Journal of Image and Signal Processing 图像与信号处理, 2026, 15(4), 605-621
Published Online October 2026 in Hans. https://www.hanspub.org/journal/jisp
https://doi.org/10.12677/jisp.2026.154053

Hans汉斯

基于SE注意力增强的双分支子像素引导高光谱图像分类网络

郑博文

成都信息工程大学统计学院，四川 成都

收稿日期：2026年9月27日；录用日期：2026年10月19日；发布日期：2026年10月29日

摘　要

高光谱图像(Hyperspectral Image, HSI)富含光谱与空间信息，在农业、环境等领域应用广泛，但高维冗余、类别不均衡和混合像元等问题使传统分类易出现混淆与误判。本文在DSNet基础上提出SE-DSNet，通过空间－光谱特征分支与子像素解混分支联合建模空间与丰度信息，增强混合像元表达；引入SE通道注意力自适应加权关键波段，结合Focal Loss缓解类别不平衡。在Indian Pines与Augsburg数据集上与CNN2D、CNN3D、GRU及DSNet对比，实验表明SE-DSNet在OA、AA和Kappa上均取得较优结果，验证了其在复杂场景下的有效性与泛化能力。

关键词

高光谱图像分类，双分支网络，子像素解混，SE注意力机制，焦点损失，深度学习

SE-Aware Dual-Branch Subpixel-Guided Network for Hyperspectral Image Classification

Bowen Zheng

School of Statistics, Chengdu University of Information Technology, Chengdu Sichuan

Received: September 27, 2026; accepted: October 19, 2026; published: October 29, 2026

Abstract

Hyperspectral images (HSIs) are rich in spectral and spatial information and are widely used in agriculture, environmental monitoring, and other fields. However, problems such as high-dimensional redundancy, class imbalance, and mixed pixels tend to cause confusion and mis-classification in traditional

文章引用: 郑博文. 基于 SE 注意力增强的双分支子像素引导高光谱图像分类网络[J]. 图像与信号处理, 2026, 15(4): 605-621. DOI: 10.12677/jisp.2026.154053

classification. This paper proposes SE-DSNet based on DSNet, which jointly models spatial and abundance information through a spatial-spectral feature branch and a subpixel unmixing branch to enhance the representation of mixed pixels. It introduces SE channel attention to adaptively weight key bands and combines Focal Loss to alleviate class imbalance. Experiments conducted on the Indian Pines and Augsburg datasets demonstrate that SE-DSNet outperforms CNN2D, CNN3D, GRU, and DSNet in terms of overall accuracy (OA), average accuracy (AA), and Kappa coefficient, verifying its effectiveness and generalization capability in complex scenarios.

Keywords

Hyperspectral Image Classification, Dual-Branch Network, Subpixel Unmixing, SE Attention Mechanism, Focal Loss, Deep Learning

Copyright © 2026 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. 绪论

1.1. 选题研究背景与意义

高光谱成像(HSI)具备高光谱分辨率和地物区分能力，被广泛用于农业、环境、资源及城市地物识别，HSI 分类已成为遥感信息处理的重要方向[1]。然而，该任务仍面临多重挑战：空间分辨率限制导致混合像元现象，边界区域光谱混叠易引发误判；波段数量多、冗余度高，加之标注成本昂贵[2]，致使训练样本有限，模型易过拟合且泛化不足。深度学习方法(如 CNN [3]、3D-CNN [4]及 RNN/GRU/Transformer [5])虽显著提升了分类性能，但多数仍以像素级特征建模为主，对混合像元内部结构缺乏显式描述。DSNet [6]引入子像素解混分支，联合空间–光谱与丰度信息，在混合区域取得较好效果，但其未能区分通道重要性，且未处理类别不均衡问题。为此，本文在 DSNet 基础上引入 SE 通道注意力机制，对关键波段自适应加权，并采用 Focal Loss 缓解类别不平衡，通过 Indian Pines 与 Augsburg 两个典型数据集验证改进模型在复杂场景下的分类性能。

1.2. 国内外研究现状

随着遥感成像技术的发展，高光谱图像分类逐渐成为遥感信息处理领域的重要研究方向。由于高光谱数据同时包含丰富的光谱信息与空间结构信息，研究方法也经历了从传统机器学习到深度学习的发展过程。近年来，研究重点逐渐由单一像素分类转向空间–光谱联合建模以及混合像元建模。

1.2.1. 基于传统机器学习的高光谱分类方法

早期方法以 SVM [7]、RF、KNN 为主，常结合 PCA 或 LDA 降维，其中 SVM 在小样本条件下表现较好。但这些方法依赖人工设计特征，对空间上下文和复杂非线性关系利用不足，边界和混合像元区域分类效果较差，整体性能随着数据复杂度增加而受限。

1.2.2. 基于深度学习的高光谱分类方法

深度学习显著提升了分类效果。2D-CNN [8]提取空间特征，3D-CNN [9] [10]联合空间–光谱建模，但参数较多、小样本易过拟合；RNN/LSTM/GRU 建模光谱序列关系；Transformer [11]及 CNN-Transformer 结合[12]引入自注意力以增强全局建模能力，但参数量大、训练不稳定[13]，Transformer 亦被用于高光谱

任务[14]。总体而言，深度学习方法提升了性能，但仍以像素级特征建模为主。

1.2.3. 子像素建模与光谱解混方法

为处理混合像元，传统解混方法(几何、统计、稀疏表示)及 NMF 被广泛采用。深度学习中，自编码器和 DSNet [15]将解混与分类结合，但 DSNet 未对不同光谱通道的重要性进行区分，存在进一步改进空间。

1.2.4. 损失函数优化策略

训练策略影响模型性能，高光谱数据存在类别不均衡，交叉熵易偏向多数类。Focal Loss [16]通过降低易分类样本损失权重，使模型更关注难分类与少数类样本，提升鲁棒性。

综上，目前高光谱图像分类研究已经在空间–光谱联合建模、注意力机制以及子像素建模等方面取得了较大进展，但在复杂场景下仍存在类别混叠、边界误分类以及小样本训练困难等问题[17]。因此，本文在 DSNet 基础上引入 SE 注意力机制，并结合 Focal Loss，对模型进行改进，以进一步提升复杂场景下的高光谱图像分类性能[18]。

1.3. 本文研究内容

针对高光谱图像分类中存在的混合像元、光谱冗余以及类别不平衡等问题，本文在双分支子像素引导网络(DSNet)基础上进行了改进。主要研究内容如下：

(1) 复现 DSNet 双分支子像素引导网络。对原始 DSNet 模型进行了复现与实验验证，该模型通过空间–光谱特征提取分支与子像素解混分支的联合建模，在混合像元区域具有较好的分类表现，为后续改进提供基线。

(2) 引入 SE 注意力机制增强通道特征表达。在特征提取阶段加入 SE 通道注意力模块，通过学习通道间依赖关系对关键波段进行自适应加权，以突出有效特征并减少冗余信息干扰。

(3) 采用 FocalLoss 缓解类别不平衡问题。使用 FocalLoss 替代标准交叉熵损失函数，通过降低易分类样本的权重，使模型更多关注难分类与少数类样本，提高模型在类别不均衡条件下的分类稳定性。

(4) 实验验证与结果分析。在 IndianPines 与 Augsburg 两个典型高光谱数据集上开展对比实验与消融实验，评估改进模块的实际效果。

此外，本文还进行了消融实验，用于分析 SE 注意力机制与 FocalLoss 对模型性能的具体影响。

1.4. 本文创新点

本文在 DSNet 基础上进行针对性优化，主要创新点如下：

(1) 引入 SE 通道注意力机制。针对原始 DSNet 未区分通道重要性的问题，在双分支网络中嵌入 SE 模块，通过通道自适应加权增强关键光谱特征。SE 模块轻量，适合高光谱小样本场景，实验表明其在玉米、大豆等光谱相似类别上分类效果有所改善。

(2) 采用 Focal Loss 缓解类别不平衡。针对高光谱数据类别分布不均、交叉熵易偏向多数类的问题，以 Focal Loss 替代交叉熵损失，降低易分类样本权重，使模型更关注难分类与少数类样本，提升不均衡条件下的分类稳定性。

(3) 验证 SE 模块对双分支结构的改进效果。在 Indian Pines 与 Augsburg 数据集上进行对比与消融实验，结果表明加入 SE 模块后 OA、AA 及 Kappa 均有提升，且分类结果图在边界区域和复杂场景中具有更好的空间连续性，验证了通道注意力机制对关键光谱特征表达能力的增强作用。

2. 数据集

为验证模型在不同场景下的分类性能与泛化能力，选取 Indian Pines (农业场景)和 Augsburg (城市场

景)两个公开高光谱数据集。两者在空间结构、类别分布和混合像元程度上差异显著，可综合评估模型的光谱区分能力与空间建模能力。

2.1. Indian Pines 数据集

Indian Pines 数据集是高光谱图像分类领域中应用最广泛的数据集之一，如图 1 所示，由 AVIRIS 传感器获取于美国印第安纳州农业区，影像尺寸 145 × 145 像素，包含 220 个光谱波段(波长 0.4~2.5 μm)，去除水汽吸收波段后保留 200 个有效波段，空间分辨率 20 m。共 16 类地物(以农作物为主)，存在类别光谱相似性高、分布严重不均衡及混合像元明显等问题，常用于验证模型在农业场景下的分类性能。

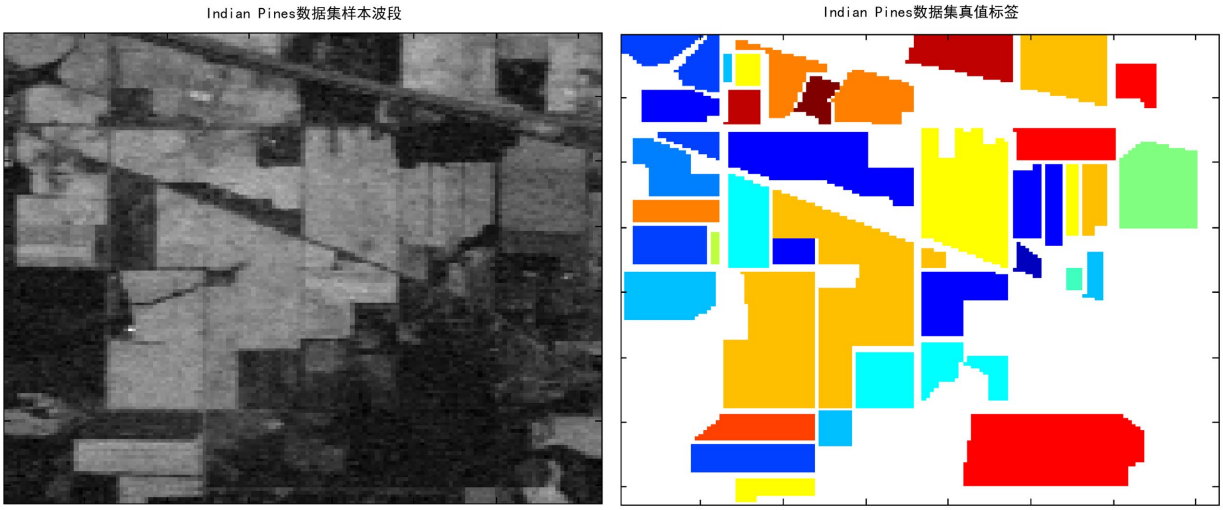


Figure 1. Image of the Indian Pines dataset
图 1. Indian Pines 数据集图像

该数据集共包含 16 类地物类别，其类别名称及对应样本数量如表 1 所示。

Table 1. Ground-Truth classes and sample numbers of Indian Pines scene
表 1. Indian Pines 场景真值类别及其对应样本数量

序号	类别名称(中文)	类别名称(英文)	样本数量
1	苜蓿	Alfalfa	46
2	免耕玉米	Corn-notill	1428
3	少耕玉米	Corn-mintill	830
4	玉米	Corn	237
5	牧场草地	Grass-pasture	483
6	乔木草地	Grass-trees	730
7	修剪牧场草地	Grass-pasture-mowed	28
8	风干干草	Hay-windrowed	478
9	燕麦	Oats	20
10	免耕大豆	Soybean-notill	972
11	少耕大豆	Soybean-mintill	2455
12	清耕大豆	Soybean-clean	593

续表

13	小麦	Wheat	205
14	林地	Woods	1265
15	建筑－草地－树木－道路混合区	Buildings-Grass-Trees-Drives	386
16	砖石－钢结构塔架	Stone-Steel-Towers	93

2.2. Augsburg 数据集

Augsburg 数据集是近年来较常用于高光谱分类研究的数据集之一，能够较好反映城市遥感场景中的复杂空间结构。如图 2 所示，该数据集获取于德国奥格斯堡，涵盖城市建筑、道路、植被、水体等 7 类地物，空间纹理复杂，地物边界数量多，混合像元现象突出，对模型空间特征提取能力要求较高，适用于评估复杂城市场景下的泛化性与鲁棒性。

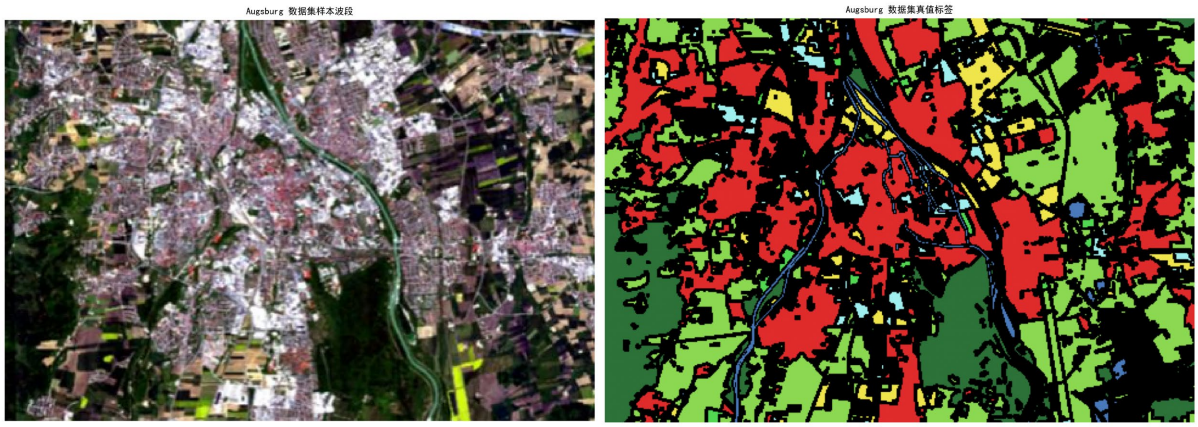


Figure 2. Image of the Augsburg dataset
图 2. Augsburg 数据集图像

Augsburg 数据集共包含 7 类地物，其类别名称及对应样本数量如表 2 所示。

Table 2. Ground-Truth classes and sample numbers of Augsburg scene
表 2. Augsburg 场景真值类别及其对应样本数量

序号	类别名称(中文)	类别名称(英文)	样本数量
1	森林	Forest	13,507
2	住宅区	ResidentialArea	30,329
3	工业区	IndustrialArea	3851
4	低矮植被	LowPlants	26,857
5	分配地	Allotment	575
6	商业区	CommercialArea	1645
7	水体	Water	1530

2.3. 数据集对比分析

为了更全面地验证模型性能，本文同时选取农业场景与城市场景数据集开展实验。两个数据集在场

景类型、类别组成和空间结构上具有明显差异：Indian Pines 侧重光谱区分能力的验证，Augsburg 侧重空间结构建模能力的评估。两者均存在类别不均衡与混合像元问题，结合使用可从不同角度综合评估模型的分类性能与泛化能力。

3. 问题概述与研究方法

3.1. 问题概述

高光谱图像(HSI)每个像元包含数十至数百个连续波段，高维特性在提升地物可分性的同时引入大量冗余与复杂性。分类任务主要面临两类挑战：一是高维特征空间导致维数灾难，波段高度相关，加之标注成本高、训练样本有限，易致过拟合；二是混合像元普遍存在，受空间分辨率限制，像元光谱可视为多个端元光谱的线性组合，即线性混合模型(LMM)，其表达形式为：

$$x=\sum_{i=1}^{K}a_i e_i+n \tag{1}$$

其中，x 表示观测像元光谱向量，e_i 表示第 i 个端元光谱，a_i 表示对应丰度系数，K 表示端元数量，n 表示噪声项。丰度系数通常满足非负约束与和为 1 约束：$a_i\geq 0$，$\sum_{i=1}^{K}a_i=1$。

此外，高光谱分类的关键在于联合建模光谱、空间与子像素混合结构。DSNet 以双分支联合空间与丰度信息，在混合区域表现良好，但未建模通道间重要性，关键光谱信息表达不足。为此，本文引入 Squeeze-and-Excitation 注意力机制，对通道特征自适应重标定以增强判别性表达，其形式可表示为：

$$\tilde{F}=\sigma\left(W_2\delta\left(W_1 z\right)\right)\otimes F \tag{2}$$

其中，F 为输入特征，z 表示全局池化后的通道描述向量，W_1 与 W_2 为全连接层参数，δ 表示 ReLU 激活函数，σ 表示 Sigmoid 函数，$\otimes$ 表示逐通道加权操作。该机制能够自适应增强关键波段响应并抑制冗余信息，从而提升特征表达能力[19]。

3.2. SE-DSNet 总体框架

本文提出一种基于 SE 注意力增强的双分支子像素引导网络(SE-DSNet)用于高光谱图像分类。整体框架如图 3 所示。该方法在原始 DSNet 基础上引入通道注意力机制，以增强光谱特征表达能力并优化空间特征表示，从而提升复杂场景下的分类性能。

网络整体由以下模块组成：数据预处理模块、子像素解混分支、空间–光谱特征提取分支、SE 注意力增强模块以及特征融合与分类模块。

3.3. 数据预处理

由于高光谱图像不同波段之间数值范围差异较大，为减小波段尺度差异对模型训练的影响，本文首先对原始数据进行归一化处理。本文首先采用最大最小归一化方法对原始数据进行标准化处理：

$$x_i'=\frac{x_i-x_{\min}}{x_{\max}-x_{\min}} \tag{3}$$

其中，x_i' 为归一化后的结果，x_i 为原始像元值，$x_{\min}$、$x_{\max}$ 分别表示当前波段最小值和最大值。

随后，采用滑动窗口方式提取局部邻域 Patch，用于构建输入样本。设窗口大小为 $P\times P$，则每个样本表示为：

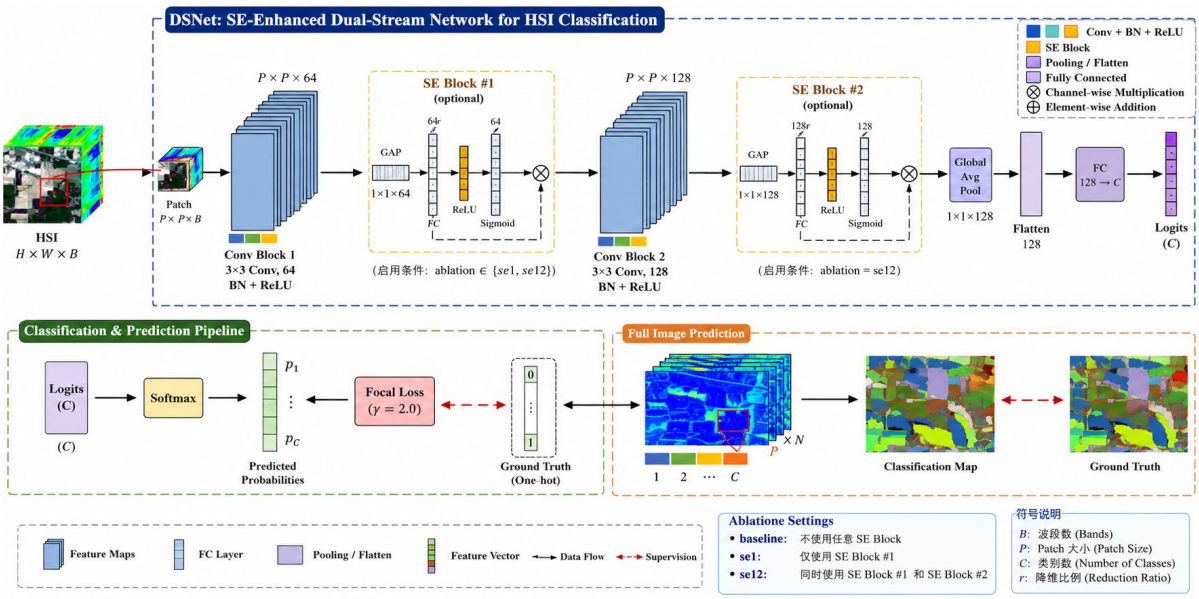


Figure 3. The proposed SE-DSNet framework
图 3. 所提出的 SE-DSNet 框架

$$X \in K^{P\times P\times B} \tag{4}$$

其中， P 表示空间窗口大小， B 表示光谱波段数。

3.4. 双分支特征提取网络

3.4.1. 子像素解混分支

为解决混合像元问题，DSNet 引入子像素解混机制。该分支利用编码器－解码器结构学习像元内部丰度信息。编码阶段：

$$A = f_{enc}\left(X\right) \tag{5}$$

其中， $f_{enc}\left(\cdot\right)$ 表示编码器， A 表示丰度矩阵。解码阶段：

$$\hat{X} = f_{dec}\left(A\right) \tag{6}$$

其中， $f_{dec}\left(\cdot\right)$ 表示解码器， $\hat{X}$ 表示重构光谱。

3.4.2. 空间－光谱特征提取分支

该分支用于学习高光谱图像中的判别性特征，包括光谱依赖关系与空间上下文信息。输入局部图像块后，通过卷积神经网络进行特征提取：

$$F_l = \phi\left(W_l * F_{l-1} + b_l\right) \tag{7}$$

其中， F_l 表示第 ℓ 层输出特征， W_l 与 b_l 分别表示卷积核参数与偏置项， $*$ 表示卷积操作， ϕ 表示非线性激活函数。通过多层卷积与非线性激活函数，模型能够逐步提取深层空间语义特征[20]。

3.5. SE 注意力增强模块

高光谱图像中存在高度相关的光谱波段和复杂的空间分布，标准卷积网络难以自适应地强调最具有判别性的特征通道。为此，我们在 DSNet 中嵌入了通道注意力机制(SE)，以动态重新校准各通道的响应强度[21]。同时，为了利用高光谱图像丰富的光谱信息，我们引入了一个无监督的光谱重建分支，通过最

小化光谱角距离(SAD)来增强特征的光谱保真度。

(1) Squeeze 操作

给定中间特征图$U \in \mathbb{R}^{C\times H\times W}$，SE 模块首先通过全局平均池化将空间信息压缩为通道描述$z \in \mathbb{R}^{C}$：

$$z_c = \frac{1}{H\times W}\sum_{i=1}^{H}\sum_{j=1}^{W} u_c(i,j) \tag{8}$$

其中，$U_c(i,j)$表示第 c 个通道位置(i,j)的特征值，H 与 W 分别表示特征图高度与宽度。该操作将每个通道映射为一个全局统计值。

(2) Excitation 操作

通过两层全连接网络学习通道权重：

$$s = \sigma\left(W_2\delta\left(W_1 z\right)\right) \tag{9}$$

其中，s 表示通道权重向量，$\delta(\cdot)$为 ReLU 函数，$\sigma(\cdot)$为 Sigmoid 函数，W_1、W_2为可学习参数；该过程用于建模通道之间的依赖关系。

(3) 特征重标定

最终对原始特征进行加权[22]：

$$\tilde{U}_c = s_c \otimes U_c \tag{10}$$

其中，s_c表示第 c 个通道权重，U_c表示对应通道特征。通过该机制，模型能够增强关键波段特征响应，抑制冗余信息，提高特征表达能力。

3.6. 特征融合与分类模块

空间分支与子像素分支提取的特征最终进行融合：

$$F_{\text{fusion}} = \text{Concat}\left(F_{\text{spatial}}, F_{\text{unmix}}\right) \tag{11}$$

其中，F_{spatial}表示空间特征，F_{unmix}表示解混特征。

融合后的特征通过卷积层或全连接层进一步映射，最终输出分类结果：

$$\hat{y} = \text{Softmax}\left(W_c F_{\text{fusion}} + b_c\right) \tag{12}$$

其中，W_c和b_c表示分类层参数与偏置项，y表示最终分类概率。

3.7. 损失函数

为了同时优化重构能力与分类性能，本文采用联合损失函数进行训练：

$$L = L_{\text{SAD}} + L_{\text{focal}} \tag{13}$$

(1) 重构损失

为了确保子像素解混分支能够有效地学习到像元内部的丰度信息，我们对该分支的输出施加重构约束。如 3.4.1 节所述，解混分支采用编码器–解码器结构，旨在将输入光谱 X 分解为端元光谱与丰度矩阵，并利用解码器将其重构成原始光谱。我们采用均方误差(MSE)作为重构损失L_{SAD}，以最小化原始输入光谱 X 与解码器重构光谱 $\hat{X}$ 之间的差异：

$$L_{\text{SAD}} = \frac{1}{B}\sum_{i=1}^{B}\left\|X_i - \hat{X}_i\right\|_2^2 \tag{14}$$

其中，B 为批次大小，X_i 和 $\hat{X}_i$ 分别为第 i 个样本的原始输入光谱和网络重构光谱。通过最小化该重构误差，网络被迫学习到能够保留原始光谱主要信息的低维丰度表示，从而为分类分支提供更具判别性的子

像素级特征。

(2) 分类损失

在基线方法中，分类任务通常采用标准的交叉熵损失函数：

$$L_{\text{loss}} = -\sum_{i=1}^{N} y_i \log p_i \tag{15}$$

其中，y_i 为第 i 个类别真实标签的编码，p_i 为模型预测第 i 个类别的概率。然而，由于高光谱图像中存在严重的类别不均衡问题(少数类样本不足)，交叉熵损失会偏向多数类，导致少数类的识别性能下降。

为此，本文引入采用焦点损失(Focal Loss)替代标准交叉熵损失，以缓解类别不平衡问题，焦点损失定义为：

$$L_{\text{focal}} = -\sum_{i=1}^{N} \alpha_i \left(1-p_i\right)^{\gamma} y_i \log\left(p_i\right) \tag{16}$$

其中，p_i 为真实类别的预测概率，α 为类别权重，γ 为聚焦参数。该损失函数能够减少易分类样本的影响，使模型更加关注难分类样本，从而缓解类别不平衡问题。

4. 实验与结果分析

4.1. 实验设置

本文所有实验基于 PyTorch 框架，所有方法在相同数据划分与训练策略下进行。数据预处理采用最大最小归一化，并以 7×7 滑动窗口提取空间邻域块作为输入。训练参数：批大小 16、轮数 200、学习率 0.001、Adam 优化器[23] [24]。损失函数为联合损失(分类 Focal Loss，$\gamma=2$，α 取类别频率倒数[25]；重构损失为 MSE)。对比方法包括 CNN2D (两层卷积 + 全连接)、CNN3D (HybridSN 结构)、GRU (两层，隐层 128)、DSNet (原始设置)及本文 SE-DSNet。每类随机取 10%训练、90%测试。评价指标采用整体精度(OA)、平均精度(AA)和 Kappa 系数。

对应计算公式如下：

(1) 整体精度(OA)

$$\text{OA} = \frac{\sum_{i=1}^{C} N_{ii}}{N} \tag{17}$$

其中，N_{ii} 表示第 i 类被正确分类的样本数量，N 表示总测试样本数量，C 表示类别总数。

(2) 平均精度(AA)

$$\text{AA} = \frac{1}{C}\sum_{i=1}^{C} \frac{N_{ii}}{N_i} \tag{18}$$

其中，N_i 表示第 i 类真实样本数量。

(3) Kappa 系数

$$\text{Kappa} = \frac{P_0 - P_e}{1-P_e} \tag{19}$$

其中，P_0 表示实际分类准确率；P_e 表示随机分类情况下的理论一致性概率。

(4) 模型选择策略

在训练过程中，为避免仅依赖单一指标带来的偏差，本文采用综合评价指标进行模型选择：

$$\text{Score} = \text{OA} + \lambda \cdot \text{AA} \tag{20}$$

其中，λ用于调节总体精度与类别均衡性之间的权重关系。上述评价指标能够从整体分类性能、类别均衡性以及分类一致性等多个角度对模型进行综合评估。

4.2. 对比实验结果分析

4.2.1. Indian Pines 数据集结果分析

在 Indian Pines 数据集上(表 3、表 4)，本文提出的 SE-DSNet 在 OA (96.72%)、AA (98.05%)、Kappa (96.24%)和综合 Score (116.33)上均取得最优，相比 DSNet，OA 提高 3.11%，Kappa 提高 3.55%。传统 CNN2D (OA 83.15%)和 CNN3D (OA 60.04%)受限于光谱利用不足或小样本训练不稳定；GRU (OA 78.02%)因缺乏空间建模能力，效果弱于双分支结构。各类别中，SE-DSNet 在苜蓿、牧场草地、风干干草等易混或小样本类别上提升显著(牧场草地由 91.39%升至 99.71%)，在乔木草地、免耕大豆等类别达到满分，说明模型对判别特征提取和复杂类别区分能力较强，整体稳定性优于对比方法。

Table 3. Classification accuracy of different methods on Indian Pines dataset (%)
表 3. 不同方法在 Indian Pines 数据集分类结果的精度(%)

类别	CNN2D	CNN3D	GRU	DSNet	SE-DSNet
苜蓿	92.49	43.50	71.79	90.03	93.21
免耕玉米	61.48	43.88	72.33	95.92	95.41
少耕玉米	100.00	84.24	83.16	98.37	98.91
玉米	91.72	83.89	73.37	96.42	98.66
牧场草地	89.24	71.59	87.02	91.39	99.71
乔木草地	100.00	94.31	81.35	97.95	100.00
修剪牧场草地	90.09	60.78	100.00	93.90	98.37
风干干草	69.73	49.13	64.76	88.96	95.45
燕麦	75.53	23.23	100.00	95.92	95.04
免耕大豆	100.00	98.15	84.86	99.38	100.00
少耕大豆	90.35	80.71	66.50	99.38	98.71
清耕大豆	100.00	90.00	64.00	99.38	97.88
小麦	100.00	95.56	98.15	99.38	100.00
林地	20.51	53.85	89.39	94.87	97.44
建筑－草地－树木－道路混合区	100.00	100.00	81.52	100.00	100.00
砖石－钢结构塔架	100.00	100.00	100.00	100.00	100.00

Table 4. Overall accuracy of different methods on Indian Pines dataset (%)
表 4. 不同方法在 Indian Pines 数据集总体结果的精度(%)

方法	OA	AA	Kappa	Score
CNN2D	83.15	86.32	80.75	100.41
CNN3D	60.04	73.30	54.79	74.70
GRU	78.02	84.30	75.11	94.88
DSNet	93.61	96.18	92.69	112.85
SE-DSNet (本文方法)	96.72	98.05	96.24	116.33

4.2.2. Augsburg 数据集结果分析

在 Augsburg 城市场景(表 5、表 6)中，SE-DSNet 的 OA 为 90.47%，AA 为 66.21%，Kappa 为 86.27%，Score 为 103.71，均优于 DSNet (OA 88.34%, Kappa 83.20%)及其他对比方法。CNN2D、CNN3D 和 GRU 的 OA 分别为 81.33%、80.67%和 79.85%，明显低于双分支结构模型。类别层面，森林(97.10%)、住宅区(96.16%)、低矮植被(93.72%)等主要类别精度显著提升；分配地(57.17%)和水体(49.64%)等困难类别也有一定改善。但工业区类别(51.51%)低于 DSNet (65.74%)，表明复杂异质性类别仍存在特征偏置。总体而言，SE-DSNet 在复杂城市场景中表现出良好的泛化能力。

Table 5. Classification accuracy of different methods on Augsburg dataset (%)
表 5. 不同方法在 Augsburg 数据集分类结果的精度(%)

类别	CNN2D	CNN3D	GRU	DSNet	SE-DSNet
森林	85.76	93.02	81.10	93.56	97.10
住宅区	87.44	92.68	85.51	94.47	96.16
工业区	69.56	21.46	84.41	65.74	51.51
低矮植被	81.39	78.29	82.25	89.93	93.72
分配地	41.68	58.70	32.10	51.43	57.17
商业区	14.10	9.46	5.56	14.59	18.19
水体	35.83	9.36	19.18	42.27	49.64

Table 6. Overall accuracy of different methods on Augsburg dataset (%)
表 6. 不同方法在 Augsburg 数据集总体结果的精度(%)

方法	OA	AA	Kappa	Score
CNN2D	81.33	59.40	73.27	93.21
CNN3D	80.67	51.85	72.01	91.04
GRU	79.85	55.59	71.41	90.97
DSNet	88.34	64.57	83.20	101.25
SE-DSNet (本文方法)	90.47	66.21	86.27	103.71

4.2.3. 综合分析

综合两个数据集，SE-DSNet 在农业和城市场景中均取得最优性能。在 Indian Pines 上，OA 达 96.72%、Kappa 达 96.24%，优于 DSNet 分别约 3.1 和 3.6 个百分点；在 Augsburg 上，OA 和 Kappa 分别达 90.47%和 86.27%。模型对混合特征明显的类别(如牧场草地、水体)提升显著，但对强异质性类别(如工业区、商业区)提升有限。与 CNN、GRU 相比，双分支结构联合空间－光谱－子像素信息优势明显，SE 注意力和 Focal Loss 的引入进一步提升了整体精度与类别均衡性，验证了模型在多种场景下的鲁棒性与适应性。

4.3. 分类结果可视化分析

为直观展示不同方法的分类效果，本文分别在 Indian Pines 数据集和 Augsburg 数据集上对各模型的分类结果进行了可视化对比，如图 4 和图 5 所示。

4.3.1. Indian Pines 数据集可视化分析

从分类结果图(图 4)可以看出，CNN2D 和 GRU 的分类图中存在大量椒盐噪声和碎片化区域，区域连续性较差；CNN3D 虽有所改善，但边界仍较模糊；DSNet 的区域完整性和边界保持能力明显优于前三者，

大面积地物内部基本一致。本文提出的 SE-DSNet 在视觉效果上最佳，分类区域最完整，噪声点最少，边界最为清晰，与真实标注最为接近。该结果与表 3、表 4 的定量实验一致。

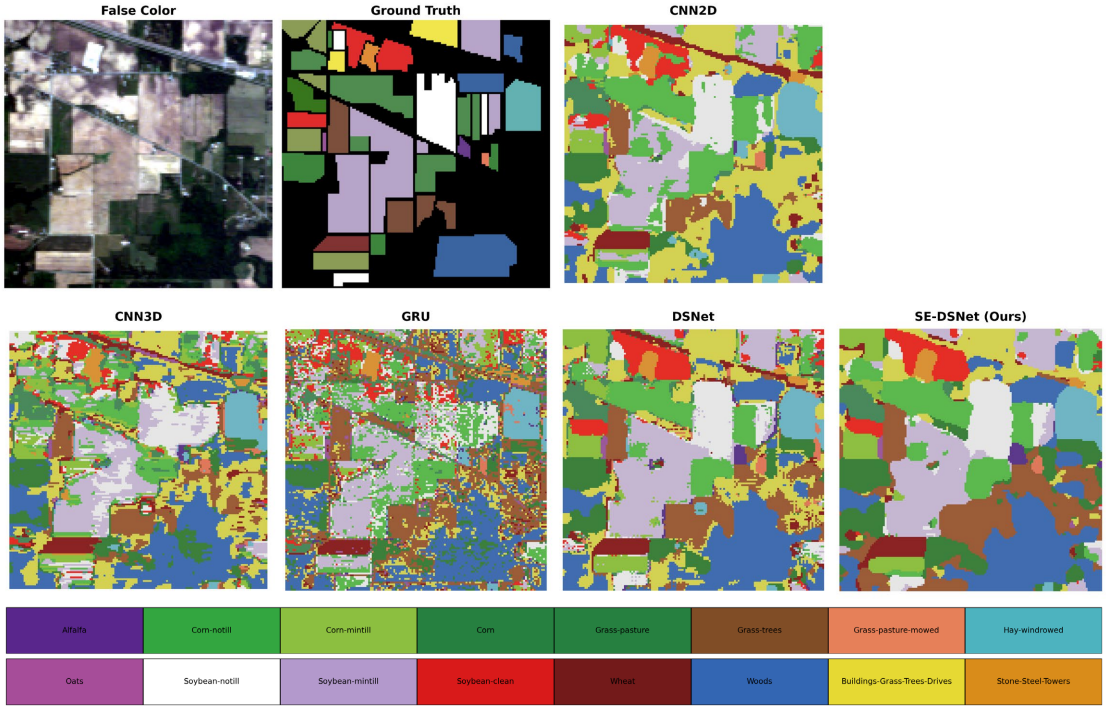


Figure 4. Classification results of different methods for Indian Pines dataset
图 4. 不同方法在 Indian Pines 数据集上的分类结果图

4.3.2. Augsburg 数据集可视化分析

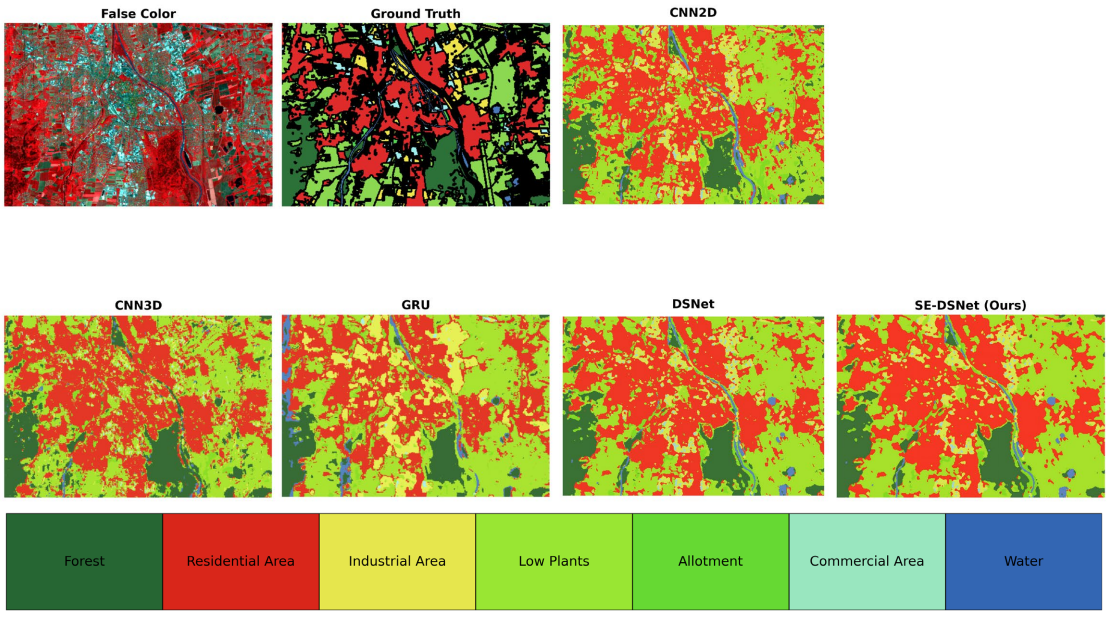


Figure 5. Classification results of different methods for Augsburg dataset
图 5. 不同方法在 Augsburg 数据集上的分类结果图

从分类结果图(图 5)可以看出，CNN2D 和 GRU 虽能大致勾勒森林、住宅区等主要地物轮廓，但商业区、水体等困难类别几乎被完全误分；CNN3D 出现严重的类别丢失现象，工业区和商业区大面积消失；DSNet 整体分类质量明显提升，主要地物区域完整，但商业区仍是明显短板。本文提出的 SE-DSNet 在森林、住宅区、低矮植被上均取得最优视觉效果，分配地和水体误分点进一步减少，分类图最接近真值图。需要注意的是，工业区区域出现比 DSNet 更多的误分现象，与表 5 中该类别精度下降一致。

4.4. 消融实验与分析

为验证 SE 注意力与 Focal Loss 的贡献，以 DSNet 为基线进行消融实验，构建以下四组对比模型：(1) DSNet：原始双分支子像素引导网络；(2) DSNet + SE：在空间 - 光谱分支中嵌入 SE 通道注意力模块；(3) DSNet + Focal Loss：以 Focal Loss 替代交叉熵损失；(4) SE-DSNet：同时引入 SE 与 Focal Loss 的完整模型。所有实验保持相同数据划分与超参数设置，仅调整对应模块，通过对比分析各改进对分类性能的单独及综合影响。

4.4.1. Indian Pines 数据集消融结果分析

消融实验(表 7、表 8)显示，在 Indian Pines 上，原始 DSNet 的 OA 为 93.61%、Kappa 为 92.69%。单独加入 SE 模块(DSNet + SE)后，OA 提升至 95.20%，Kappa 升至 94.51%；单独使用 Focal Loss (DSNet + Focal Loss)则 OA 为 95.01%，Kappa 为 94.25%。完整模型 SE-DSNet 取得最佳效果(OA 96.72%, Kappa 96.24%)。各类别中，SE 模块显著改善了牧场草地(91.39% → 99.71%)和修剪牧场草地(93.90% → 98.37%)等类别；Focal Loss 对苜蓿、燕麦等小样本类别亦有提升。两者结合后，模型在多数类别上达到或接近满分，整体性能优于任一单模块，验证了两种改进策略在农业场景下的协同有效性。

Table 7. Overall and ablation experiment results for Indian Pines dataset (%)
表 7. Indian Pines 数据集总体结果消融实验结果(%)

方法	OA (%)	AA (%)	Kappa
DSNet	93.61	96.18	92.69
DSNet + SE	95.20	97.74	94.51
DSNet + Focal Loss	95.01	97.14	94.25
SE-DSNet (本文方法)	96.72	98.05	96.24

Table 8. Classification and ablation experiment results for Indian Pines dataset (%)
表 8. Indian Pines 数据集分类结果消融实验结果(%)

类别	DSNet	DSNet + SE	DSNet + Focal Loss	SE-DSNet
苜蓿	90.03	90.90	86.92	93.21
免耕玉米	95.92	91.45	96.56	95.41
少耕玉米	98.37	100.00	100.00	98.91
玉米	96.42	97.99	97.54	98.66
牧场草地	91.39	99.71	99.00	99.71
乔木草地	97.95	100.00	100.00	100.00
修剪牧场草地	93.90	98.47	86.06	98.37
风干干草	88.96	91.77	96.73	95.45
燕麦	95.92	96.99	96.10	95.04

续表

免耕大豆	99.38	100.00	100.00	100.00
少耕大豆	99.38	91.77	97.83	98.71
清耕大豆	99.38	96.99	100.00	97.88
小麦	99.38	100.00	100.00	100.00
林地	94.87	100.00	97.44	97.44
建筑－草地－树木－道路混合区	100.00	100.00	100.00	100.00
砖石－钢结构塔架	100.00	100.00	100.00	100.00

4.4.2. Augsburg 数据集消融结果分析

在 Augsburg 数据集(表 9、表 10)上，DSNet 的 OA 为 88.34%、AA 为 64.57%。加入 SE 后 OA 提升至 91.20%，Kappa 升至 87.35%，主类别(森林、低矮植被)精度明显提高，但 AA 略降至 63.71%。单独使用 Focal Loss 使 AA 提升至 69.08% (最高)，水体(42.27% → 50.70%)和分配地(51.43% → 60.04%)等困难类别改善明显，但 OA 提升有限。完整 SE-DSNet 的 OA 为 90.47%，Kappa 为 86.27%，在森林(97.10%)、住宅区(96.16%)上保持高精度，但工业区精度(51.51%)低于 DSNet + Focal Loss (81.04%)，说明多模块联合训练在极端复杂类别上存在权衡。总体而言，SE 侧重主类特征，Focal Loss 利于类别均衡，两者结合在整体性与均衡性间取得折中。

Table 9. Overall and ablation experiment results for Augsburg dataset (%)
表 9. Augsburg 数据集总体结果消融实验结果(%)

方法	OA	AA	Kappa
DSNet	88.34	64.57	83.20
DSNet + SE	91.20	63.71	87.35
DSNet + Focal Loss	90.32	69.08	86.24
SE-DSNet (本文方法)	90.47	66.21	86.27

Table 10. Classification and ablation experiment results for Augsburg dataset (%)
表 10. Augsburg 数据集分类结果消融实验结果(%)

类别	DSNet	DSNet + SE	DSNet + Focal Loss	SE-DSNet
森林	93.56	98.66	95.34	97.10
住宅区	94.47	95.35	92.43	96.16
工业区	65.74	75.07	81.04	51.51
低矮植被	89.93	95.07	94.58	93.72
分配地	51.43	55.83	60.04	57.17
商业区	14.59	4.64	9.40	18.19
水体	42.27	21.43	50.70	49.64

4.4.3. 消融实验结果分析

综合两数据集，消融实验表明：SE 注意力机制主要增强主类别特征表达，提升整体精度和区域连续性；Focal Loss 更有效改善小样本和难分类类别，提升类别均衡性；两者联合在农业场景中协同良好，精

度稳定提升；在城市场景中，由于类别结构复杂且极不均衡，模块间存在一定性能权衡(如工业类)。尽管如此，SE-DSNet 在两个场景中均取得最优综合性能，验证了 SE 和 Focal Loss 在增强特征判别与缓解不平衡方面的有效性，同时也提示后续需进一步优化多模块融合策略以应对极端异质性类别。

5. 总结与展望

5.1. 总结

本文针对高光谱图像分类中混合像元、光谱冗余和类别不平衡等问题，在 DSNet 的基础上引入了 SE 注意力机制和 Focal Loss，提出了 SE-DSNet 模型。主要工作及结果如下：

(1) 复现了 DSNet 模型，验证了其双分支结构对空间特征和丰度信息的联合建模能力。在 Indian Pines 和 Augsburg 两个数据集上，DSNet 的表现均优于传统 CNN 和 GRU 方法。

(2) 在双分支网络中嵌入 SE 通道注意力模块，使模型能够自适应增强关键波段特征并抑制冗余信息。消融实验表明，加入 SE 模块后，模型在主类别上的分类精度和区域完整性均有明显提升。

(3) 采用 Focal Loss 替代交叉熵损失，有效缓解了类别不均衡问题。单独使用 Focal Loss 时，模型在少数类和难分类样本上的平均精度显著提高。

(4) 完整模型 SE-DSNet 在两个数据集上均取得了最优性能。相比原始 DSNet，整体精度、平均精度和 Kappa 系数均有稳定提升。可视化结果也显示，SE-DSNet 的分类图最接近真实标注，区域连续性和边界清晰度均优于对比方法。

消融实验进一步表明，SE 模块侧重提升主类别的特征表达，Focal Loss 更有利于类别均衡，两者结合后在整体性能和类别平衡之间取得了较好的折中。

5.2. 展望

本文方法仍存在一些不足，未来可以从以下几个方面继续改进：

(1) 轻量化设计。当前模型参数较多，可尝试模型剪枝、深度可分离卷积或知识蒸馏，便于在边缘设备上部署[26]。

(2) 多尺度特征提取。固定尺寸的感受野难以适应不同大小的地物，可引入空洞卷积或特征金字塔，增强模型对多尺度目标的适应能力。

(3) 自监督与对比学习。高光谱数据标注成本高，探索利用无标签数据进行自监督预训练，减少对标注样本的依赖。

(4) 跨场景迁移。不同数据集之间存在光谱漂移，研究域自适应或域泛化方法，提高模型在不同场景下的迁移能力。

(5) 对极端类别的鲁棒性。Augsburg 数据集中工业区的分类精度波动较大，后续可设计类别自适应的损失函数或注意力机制，增强对高异质性类别的识别能力。

(6) 多模态融合。结合高分辨率全色图像、LiDAR 或 SAR 数据，利用多模态信息进一步提升分类精度，尤其在复杂城市场景中。

参考文献

[1] Chen, Y., Lin, Z., Zhao, X., Wang, G. and Gu, Y. (2014) Deep Learning-Based Classification of Hyperspectral Data: A Comprehensive Review. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **7**, 2094-2107. https://doi.org/10.1109/jstars.2014.2329330

[2] Wu, H. and Prasad, S. (2018) Semi-Supervised Deep Learning Using Pseudo Labels for Hyperspectral Image Classification. *IEEE Transactions on Image Processing*, **27**, 1259-1270. https://doi.org/10.1109/tip.2017.2772836

[3] Zhang, L., Zhang, L. and Du, B. (2016) Deep Learning for Remote Sensing Data: A Technical Tutorial on the State of the Art. *IEEE Geoscience and Remote Sensing Magazine*, **4**, 22-40. https://doi.org/10.1109/mgrs.2016.2540798

[4] Roy, S.K., Krishna, G., Dubey, S.R. and Chaudhuri, B.B. (2020) Hybridsn: Exploring 3-D-2-D CNN Feature Hierarchy for Hyperspectral Image Classification. *IEEE Geoscience and Remote Sensing Letters*, **17**, 277-281. https://doi.org/10.1109/lgrs.2019.2918719

[5] Hang, R., Liu, Q., Hong, D. and Ghamisi, P. (2019) Cascaded Recurrent Neural Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **57**, 5384-5394. https://doi.org/10.1109/tgrs.2019.2899129

[6] Ghaderizadeh, S., Abbasi-Moghadam, D., Sharifi, A., Tariq, A. and Qin, S. (2022) Multiscale Dual-Branch Residual Spectral-Spatial Network with Attention for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **15**, 5455-5467. https://doi.org/10.1109/jstars.2022.3188732

[7] Rasti, B., Hong, D., Hang, R., Ghamisi, P., Kang, X., Chanussot, J., *et al*. (2020) Feature Extraction for Hyperspectral Imagery: The Evolution from Shallow to Deep: Overview and Toolbox. *IEEE Geoscience and Remote Sensing Magazine*, **8**, 60-88. https://doi.org/10.1109/mgrs.2020.2979764

[8] Li, W., Wu, G., Zhang, F. and Du, Q. (2017) Hyperspectral Image Classification Using Deep Pixel-Pair Features. *IEEE Transactions on Geoscience and Remote Sensing*, **55**, 844-853. https://doi.org/10.1109/tgrs.2016.2616355

[9] Yang, X., Ye, Y., Li, X., Lau, R.Y.K., Zhang, X. and Huang, X. (2018) Hyperspectral Image Classification with Deep Learning Models. *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 5408-5423. https://doi.org/10.1109/tgrs.2018.2815613

[10] He, K., Zhang, X., Ren, S. and Sun, J. (2016) Deep Residual Learning for Image Recognition. 2016 *IEEE Conference on Computer Vision and Pattern Recognition* (*CVPR*), Las Vegas, 27-30 June 2016, 770-778. https://doi.org/10.1109/cvpr.2016.90

[11] Zhang, L., Wang, Y., Yang, L.Z., *et al*. (2023) A Multi-Range Spectral-Spatial Transformer for Hyperspectral Image Classification. *Infrared Physics & Technology*, **135**, Article ID: 104983. https://doi.org/10.1016/j.infrared.2023.104983

[12] Hong, D., Gao, L., Yao, J., Zhang, B., Plaza, A. and Chanussot, J. (2021) Graph Convolutional Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **59**, 5966-5978. https://doi.org/10.1109/tgrs.2020.3015157

[13] Jia, S., Jiang, S., Lin, Z., Li, N., Xu, M. and Yu, S. (2021) A Survey: Deep Learning for Hyperspectral Image Classification with Few Labeled Samples. *Neurocomputing*, **448**, 179-204. https://doi.org/10.1016/j.neucom.2021.03.035

[14] Dobigeon, N., Tourneret, J., Richard, C., Bermudez, J.C.M., McLaughlin, S. and Hero, A.O. (2014) Nonlinear Unmixing of Hyperspectral Images: Models and Algorithms. *IEEE Signal Processing Magazine*, **31**, 82-94. https://doi.org/10.1109/msp.2013.2279274

[15] Han, Z., Yang, J., Gao, L.R., *et al.* (2024) Dual-Branch Subpixel-Guided Network for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **62**, 1-13. https://doi.org/10.1109/tgrs.2024.3418583

[16] Lin, T., Goyal, P., Girshick, R., He, K. and Dollar, P. (2017) Focal Loss for Dense Object Detection. 2017 *IEEE International Conference on Computer Vision* (*ICCV*), Venice, 22-29 October 2017, 2980-2988. https://doi.org/10.1109/iccv.2017.324

[17] Zhao, L., Luo, W., Liao, Q., Chen, S. and Wu, J. (2022) Hyperspectral Image Classification with Contrastive Self-Supervised Learning under Limited Labeled Samples. *IEEE Geoscience and Remote Sensing Letters*, **19**, 1-5. https://doi.org/10.1109/lgrs.2022.3159549

[18] Wu, Q., He, M., Liu, Z. and Liu, Y. (2024) Multi-Scale Spatial-Spectral Residual Attention Network for Hyperspectral Image Classification. *Electronics*, **13**, Article 262. https://doi.org/10.3390/electronics13020262

[19] Chen, J., Li, W. and El-Askary, H. (2025) Spectral-Spatial Transformer with Multiscale Convolutional Attention for Hyperspectral Image Classification. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, **18**, 24223-24241. https://doi.org/10.1109/jstars.2025.3608699

[20] Xu, Y., Zhang, L., Du, B. and Zhang, F. (2018) Spectral-Spatial Unified Networks for Hyperspectral Image Classification. *IEEE Transactions on Geoscience and Remote Sensing*, **56**, 5893-5909. https://doi.org/10.1109/tgrs.2018.2827407

[21] Meng, Z., Yan, Q., Zhao, F. and Liang, M. (2023) Hyperspectral Image Classification with Dynamic Spatial-Spectral Attention Network. 2023 13*th Workshop on Hyperspectral Imaging and Signal Processing*: *Evolution in Remote Sensing* (*WHISPERS*), Athens, 31 October 31-2 November 2023, 1-4. https://doi.org/10.1109/whispers61460.2023.10431375

[22] Shu, Z., Wang, Y. and Yu, Z. (2024) Dual Attention Transformer Network for Hyperspectral Image Classification. *Engineering Applications of Artificial Intelligence*, **127**, Article ID: 107351. https://doi.org/10.1016/j.engappai.2023.107351

[23] Loshchilov, I. and Hutter, F. (2019) Decoupled Weight Decay Regularization. arXiv: 1711.05101

[24] Kingma, D.P. and Ba, J. (2015) Adam: A Method for Stochastic Optimization. arXiv: 1412.6980 https://arxiv.org/abs/1412.6980

[25] Ioffe, S. and Szegedy, C. (2015) Batch Normalization: Accelerating Deep Network Training by Reducing Internal Covariate Shift. 32*nd International Conference on Machine Learning*, Lille, 6-11 July 2015, 448-456.

[26] Chen, L., Wei, Z. and Xu, Y. (2020) A Lightweight Spectral-Spatial Feature Extraction and Fusion Network for Hyperspectral Image Classification. *Remote Sensing*, **12**, Article 1395. https://doi.org/10.3390/rs12091395