
7	99.92±0.04	99.86±0.10	99.78±0.20	99.96±0.03*	99.28±0.30	99.44±0.14	99.78±0.11	99.75±0.10
8	65.23±2.77	81.04±2.03	78.45±3.71	71.77±2.42	73.51±4.08	80.32±3.57	78.73±2.12	81.68±1.30*
9	99.17±0.24	99.48±0.48	99.95±0.11	100.00±0.00*	96.34±0.85	98.56±0.29	98.26±0.88	98.24±0.81
10	93.41±1.52	95.69±1.28	95.40±2.73	97.21±0.50*	93.09±2.49	93.26±1.82	93.35±1.89	92.64±2.11
11	97.74±0.71	99.39±0.25*	98.32±1.60	98.79±0.61	96.16±1.43	97.02±1.20	96.57±1.75	97.25±1.40
12	99.44±0.89	100.00±0.00*	99.74±0.31	99.99±0.03	97.42±1.64	98.06±3.07	97.84±1.41	97.23±1.12
13	100.00±0.00*	99.91±0.26	100.00±0.00*	100.00±0.00*	99.04±1.03	99.98±0.07	98.18±1.55	97.13±2.13
14	99.07±0.76	99.81±0.33	100.00±0.00*	100.00±0.00*	99.61±0.45	99.55±0.28	97.87±1.43	97.48±1.16
15	63.78±2.37	64.22±3.33	62.27±2.97	67.52±3.04	47.42±2.18	38.27±4.87	73.79±3.35*	71.48±2.14
16	98.13±0.34	97.72±0.51	98.48±0.85	98.67±0.45*	96.84±0.74	97.76±0.16	97.34±1.15	97.86±0.45
评价指标	3D CNN	SSRN	SSTN	BS2T	1D CNN	RNN	DPPF	Faster-DPPF
整体准确率/%	84.60±1.03	88.07±0.95	87.73±1.44	86.57±0.84	83.06±1.05	83.76±1.13	89.23±0.85	89.40±1.17*
平均准确率/%	90.79±0.52	91.96±0.67	92.13±1.21	91.46±0.62	88.99±1.04	89.20±0.92	92.79±0.81	92.80±0.97*
Kappa 系数	82.90±1.13	86.72±1.05	86.34±1.57	85.07±0.93	81.13±1.24	81.88±1.26	88.02±0.94	88.19±1.30*
训练时间/s	762.86	157.05	215.26	812.06	100.03*	150.32	3218.45	810.36
测试时间/s	262.94	112.37	19.59	377.30	5.16*	12.36	163.19	88.36

注: \*为最佳结果。

表 3 在 KSC 数据集上的不同方法的实验结果

Tab. 3 Quantitative results of different methods on the KSC dataset

类别编号	各类别整体准确率/%							
	3D CNN	SSRN	SSTN	BS2T	1D CNN	RNN	DPPF	Faster-DPPF
1	93.14±2.21	94.14±2.34	96.76±1.51*	95.37±1.74	86.39±2.93	92.15±3.04	91.27±5.46	90.81±4.05
2	33.82±4.99	38.38±5.05	43.33±9.06	41.63±6.01	65.85±7.23	66.83±3.90	75.20±5.53*	72.38±4.63
3	72.79±6.65	87.68±8.24	95.04±4.13*	88.29±11.10	86.92±14.25	73.46±15.69	88.60±5.99	85.98±7.43
4	13.47±4.86	21.50±11.37	18.11±13.73	47.01±15.42	50.39±16.23	52.68±8.88*	41.85±12.49	47.20±11.57
5	25.49±10.40	58.17±8.09	59.03±6.53	58.29±5.71	65.43±21.11	61.24±10.05	71.84±7.99	76.24±2.84*
6	69.31±7.36	77.76±14.90	71.04±13.46	81.55±11.67*	59.13±7.65	59.22±2.47	69.90±4.46	73.11±3.88
7	76.30±10.73	81.48±6.25	90.74±11.14	95.37±8.12*	54.72±21.13	61.13±9.13	76.47±8.40	81.53±8.88
8	14.15±4.02	17.05±0.21	17.05±0.00	17.05±0.00	57.60±1.09*	57.42±1.83	55.23±2.70	54.55±1.58
9	73.18±8.10	73.79±9.88	80.57±13.40*	80.46±6.62	64.75±8.81	76.63±7.11	77.44±5.47	77.24±4.46
10	86.26±4.03	90.69±3.05	90.89±5.28	95.37±3.05*	92.12±3.50	92.17±3.44	88.33±8.18	90.16±7.82
11	100.00±0.00*	100.00±0.00*	100.00±0.00*	100.00±0.00*	99.52±1.43	99.00±0.75	99.83±0.20	99.88±0.26
12	73.24±5.47	78.58±5.03	81.78±5.54	90.04±4.24*	76.68±5.10	82.93±3.95	86.29±2.57	85.59±2.99
13	99.70±0.63	98.73±2.01	99.70±0.59	100.00±0.00*	100.00±0.53	99.63±0.52	99.30±1.93	99.93±0.16
评价指标	3D CNN	SSRN	SSTN	BS2T	1D CNN	RNN	DPPF	Faster-DPPF
整体准确率/%	72.73±0.93	76.69±1.14	78.60±1.53	81.11±0.94	79.65±1.64	81.65±0.89	83.45±1.80	83.85±1.28*
平均准确率/%	63.91±1.11	70.61±1.23	72.62±1.59	76.19±0.90	73.81±2.73	74.96±0.77	78.58±2.26	79.58±1.72*
Kappa 系数	69.62±1.02	74.06±1.26	76.16±1.69	78.96±1.04	77.39±1.82	79.57±0.98	81.54±1.99	81.98±1.42*
训练时间/s	39.41	11.86	19.89	67.70	8.02*	10.43	25.46	14.91
测试时间/s	613.48	274.00	53.66	947.22	13.06*	29.92	405.84	214.26

注: \*为最佳结果。

此外,由表3可见,某些类别的分类准确率存在较大的标准差,有些标准差甚至达到了10。这是由于KSC数据集中样本量较小,每个类别仅有20个样本。

SV和KSC数据集上不同方法的可视化分类图如附录A图A1和A2所示。

### 3.4 不同样本划分方式带来的影响

对于基于空间-光谱联合特征的方法,由于使用了强分离抽样策略,得到的实验结果明显低于文献中报告的结果。但这些较差的结果比通过随机抽样策略获得的结果更可信,因为在随机抽样策略中,大部分测试样本会无意地被包含在训练集中。为了探讨不同抽样策略之间的分类准确率差距,以IP数据集为例,进行了3种抽样策略下的实验,并对结果进行了分析。

实际上,这3种抽样策略对应于训练集和测试集间不同的空间重叠度。例如,在随机抽样策略下,由于训练集和测试集都来自整个图像,两者覆盖的空间范围几乎完全重叠。而在强分离抽样策略下,训练集的空间范围主要集中在图像的上半部分,而测试集的空间范围主要集中在图像的下半部分,如图4(d)~(e)所示。对于基于空间-光谱联合特征的方法,不同的抽样策略还意味着不同的重叠率(overlap ratio),重叠率定义为与训练样本有重叠的测试样本占总测试样本的百分比。3种抽样策略下的实验结果如表4所示。

表4 3种抽样策略下的实验结果

Tab. 4 Experimental results under three evaluation strategies

方法	随机抽样策略		弱分离抽样策略		强分离抽样策略	
	重叠率/%	整体准确率/%	重叠率/%	整体准确率/%	重叠率/%	整体准确率/%
3D CNN	99.05	94.40	39.48	85.07	22.13	68.70
SSRN	99.05	98.99	39.48	91.60	22.13	75.34
SSTN	99.05	98.77	39.48	92.92	22.13	79.64
BS2T	99.05	99.48	39.48	95.03	22.13	80.48
1D CNN	0	77.38	0	76.44	0	75.70
RNN	0	79.66	0	78.26	0	76.81
Faster-DPPF	0	94.49	0	91.07	0	87.76

图6为3种抽样策略下不同方法OA的折线图。综合表4与图6可见,对于所有涉及的方法,OA通常随重叠率的降低而下降;基于空间-光谱联合特征的方法(3D CNN、SSRN、SSTN、BS2T)的OA随重叠率的下降而迅速下降,而基于光谱特征的方法(1D CNN、RNN、Faster-DPPF)的OA下降速度较慢。实验结果表明,基于空间-光谱联合特征的方法在使用随机抽样策略时,高重叠率可能导致虚高的分类精度和较差的泛化能力。

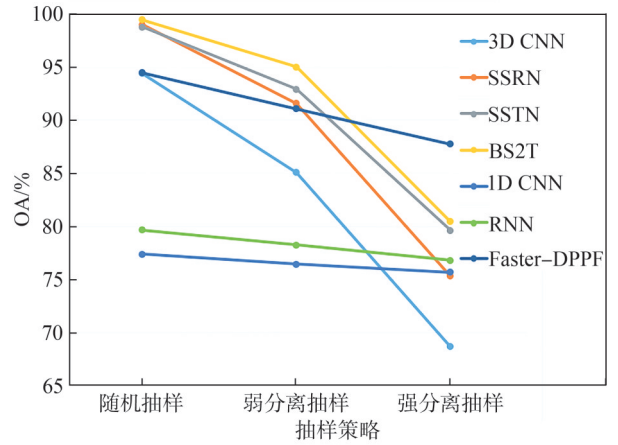


图6 3种抽样策略下不同方法的OA折线图

Fig. 6 Line graph of experimental results under three evaluation strategies

### 3.5 消融实验

本节进行了不同方案的消融实验,以评估本文的两个创新对最终性能的贡献。实验方案包括:1)没有任何优化的原始DPPF;2)仅使用距离约束的DPPF-1;3)仅使用重用策略的DPPF-2;4)最终的Faster-DPPF,包含了两种优化。

以IP数据集为例,表5为消融实验结果。由表5可见,通过应用距离约束和重用策略可以显著改善性能。具体来说,这两个创新从不同的角度改进了原始DPPF。首先,距离约束可以显著减少训练时间,从1 854.26 s减少到465.71 s。然而OA轻微增加而不是减少。这表明,距离约束不仅消除了冗余的像素对,还减轻了无关信息的影响。其次,重用策略可以进一步减少整个图像的测试时间,从31.37 s减少到16.92 s,同时OA几乎保持不变。

表5 消融实验结果

Tab. 5 Results of ablation experiments

实验方案	整体准确率/%	训练时间/s	测试时间/s
原始DPPF	87.39	1 854.26	31.37
DPPF-1	87.65	465.71	31.86
DPPF-2	87.44	1 819.62	16.92
Faster-DPPF	87.76	469.55	17.00

### 3.6 参数K对性能的影响

Faster-DPPF的性能受到基于距离约束的像素对生成器中使用的参数K的影响。实验过程中,Faster-DPPF的训练过程与DPPF一致。本节以IP数据集为例,评估参数K对分类准确率和训练时间的影响。参数K设置为1、3、5、10、20、50、100和199。图7为参数K对OA和训练时间的影响。由图7(a)可见,随着K的增加,OA一开始呈上升趋势,但在K=5之后趋于稳定,并在K=50时达到峰值,由图7(b)可见,随着K的增加,训练时

间逐渐增加。这证实了在训练集中不需要配对所有像素,尤其是那些空间距离较远的两个像素。

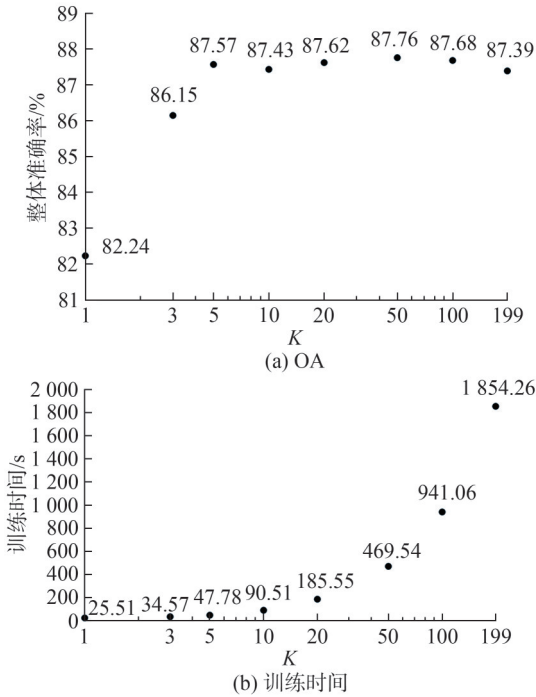


图7  $K$ 值对OA和训练时间的影响

Fig. 7 Influence of parameter  $K$  on OA and train time

## 4 结 论

本文指出并确认了被空间-光谱联合特征方法长期忽视的重叠问题,它可能导致分类精度虚高。在此基础上,重新审视了基于光谱特征的方法,并提出了一种鲁棒的高光谱图像分类方法Faster-DPPF。在Faster-DPPF中,为了减少训练时间并消除冗余和不合格的像素对,引入了基于距离约束的像素对生成方法。此外,考虑到邻接关系的对称性质,在基于投票策略的联合分类中采用了重用策略,该策略可以避免不必要的模型调用,因此显著减少了分类时间。在公共数据集上的实验结果表明,与目前先进的基于深度学习的HSIC方法相比,本文方法具有明显的优越性。

在未来的工作中,将从两个方向处理实际用例中的HSIC任务:一是探索更适用的深度网络来提取光谱特征,二是研究基于深度空间-光谱联合特征的少样本分类在高光谱数据上的应用。

附录见本刊网络版,扫描标题旁的二维码可阅读网络全文。

### 参考文献:

[1] Landgrebe D. Hyperspectral image data analysis[J]. IEEE Signal Processing Magazine, 2002, 19(1): 17-28.  
 [2] Malthus T J, Mumby P J. Remote sensing of the coastal zone: An overview and priorities for future research[J]. In-

ternational Journal of Remote Sensing, 2003, 24(13): 2805-2815.

- [3] Bioucas-Dias J M, Plaza A, Camps-Valls G, et al. Hyperspectral remote sensing data analysis and future challenges [J]. IEEE Geoscience and Remote Sensing Magazine, 2013, 1(2): 6-36.  
 [4] Du Bo, Zhang Liangpei. Random-selection-based anomaly detector for hyperspectral imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 49(5): 1578-1589.  
 [5] Bandos T V, Bruzzone L, Camps-Valls G. Classification of hyperspectral images with regularized linear discriminant analysis[J]. IEEE Transactions on Geoscience and Remote Sensing, 2009, 47(3): 862-873.  
 [6] Melgani F, Bruzzone L. Classification of hyperspectral remote sensing images with support vector machines[J]. IEEE Transactions on Geoscience and Remote Sensing, 2004, 42(8): 1778-1790.  
 [7] Villa A, Benediktsson J A, Chanussot J, et al. Hyperspectral image classification with independent component discriminant analysis[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011, 49(12): 4865-4876.  
 [8] Zhong Yanfei, Zhang Liangpei. An adaptive artificial immune network for supervised classification of multi-/hyperspectral remote sensing imagery[J]. IEEE Transactions on Geoscience and Remote Sensing, 2012, 50(3): 894-909.  
 [9] Audebert N, Le Saux B, Lefevre S. Deep learning for classification of hyperspectral data: A comparative review[J]. IEEE Geoscience and Remote Sensing Magazine, 2019, 7(2): 159-173.  
 [10] Ghamisi P, Maggiori E, Li Shutao, et al. New frontiers in spectral-spatial hyperspectral image classification: The latest advances based on mathematical morphology, Markov random fields, segmentation, sparse representation, and deep learning[J]. IEEE Geoscience and Remote Sensing Magazine, 2018, 6(3): 10-43.  
 [11] Li Shutao, Song Weiwei, Fang Leyuan, et al. Deep learning for hyperspectral image classification: An overview[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(9): 6690-6709.  
 [12] Licciardi G, Marpu P R, Chanussot J, et al. Linear versus non-linear PCA for the classification of hyperspectral data based on the extended morphological profiles[J]. IEEE Geoscience and Remote Sensing Letters, 2012, 9(3): 447-451.  
 [13] He Kaiming, Zhang Xiangyu, Ren Shaoqing, et al. Deep residual learning for image recognition[C]//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas: IEEE, 2016: 770-778.  
 [14] Szegedy C, Liu Wei, Jia Yangqing, et al. Going deeper with convolutions[C]//Proceedings of the 2015 IEEE Conference on Computer Vision and Pattern Recognition

- (CVPR).Boston:IEEE,2015:1–9.
- [15] Krizhevsky A, Sutskever I, Hinton G E. ImageNet classification with deep convolutional neural networks[J]. Communications of the ACM, 2017, 60(6):84–90.
- [16] Ben Hamida A, Benoit A, Lambert P, et al. 3–D deep learning approach for remote sensing image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(8):4420–4434.
- [17] Roy S K, Krishna G, Dubey S R, et al. HybridSN: Exploring 3–D–2–D CNN feature hierarchy for hyperspectral image classification[J]. IEEE Geoscience and Remote Sensing Letters, 2020, 17(2):277–281.
- [18] Zhong Zilong, Li J, Luo Zhiming, et al. Spectral-spatial residual network for hyperspectral image classification: A 3–D deep learning framework[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(2):847–858.
- [19] Paoletti M E, Haut J M, Fernandez–Beltran R, et al. Deep pyramidal residual networks for spectral-spatial hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(2):740–754.
- [20] Roy S K, Manna S, Song Tiecheng, et al. Attention-based adaptive spectral-spatial kernel ResNet for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(9):7831–7843.
- [21] Zhu Minghao, Jiao Licheng, Liu Fang, et al. Residual spectral-spatial attention network for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 59(1):449–462.
- [22] Fang Leyuan, Li Shutao, Duan Wuhui, et al. Classification of hyperspectral images by exploiting spectral-spatial information of superpixel via multiple kernels[J]. IEEE Transactions on Geoscience and Remote Sensing, 2015, 53(12):6663–6674.
- [23] Li Shutao, Lu Ting, Fang Leyuan, et al. Probabilistic fusion of pixel-level and superpixel-level hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(12):7416–7430.
- [24] Zhong Zilong, Li Ying, Ma Lingfei, et al. Spectral-spatial transformer network for hyperspectral image classification: A factorized architecture search framework[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60:5514715.
- [25] Song Ruoxi, Feng Yining, Cheng Wei, et al. BS2T: Bottleneck spatial-spectral transformer for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60:5532117.
- [26] Sun Le, Zhao Guangrui, Zheng Yuhui, et al. Spectral-spatial feature tokenization transformer for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60:5522214.
- [27] Roy S K, Deria A, Hong Danfeng, et al. Multimodal fusion transformer for remote sensing image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2023, 61:5515620.
- [28] Hong Danfeng, Han Zhu, Yao Jing, et al. SpectralFormer: Rethinking hyperspectral image classification with transformers[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021, 60:5518615.
- [29] He Ji, Zhao Lina, Yang Hongwei, et al. HSI–BERT: Hyperspectral image classification using the bidirectional encoder representation from transformers[J]. IEEE Transactions on Geoscience and Remote Sensing, 2020, 58(1):165–178.
- [30] Yu Haoyang, Xu Zhen, Zheng Ke, et al. MSTNet: A multi-level spectral-spatial transformer network for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60:5532513.
- [31] Gao Liang, Liu Hui, Yang Minhang, et al. STransFuse: Fusing swin transformer and convolutional neural network for remote sensing image semantic segmentation[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2021, 14:10990–11003.
- [32] Lange J, Cavallaro G, Götz M, et al. The influence of sampling methods on pixel-wise hyperspectral image classification with 3D convolutional neural networks[C]//Proceedings of the IGARSS 2018–2018 IEEE International Geoscience and Remote Sensing Symposium. Valencia: IEEE, 2018:2087–2090.
- [33] Chen Yushi, Lin Zhouhan, Zhao Xing, 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.
- [34] Zhao Chunhui, Wan Xiaoqing, Yan Yiming. Spectral-spatial classification of hyperspectral images based on joint bilateral filter and stacked sparse autoencoder[C]//Proceedings of the 2017 First International Conference on Electronics Instrumentation & Information Systems(EIIS). Harbin: IEEE, 2017:1–5.
- [35] Yang Xiaofei, Ye Yunming, Li Xutao, et al. Hyperspectral image classification with deep learning models[J]. IEEE Transactions on Geoscience and Remote Sensing, 2018, 56(9):5408–5423.
- [36] Hang Renlong, Liu Qingshan, Hong Danfeng, et al. Cascaded recurrent neural networks for hyperspectral image classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019, 57(8):5384–5394.
- [37] Zhou Feng, Hang Renlong, Liu Qingshan, et al. Hyperspectral image classification using spectral-spatial LSTMs[J]. Neurocomputing, 2019, 328:39–47.
- [38] Li Wei, Wu Guodong, Zhang Fan, et al. Hyperspectral image classification using deep pixel-pair features[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017, 55(2):844–853.

## Optimization of Hyperspectral Image Classification Using Spectral-feature Based Approaches

LI Xunfeng, LI Xiaohua\*

(Collage of Computer Science, Sichuan University, Chengdu 610065, China)

### Abstract:

**Objective** Hyperspectral imaging technology captures image data across a wide range of wavelengths, providing rich spectral information for each pixel in the scene. This detailed spectral information enables precise identification and classification of various materials and land cover types, making HSIC a critical task in remote sensing. The vast amount of data contained in hyperspectral images presents significant challenges and opportunities for advanced image processing techniques, particularly those involving deep learning. Over the past decade, deep learning has revolutionized numerous fields, including image processing and classification. In the context of HSIC, deep learning techniques, especially those utilizing convolutional neural networks (CNNs) and recurrent neural networks (RNNs), have demonstrated remarkable improvements in performance. These methods effectively capture the complex and high-dimensional nature of hyperspectral data, extracting both spatial and spectral features to enhance classification accuracy. Methods based on spatial-spectral features have gained substantial attention due to their ability to integrate spatial context with spectral information. These methods can better discriminate between different classes by combining these two types of features. However, a limitation arises when applying sample partitioning strategies from natural image classification directly to hyperspectral images. This approach often leads to unintended overlap between training and test samples, particularly in spatially contiguous regions, which can artificially inflate classification accuracy and reduce the model's generalization capability.

**Methods** This study proposed a unique approach that reduced the size of input image patches in spatial-spectral-based methods, enabling an effective separation of the training and test sets. When the image patches were reduced to a specific size, their features gradually transformed into pure spectral features. Considering this observation, this study refocused on the HSIC method based on spectral features and proposed an improved method that relied on deep pixel pair features for faster and more robust hyperspectral image classification. The proposed method improved the original pixel pair feature method in two main aspects. The first key improvement introduced was the distance-constrained pixel pair generation method. Traditional pixel pair methods often suffered from inefficiencies and the inclusion of redundant or irrelevant pixel pairs. The proposed method ensured that only meaningful and diverse pixel pairs were selected by incorporating distance constraints, enhancing the training process and the robustness of the model. This method not only efficiently constructed a large-scale pixel pair training set but also eliminated redundant and unreasonable pixel pairs, leading to more accurate and efficient model training. The second major enhancement was the result-reuse voting strategy combined with a classification method. This strategy optimized the classification process by reusing intermediate results, reducing the computational burden, and accelerating the overall classification time. This approach enhanced efficiency while maintaining high classification accuracy by avoiding redundant model calls. The result-reuse voting strategy ensured that the classification process remained both effective and efficient, making it suitable for large-scale hyperspectral datasets.

**Results and Discussions** Extensive experiments were conducted on public hyperspectral image datasets to validate the effectiveness of the proposed method. The results demonstrated that the proposed method outperformed existing spatial-spectral feature-based methods in terms of classification accuracy. In addition, compared to standard pixel pair methods, the proposed approach significantly reduced computational time, making it more suitable for practical applications. The experimental results confirmed that the proposed method achieved higher classification accuracy in practical applications and significantly improved time efficiency compared to existing methods.

**Conclusions** Accordingly, this study addresses a critical challenge in hyperspectral image classification by proposing a novel approach that shifts the emphasis from spatial-spectral features to purely spectral features through the reduction of the input image block size. The introduction of a distance-constrained pixel-pair generation method and a result-reuse voting strategy significantly enhances classification efficiency and robustness. The experimental results validate the superiority of the proposed method, highlighting its potential for broader applications in remote sensing. This research contributes to the development of more accurate and generalizable HSIC models, paving the way for future advancements in hyperspectral image analysis by addressing the overlap issue between training and test samples. The study highlights the importance of considering the unique characteristics of hyperspectral data and provides a robust framework for leveraging spectral features to achieve high-precision classification. The proposed method represents a significant advancement in HSIC, providing a practical and efficient solution to challenges inherent in existing methods. Future research will explore the integration of additional constraints and optimizations to improve the performance and applicability of hyperspectral image classification techniques.

**Key words:** hyperspectral image classification; deep learning; pixel-pair feature; sample division strategy

(编辑 张 琼)

引用格式: Li Xunfeng, Li Xiaohua. Optimization of hyperspectral image classification using spectral-feature based approaches[J]. *Advanced Engineering Sciences*, 2025, 57(6): 323–334. [李洵枫, 李晓华. 基于光谱特征的高光谱图像分类的优化方法[J]. *工程科学与技术*, 2025, 57(6): 323–334.]