

摘要

由于高光谱图像(Hyperspectral Image, HSI)具有丰富的光谱信息，已成为遥感(Remote Sensing, RS)领域的研究热点。然而空间分辨率低导致的混合像元问题不仅严重限制了HSI的处理精度，还阻碍了HSI分析技术的应用及发展。高光谱解混(Hyperspectral Unmixing, HU)是处理混合像元的有效方法，主要目的是突破图像空间分辨率限制，提取端元信息并估计其相应丰度。

随着深度学习在HU领域的广泛应用，各种基于深度学习的解混算法随之提出，然而这些方法都存在对光谱-空间特征信息运用不充分的问题。为此，本文提出一种基于双注意力卷积网络(Dual Attention Convolutional Network, DACN)的端到端解混算法，主要工作包括以下三个方面：

- (1)引入通道注意力模块(Channel Attention Module)与空间注意力模块(Spatial Attention Module)以更好地提取光谱-空间特征信息；
- (2)考虑到HSI中单个像素光谱波段的序列性，采用层标准化(Layer Normalization, LN)在每层对单个样本的各个维度特征做标准化处理，避免过拟合现象的同时提高训练速度与精度；
- (3)由于激活函数ReLU的输入值为负时，其输出与一阶导数始终为零，导致神经元参数不能正常更新。为此，采用Leaky ReLU来避免神经元失活。

在Jsaper Ridge和Urban两个高光谱数据集上进行了完整试验表明，与当前已有的解混算法相比，本文提出的DACN能更有效地提取光谱-空间特征信息，解混精度有较为显著的提高。

关键词：高光谱解混；注意力机制；光谱-空间信息；卷积神经网络

ABSTRACT

Hyperspectral Image (HSI) have rich spectral information and have become a research frontier in Remote Sensing (RS). However, the mixed pixels element problem caused by low spatial resolution not only severely limits the processing accuracy of HSI, but also hinders the application and development of HSI analysis technology. Hyperspectral Unmixing (HU) is an effective method for processing mixed pixels, with the main purpose of breaking the spatial resolution limitation of images, extracting end member information and estimating the corresponding abundance.

In recent years, with the wide application of deep learning in the field of HU, various deep learning-based unmixing algorithms have emerged. However, all these methods suffer from the problem of inadequate use of spectral-spatial feature information. To this end, an end-to-end unmixing algorithm based on Dual Attention Convolutional Network (DACN) is proposed in this paper, and the main work include the following three aspects.

(1) Better extraction of spectral-spatial feature information by introducing Channel Attention Module and Spatial Attention Module.

(2) Considering the sequential nature of individual pixel spectral bands in HSI, Layer Normalization is used to normalize each dimensional feature of a single sample at each layer to avoid overfitting and improve the training speed and accuracy.

(3) Since the output of the traditional activation function ReLU is always zero when its input value is negative, and the first-order derivative is always zero, the neuron parameters cannot be updated properly. For this reason, Leaky ReLU is used to avoid neuron deactivation.

Complete experiments on two hyperspectral datasets, Jasper Ridge and Urban, show that the DACN proposed in this paper can extract spectral-spatial feature information more effectively and has a more significant improvement in unmixing accuracy compared with currently available unmixing algorithms.

Key words: Hyperspectral Unmixing; Attention Mechanism; Spectral-Spatial Information; CNN

目录

1. 绪论.....	1
1.1 研究背景及意义.....	1
1.2 国内外研究现状.....	2
1.2.1 基于线性光谱混合模型的高光谱解混.....	3
1.2.1 基于非线性光谱混合模型的高光谱解混.....	4
1.3 主要研究内容及论文结构.....	5
1.3.1 主要研究内容.....	5
1.3.2 论文组织结构.....	5
2. 相关理论基础.....	7
2.1 高光谱混合模型.....	7
2.2.1 线性光谱混合模型.....	7
2.2.2 非线性光谱混合模型.....	8
2.2 卷积神经网络.....	8
2.1.1 卷积神经网络特点.....	9
2.1.2 卷积神经网络结构.....	9
2.3 注意力机制.....	9
2.2.1 注意力机制分类.....	11
2.2.2 相关注意力机制.....	11
2.4 基于深度学习的高光谱解混.....	13
2.4.1 基于 AANN 的高光谱解混.....	13
2.4.2 基于 CNN 的高光谱解混.....	13
2.4.3 基于 FNNC 的高光谱解混.....	14
2.5 本章小结.....	15
3. 基于双注意力卷积网络的高光谱解混.....	17
3.1 引言.....	17
3.2 双注意力卷积网络.....	17
3.2.1 网络结构.....	18
3.2.2 网络设置.....	20
3.3 本章小结.....	21
4. 试验设计与结果分析.....	23
4.1 性能评价指标.....	23
4.2 试验数据及环境介绍.....	23
4.2.1 数据集介绍.....	24

4.2.2 实验环境及参数设置.....	24
4.3 试验分析.....	25
4.4 本章小结.....	28
5.总结与展望.....	29
5.1 总结.....	29
5.2 展望.....	29
致谢.....	31
参考文献.....	33

1.绪论

1.1 研究背景及意义

高光谱遥感成像技术的不断完善与飞速发展，加快了利用遥感技术探测与理解事物能力的速度。高光谱遥感作为光学遥感领域的前沿技术，与其他传统遥感探测方法相比，它将检测对象的光谱信息与空间信息高度融合，实现了“图谱合一”，并解决了“光谱不成像”与“成像无光谱”的问题^[1]。在高光谱成像中，高光谱的各个波段分别成像并叠加在一起以形成三维立方体数据，即高光谱图像(Hyperspectral Image, HSI)^[2]，如图 1-1 所示。由于 HSI 需要在保持二维空间中像元分布(光谱分辨率高)的同时，提供每个像元在光谱维度上的散射特征(光谱波段多)，所以 HSI 具有立方体式的数据结构，因此能够同时获取地物的光谱-空间信息。

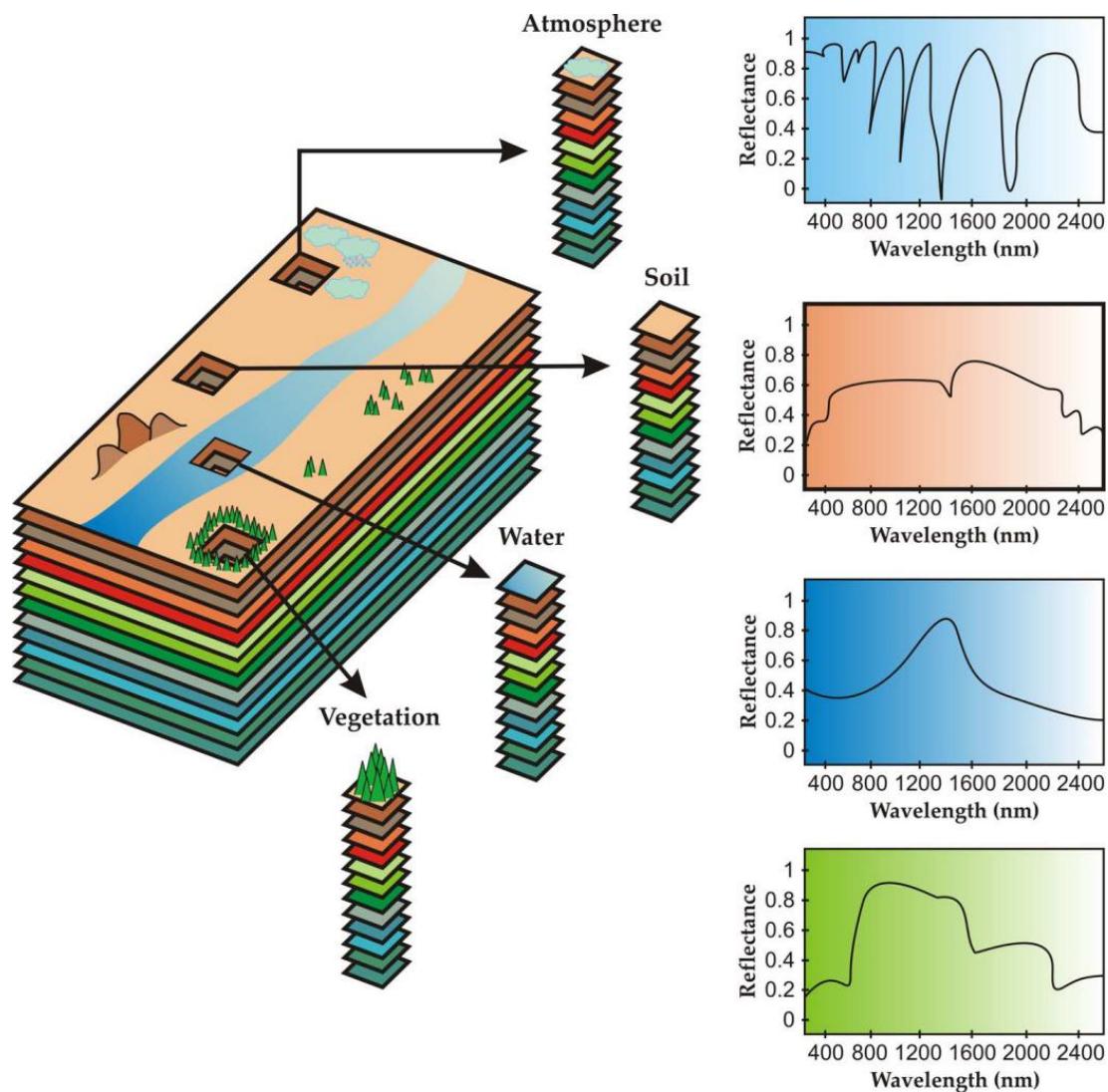


图 1-1 高光谱成像示意图

虽然 HSI 具有较高的光谱分辨率、并且能够获得地物更为详细的光谱信息，

但是由于 HSI 的空间分辨率较低，使得在实际成像过程中图像分辨率不足，直接导致像元中包含多种地物信息，这样的像元被称为混合像元^{[3][4]}。其中，纯像元(Pure Pixel)中只含有一种类型的地物信息。每个端元具有不同的光谱特征，通常称为端元特征。端元作为像元的基本组成部分，其比例信息通常称为丰度值。

混合像元的存在限制了 HSI 的精确分析及应用，对 HSI 像元级别的分类精度、目标探测效果与解码质量产生了很大的影响，并严重妨碍了高光谱数据量化的发展。因而，如何处理混合像元分解问题，即高光谱解混(Hyperspectral Unmixing, HU)问题，成为 HSI 分析与应用的关键。目前对 HSI 处理的研究，如 HSI 分类、地物识别等，都依赖于 HU 问题的有效解决。

1.2 国内外研究现状

HU 是处理混合像元最有效的方式，它能够突破 HSI 空间分辨率的限制，展示出混合像元的真实属性，从而提高 HSI 解混精度，并将其应用到遥感领域。此外，提取混合像元中所蕴含的基本特征称为端元提取，计算每个基本特征的比例称为丰度估计。如图 1-2 所示为 HU 的整个流程示意图。

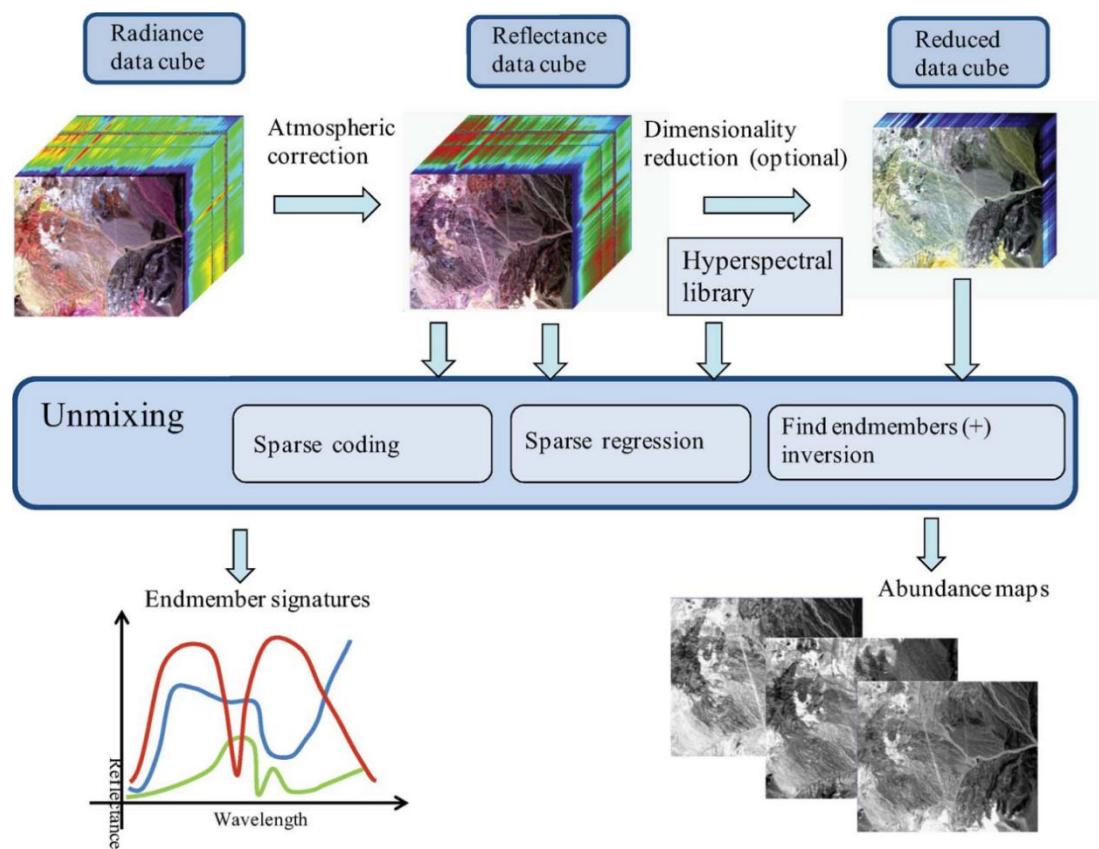


图 1-2 高光谱解混流程图

为了构建 HU 模型，根据高光谱的混合模式，高光谱混合模型可分为线性光谱混合模型(Linear Spectral Mixture Model, LSMM)^{[8][9]}与非线性光谱混合模型(Nonlinear Spectral Mixture Model, NLSMM)^{[10][11][12]}。然而在实际应用中，

NLSMM 存在构建过程复杂、分析难度大等问题，因此通常应用于特定复杂场景下的建模。LSMM 具有明确的物理概念表述以及简便快捷等优势，能较好地描述实际光谱混合现象。因此，LSMM 在 HU 中得到了普遍的应用。

1.2.1 基于线性光谱混合模型的高光谱解混

在所有的基于 LSMM 方法中，几何学方法最为丰富，研究者提出了很多经典算法。Boardman^[13]提出的像元纯度指数 (Pixel Purity Index, PPI) 算法，首先将最大噪声分量变换作为降维与提高信噪比的预处理步骤，然后将图像中每个像元看作单个向量，并将其投影到随机向量上，之后将投影数量最多的向量选择为图像的端元。虽然该方法实施方便、效率较高，然而 PPI 在对随机向量的选择上有很大的随机性，具有很强的主观性。

在基于单形体的几何方法中，由于单形体顶点与端元之间相互对应，端元提取相当于寻找单形体的顶点。Winter^[14]提出的基于单形体的内部最大体积分析 (N-FINDR) 算法，首先在光谱维度上计算得到所有单纯形的体积，然后求出单形体的最大体积。由于单形体的顶点是纯像元，从而可以自动获得图像端元。随后 Nascimento^[15]提出的顶点成分分析 (Vertex Component Analysis, VCA) 算法，首先确定一个端元，然后将像元向量投影到已选端元的正交子空间上，则投影向量的端元即为新端元。虽然这些解混算法简单易用，但是需要 HSI 满足纯像元的假设，然而在实际应用中这种假设通常是不存在的。

统计分析方法对于 HSI 具有较高的解混精度，是近年来国内外前沿研究方向之一。贝叶斯分析^[16]方法是基于假设的先验概率，事先假定端元组分符合某种概率分布，并把这种先验概率分布输入到后验概率模型中，进而获得端元组分。独立成分分析 (Independent Component Analysis, ICA)^[17]通过将光谱向量与能使信道间相互信息最小化的解混矩阵相乘得到端元，ICA 假设各地物之间组分相互独立，则它能获得准确的解混结果，因为相互信息的最小值对应且仅对应于独立的组分。然而，真实的地物分布并不满足这个假设。为了克服 ICA 的局限性，Nascimento^[18]提出相关成分分析 (Dependent Component Analysis, DECA)，DECA 假设每个像元是端元信号的线性组合，且用相应的丰度分数加权。这些丰度系数满足狄利克雷分布，从而使丰度系数满足“非负性”及“和为一性”。

Lee^[20]提出的非负矩阵分解技术 (Nonnegative Matrix Factorization, NMF) 能够将高阶信号矩阵分解为两个低阶非负矩阵相乘的形式。即 NMF 可以同时实现端元提取与丰度估计，因而该算法得到了广泛的关注。然而由于函数本身所具有的非凸性，使得 NMF 容易陷入局部最优的尴尬局面，这将使得解混结果产生影响。为此，通常将结构性约束加入到 NMF 模型中。Hoyer^[22]将稀疏性约束引入到 NMF 的算法中，提出了带有稀疏约束的 NMF。Qian Y^[23]将 $L_{1/2}$ 正则化项引入

NMF 中，提出了一种 $L_{1/2}$ 稀疏约束的 NMF 的解混算法。该算法获得的结果比使用 L_1 范数获得的结果更加稀疏和准确。

近年来，基于空间相关的 HU 取得了巨大进展。祝伟^[24]在 NMF 模型中引入全变差正则化项，提出了结合稀疏与全变差的 NMF 的解混算法。为了描述高光谱数据内部的几何结构，甘玉泉^[25]引入了图正则化理论，进而提出了一种稀疏图正则化 NMF 的解混算法。上面两种算法由于引入了空间信息，因而获得了比仅仅加入稀疏约束更好的性能。然而，这些算法没有充分利用光谱和空间两方面的特征，因此还有一定的改进空间。彭倩^[26]提出了联合光谱空间预处理与约束 NMF 的解混算法，有效地提高了约束 NMF 的解混精度。结合光谱相关性特征，袁博^[27]提出了空-谱非负矩阵分解的解混算法。该算法可以实现对高光谱数据进行高精度解混，然而准确描述高光谱数据的结构信息，必须建立光谱-空间之间的关系，这仍然是研究者面临的难题^[8]。

1.2.1 基于非线性光谱混合模型的高光谱解混

辐射传输理论 (Radiative Transfer Theory, RTT)^[28] 是一种光子与场景材料相互结合的能量转移模型。然而，由于问题的复杂性使得其在真实场景中很难获取参数。另一类被广泛应用的策略是使用更简单但受物理启发的模型来避免复杂的物理模型，基于核的方法就是此类方法。Broadwater^{[29][30][31][32]}等提出了几种基于核的解混算法来专门解决致密混合问题，其中一些核被设计的足够灵活以允许几中非线性等级(例如使用径向基函数或多项式展开)，而另一些则是受物理启发而设计的核。相反，在[33][34][35]中，双线性光谱混合模型被相继提出用以处理类似于在多层场景中发生的散射产生的影响。通过附加条件来扩展线性模型，它们之间的主要通过可加性约束相互区分。

为了使算法更灵活，一些研究者采取了机器学习策略，例如基于神经网络的解混算法，其以有监督方式从一组样例中进行非线性降维或学习模型参数。Altmann^[43]在文中引入的多项式后 NLSMM 也具有足够的通用性，其可涵盖一大类非线性模型。然而，这些算法假定了端元的先验知识或事先进行了端元提取。受限于问题的难度，很少有研究人员尝试来解决完全无监督的非线性解混问题。然而，Heylen^[44]引入了一种策略来提取非线性混合的端元，该算法在许多方面与 N-FINDR 算法很相似。其关键思想在于通过计算数据流形上的测地距离来最大化单纯形体积。在这项工作中，精确的测地距离由最近邻图中的最短路径距离近似。在他们后来的研究中显示，通过 NLSMM，例如广义双线性模型的诱导可以在任意数据流形上推导出精确的测地距离。

Close^[46]与 Gader^[47]根据上述平均反照率模型，提出了两种完全无监督的非线性解混算法，用于致密性混合物的场景中。在[31][32][48][49]提供的评价指标下，

该方法可以更准确地估算物理混合参数。此研究中仍有大量工作需要做,包括评估将双线性模型与平均反照率模型相结合的有效性。尽管研究人员已对 NLSMM 做了不少的研究,但是与 LSMM 相比,其模型复杂、参数众多且研究还不够成熟。

1.3 主要研究内容及论文结构

1.3.1 主要研究内容

本文主要以 HU 为基本任务要点,致力于将深度学习与 HU 相结合,提高解混精度,提出适用于 HSI 的双注意力卷积网络,通过结合光谱信息以及空间信息共同提高解混精度。本文的主要工作安排如下:

- (1) 提出了一种基于双注意力卷积网络(DACN)的端到端解混算法,该方法通过在卷积神经网络提取特征的基础上增加了两种类型的注意力模块,对光谱维度和空间维度上的语义信息进行建模,以自适应地融合局部特征与全局特征;
- (2) 在 Jsaper Ridge 与 Urban 两个高光谱数据集上进行了完整的性能评估,与现有的解混算法进行对比,并分析试验效果。

1.3.2 论文组织结构

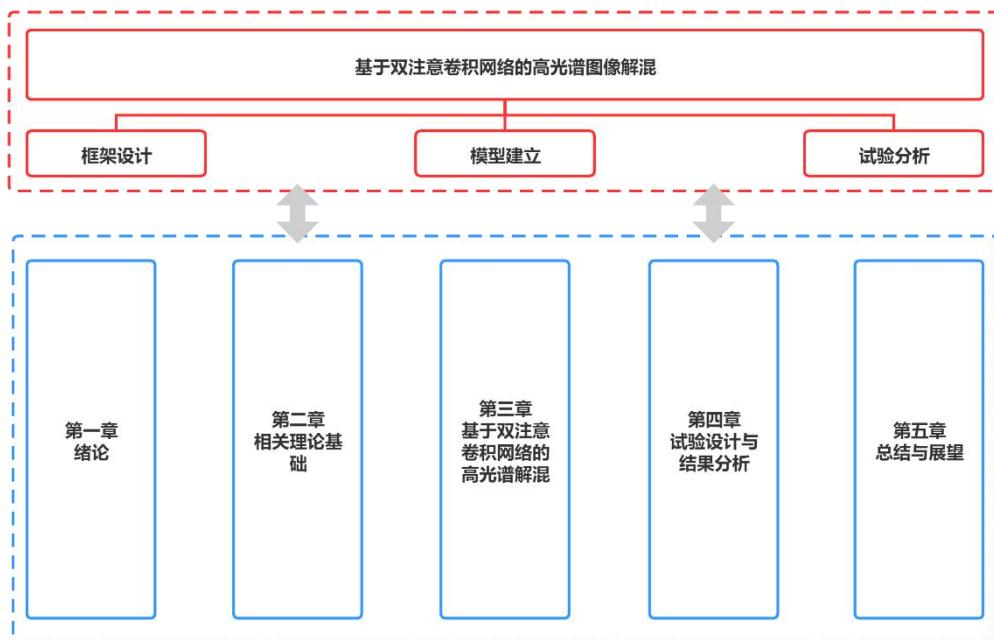


图 1-3 论文组织结构示意图

如图 1-3 所示全文共分为五个章节,各个章节具体细节如下:

第一章: 主要阐述了 HSI 的研究背景、意义以及国内外在 HU 领域的发展现状,并在章节末尾介绍了论文结构与各章的内容;

第二章: 主要阐述了高光谱混合模型的基本概念以及深度学习的相关概念,并在章节末尾介绍了三种基于深度学习的解混算法;

第三章：提出了基于双注意力卷积网络(DACN)的端到端解混算法，针对 HSI 的特性分别在光谱维度以及空间维度上引入注意力模块，充分利用 HSI 的光谱-空间信息以提高解混精度；

第四章：介绍了试验所用的性能评价指标、数据集、试验环境以及参数设置等。通过在 Jsaper Ridge 和 Urban 两个高光谱数据集上进行了完整的性能评估。与现有的解混算法相比，DACN 能更有效地提取光谱-空间特征信息，解混精度有较为显著的提高；

第五章：总结了本论文的所提出的方法及所做的工作，并对下一步需要重点关注的改进方向进行展望。

2.相关理论基础

2.1 高光谱混合模型

HU 算法依赖于特定的高光谱混合模型，该模型用来描述场景中的物质如何相互作用并形成像元内的合成光谱^[8]。设 $\{y_i\}_{i=1}^n$ 为 HSI 中一组具有 n 个像元的向量集，且每个像元 y_i 中记录了观测到的光谱向量，由 $[0, 1]$ 范围内的 L 个光谱段组成，即 $\forall i: y_i \in [0, 1]^L$ 。 $\{a_j\}_{j=1}^m$ 表示对任意给定光谱向量 y 可由 m 个端元的基本光谱组成，其中每个端元 a_j 都具有一个关联丰度 x_j 表示对应端元在混合像元中所占的百分比。混合像元由端元及其对应丰度构成，其构成关系可以使用不同的约束条件。混合模型通常可表示为：

$$y = F(A, x) \quad (2-1)$$

式中，矩阵 A 由端元与对应丰度列向量组成，函数 F 表示通过端元及其丰度估计构建光谱信号。

在经典的混合场景中，通常认为端元是整幅图像的同一组光谱信号，而丰度将基于不同混合像元而变化。一般存在以下几种通用方法：如果通过其它研究、地物真实数据或从光谱数据库中获得了端元，则解混可以归结为反演混合模型。根据混合函数 F 的细节，有若干种方法可供采用。另一方面，如果端元未知，则必须从 HSI 本身派生它们。大多数解混算法都是先采用端元提取算法提取端元光谱，然后执行反演步骤，然而某些算法也可同时找到端元和丰度。HU 问题的另一个问题是端元的数目通常是未知的，并且必须通过某种方式对其进行估计。

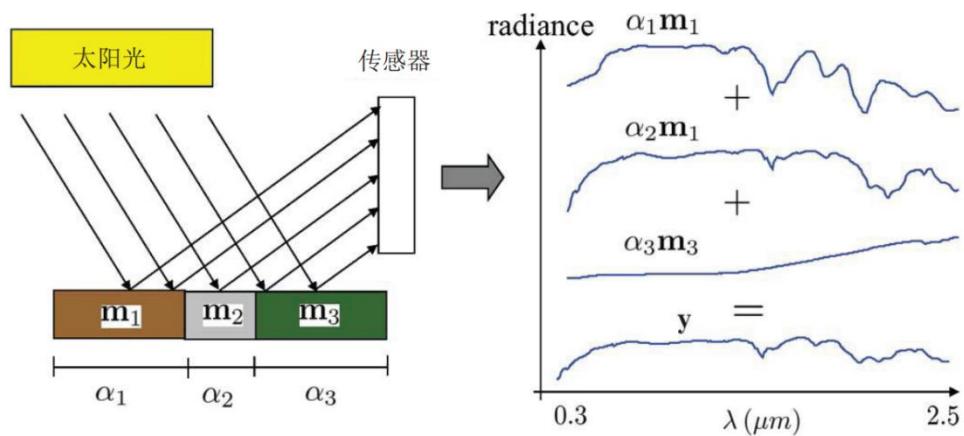


图 2-1 线性光谱混合模型

2.2.1 线性光谱混合模型

根据物质的混合程度以及地物分布的空间尺度，现有的光谱混合模型主要分为线性光谱混合模型 (Linear Spectral Mixing Model, LSMM) 与非线性光谱混合

模型 (Nonlinear Spectral Mixing Model, NLSMM)^[50]。LSMM 是一种备受欢迎且可简单易用的光谱混合模型，LSMM 假设混合规模是宏观的且入射光仅与一种物质发生相互作用。然而，由物质反射的光尽管几乎完全分开，但在成像光谱仪中混合在一起。线性光谱混合模型相关描述如图 2-1 所示：该模型将三种物质的散射光入射到含有 L 个光谱波段的传感器上。设 $Y = (y_1, y_2, \dots, y_k) \in R^{L \times K}$ 为 K 个像素 L 个光谱波段下的光谱图像，其中光谱信号 y 是物质光谱的加权平均值。 $\theta = (\theta_1, \theta_2, \dots, \theta_N) \in R^{L \times N}$ 为含有 N 个端元的端元谱库，其中 $\theta_i (i = 1, 2, \dots, N)$ 为第 i 个端元的 L 个波段的光谱，线性光谱混合模型可描述为：

$$Y = \theta W + E \quad (2-2)$$

式中 E 为高斯噪声， Y 为观测谱矩阵， θ 为端元谱矩阵， W 为丰度矩阵。

2.2.2 非线性光谱混合模型

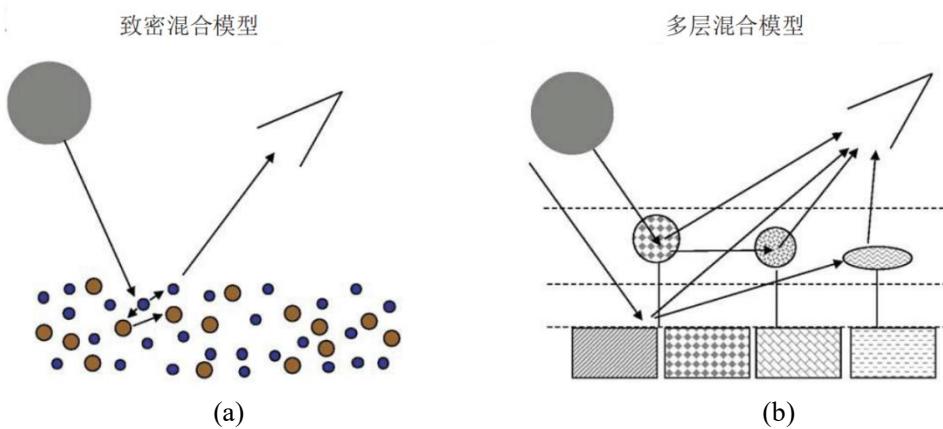


图 2-2 非线性光谱混合模型

然而，NLSMM 通常是由于场景中多种物质散射光之间的相互作用造成的。当光从一个或多个目标物质散射时，被其它目标物质反射并最终由成像光谱仪接收时，就会发生混合现象。Borel^[52]证明了多层模型可表示为光谱反射率乘积的无穷次幂。当两种物质均匀混合时会发生微观混合现象。在这种情况下，相互作用包括从一种物质发出的光子被另一种物质吸收，后者又反射更多的光子。图 2-2 对这两种非线性光谱混合模型进行了说明，图 2-2(a) 表示致密混合现象，其场景中的物质材料非常相似。图 2-2(b) 显示了一种多层混合现象，其中不同层的散射之间存在多种相互作用。

LSMM 具有模型简单、物理意义明确、计算易处理和灵活性好等优势，是目前使用最广泛的解混模型。尽管如此，但在许多实际场景中仍是一个可接受的光散射机制的近似值。

2.2 卷积神经网络

人工神经网络 (Artificial Neural Networks, ANN)^[54] 是通过调整内部神经元之间相互作用方式对信息进行处理，卷积神经网络 (Convolutional Neural Network,

CNN) 则是对 ANN 的进一步优化。Hubel^[55]首次提出了感受野 (Receptive Field, RF) 的概念, 它是 CNN 每层输出的特征图上的像素点在输入图片上映射的区域大小。基于此, Fukushima^[56]提出了感知机与权重共享的卷积层, 被广泛认为是 CNN 的雏形。LeCun^[57]通过将反向传播算法与卷积层相结合提出了 CNN, 并首次应用于手写数字的识别。

2.1.1 卷积神经网络特点

CNN 由多层感知机 (Multilayer Perceptron, MLP) 演化所得到的, 由于其具有两个关键优势: 局部连接与权值共享, 使得 CNN 在图像处理领域表现突出。局部连接有别于传统神经网络, 每个神经元只与上一层的局部区域连接, 连接的空间大小可视为神经元的感受野。权值共享使得当前层在深度方向上每层神经元都使用同样的权重和偏差。局部连接与权值共享相结合, 有效降低了参数数量, 同时使得训练复杂度骤降并且降低了过拟合现象发生的概率。

2.1.2 卷积神经网络结构

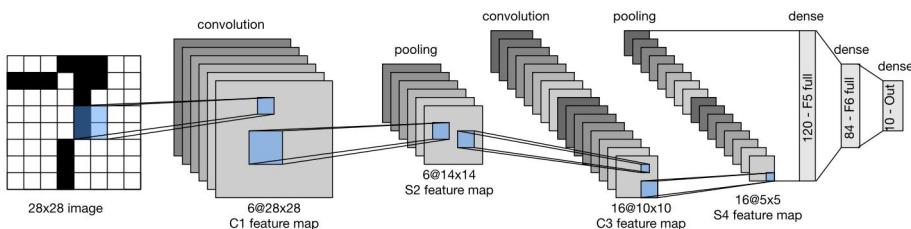


图 2-3 LeNet 基本结构示意图

以 LeNet 为例对 CNN 结构进行介绍。LeNet^[58]主要由卷积层块与全连接层块组成, 如图 2-3 所示。前者则是由多个卷积层与池化层相互叠加构成, 卷积层用于获取输入的空间特征, 池化层用于降低卷积层对位置的敏感性。后者则由三个全连接层构成, 输出量分别为 120、84 和 10, 其中 10 为类别数。

2.3 注意力机制

```
pork belly = delicious . || scallops? || I don't even
like scallops, and these were a-m-a-z-i-n-g . || fun
and tasty cocktails. || next time I in Phoenix, I will
go back here. || Highly recommend.
```

图 2-4 Yelp 评论的情绪分类中注意力机制建模示例图

注意力机制 (Attention Mechanism) 最初被应用于机器翻译^[59], 现已成为深度学习领域的一个重要概念, 并广泛应用于自然语言处理、计算机视觉等领域。注意力机制可以用视觉机制直观地解释事物^[60]。例如, 人类视觉系统往往关注图像

中的某些信息，而忽略了不相关的信息。同样，在语言或视觉问题中，输入的某些部分可能比其他部分起着决定性的作用。注意力机制结合了这种相关性概念，允许模型动态地关注输入中对当前任务有贡献的某些部分。如图 2-4 中显示了一个使用注意力机制对 Yelp 评论进行分类的示例。在此样例中^[61]，注意力机制了解到在以下句子中，第一句与第三句子更相关。此外，这些句子中"delicious"和"amazing"两个词在对评论情绪方面起到决定性作用。

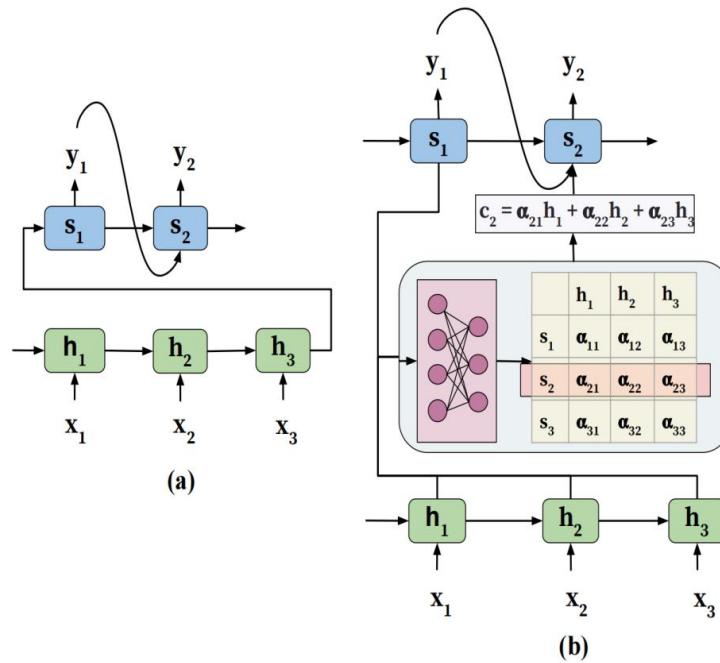


图 2-5 (a) 编码器-解码器架构；(b) 具有注意力机制编码器-解码器架构

以图 2-5(a) 模型为例，编码器为 RNN 结构，序列 $\{x_1, x_2, \dots, x_T\}$ 为其输入， T 为输入序列的长度，并将该序列编码成固定长度的向量集合 $\{h_1, h_2, \dots, h_T\}$ 。解码器同为 RNN 结构，将固定长度向量 h_T 作为输入，并且迭代地生成输出序列 $\{y_1, y_2, \dots, y_{T'}\}$ ， T' 表示输出序列的长度。在每个时刻 t ， h_t 和 s_t 分别表示编码器和解码器的隐藏层状态。

注意力机制旨在通过允许解码器访问整个编码的输入序列 $\{h_1, h_2, \dots, h_T\}$ 来解决上述的问题。其本质是在输入序列上引入注意力权重 α ，以便在生成下一个输出标记时优先考虑集合中的相关信息。

图 2-5(b) 为具有注意力机制的编码器-解码器架构。架构中的注意力模块对权重 α_{ij} 进行学习，其表示 h_i 和 s_j 之间的相关性，其中 h_i 为编码器隐藏状态， s_j 为解码器隐藏层状态。这些权重被用来建立上下文向量 c ，该向量作为输入传递给解码器。在每个解码位置 j ，上下文向量 c_j 是所有编码器的隐藏层状态与对应的

注意力权重的加权和，即 $c_j = \sum_{i=1}^T \alpha_{ij} h_i$ 。该机制的优势使得解码器不仅考虑了整个输入序列的信息，而且重点关注输入序列中相关的信息。

2.2.1 注意力机制分类

按注意力的可微性可分为：软注意力机制与硬注意力机制。按注意力所关注的域可分为：空间域、通道域、层域、混合域以及时间域。

软注意力机制的更关注区域信息或通道信息，并且具有确定性，训练完成后可直接通过网络生成。同时软注意力机制也具有可微性，它可以通过前向传播与反向传播来学习参数权重。然而，在实际应用中过分耗费运算资源。强注意力机制主要关注图像中各个像素的信息，即图像中的各个像素都可能延伸注意力。同时，强注意力机制是一个更强调动态变化的随机预测过程且具有不可微性。

2.2.2 相关注意力机制

(1) 空间域注意力

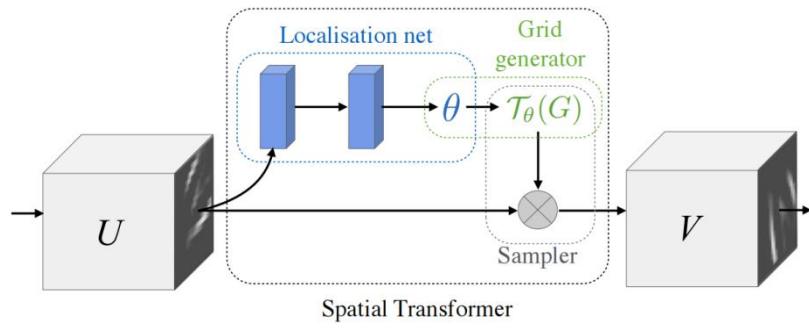


图 2-6 STN 结构示意图

Jaderberg^[62]提出空间转换网络 (Spatial Transformer Networks, STN)，STN 通过注意力机制将原始图片中的空间信息转换到另一空间并保留关键信息。由于传统的池化方式无法识别关键信息。为此，Jaderber 提出一种空间转换模块，通过对图像中的空间域信息进行变换，来提取关键信息。如图 2-6 为 STN 结构示意图。

(2) 混合域注意力

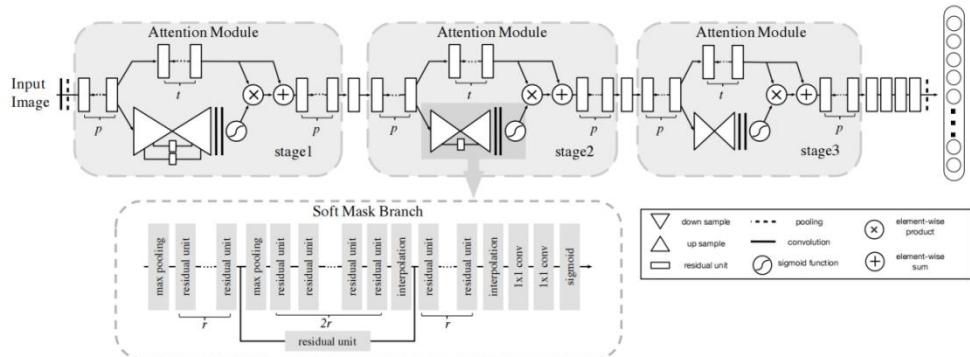


图 2-7 Residual Attention Network 结构示意图

Fe^[63]在提出的残差注意力网络(Residual Attention Network)中引入一种软注意力机制,如图2-7所示。这种注意力机制掩码(Mask)不仅适用于空间域或者通道域,而且可以作为每个特征元素的权重。注意力掩码的独特之处在于引入了残差注意力机制进行学习,不仅将掩码后的特征向量作为下一层输入,而且将掩码前的特征向量也作为下一层的输入。此时,可以获得更丰富的特征,从而能够更好地关注关键特征。

(3) 多关注域注意力

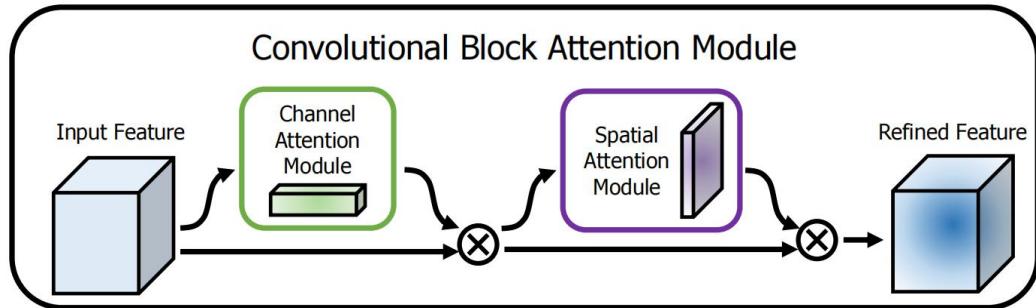


图 2-8 CBAM 结构示意图

Woo^[64]提出了(Convolutional Block Attention Module, CBAM)其实质是对SE-Net的进一步拓展。具体来说,首先通过通道注意力模块(Channel Attention Module)是沿着通道轴寻找“什么”是有意义的输入,之后使用空间注意力模块(Spatial Attention Module)是用来寻找空间维度上“哪里”是有用的信息,是对通道注意力的补充,图2-8为CBAM结构示意图。

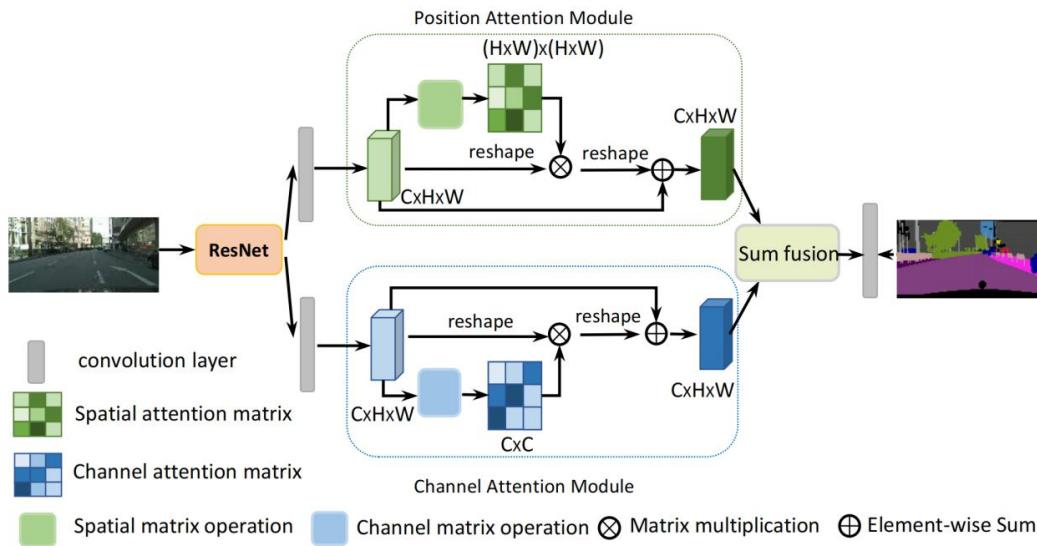


图 2-9 DANet 结构示意图

Jie^[66]提出的(Dual Attention NetWork, DANet)实质是CBAM与Non-local Neural Networks的融合变形。如图2-9所示,通过ResNet提取的特征图分别在空间注意力模块与通道注意力模块中捕获空间与通道维度之间的视觉特征依赖,最后将两个结果进行融合。

2.4 基于深度学习的高光谱解混

2.4.1 基于 AANN 的高光谱解混

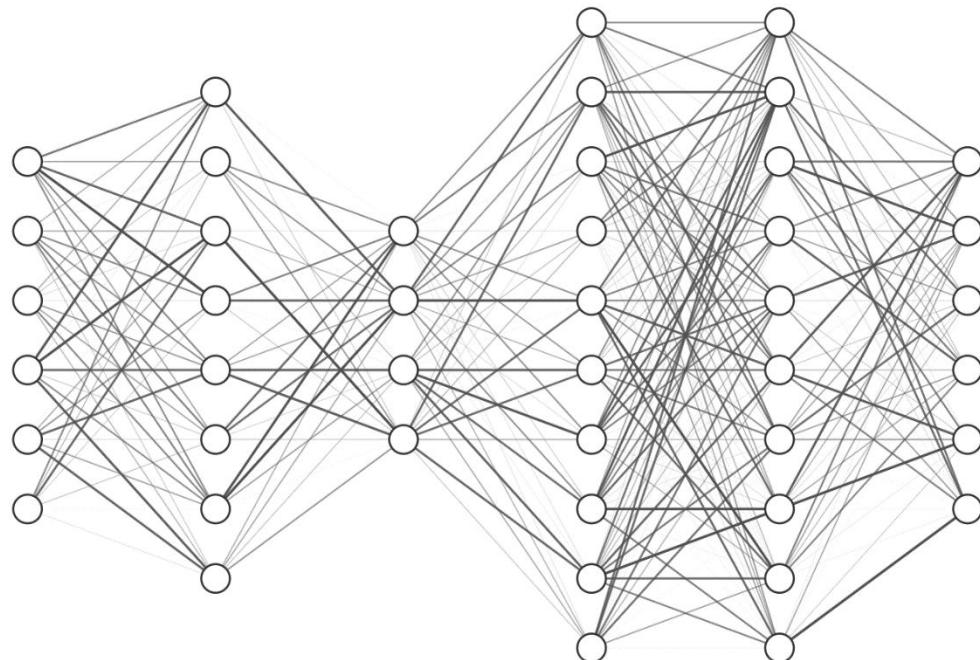


图 2-10 AANN 高光谱解混示意图

Licciardi^[41]提出从高光谱数据中提取丰度的端元自关联神经网络(Auto-Associative Neural Network, AANN)，扩展了之前文献中关于非线性非混合使用 NN 的研究。如图 2-10 所示，神经网络结构由两个阶段组成：在第一阶段提取输入向量的特征，在第二阶段从提取的特征空间进行映射以获得丰度。然而，该方法并没有考虑光谱-空间维度之间的联合信息，且特征提取和混合必须单独训练。

2.4.2 基于 CNN 的高光谱解混

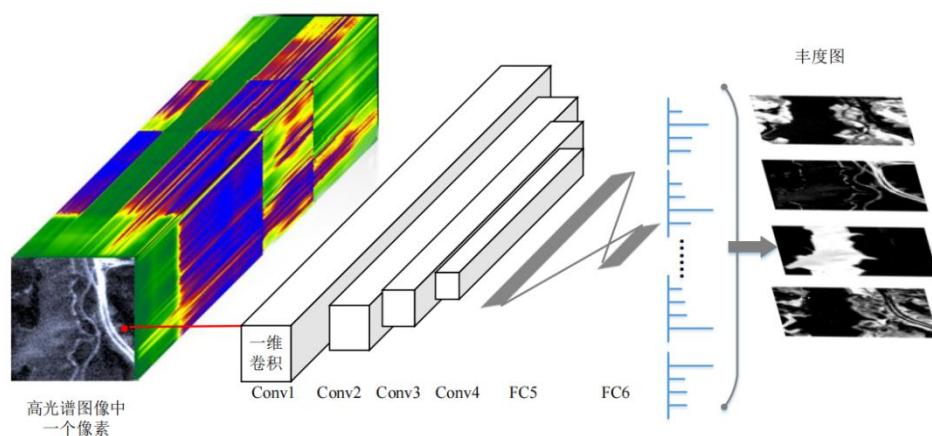


图 2-11 CNN 高光谱解混示意图

Zhang X^[67]提出通过 CNN 将高光谱图像的特征提取与解混集成到一个网络中。如图 2-11 所示，通过对输入图像提取的特征信息进行建模，利于后续的解

混。通过深度卷积神经网络提取的特征作为多层感知器结构的输入，用于网络后续的模糊分类过程。该模糊分类过程可以看作模糊解混的过程，该过程将卷积提取到的特征映射为该像素属于输出层每个神经元的概率，每个神经元所代表的端元的概率也就是丰度值。通过对网络的丰度映射进行归一化来获得像素的丰度值。

基于 CNN 的解混算法比传统的解决方案方法更易于实现，并且可以减少后续人工处理，使模型完整的地从原始输入到最终输出丰度值。然而，此类基于分类模型的解混方法需要具有标记的 HSI 训练样本集，而对于 HSI 而言，标记样本太难以获取，这限制了此类方法在解混任务中的精度。

2.4.3 基于 FNNC 的高光谱解混

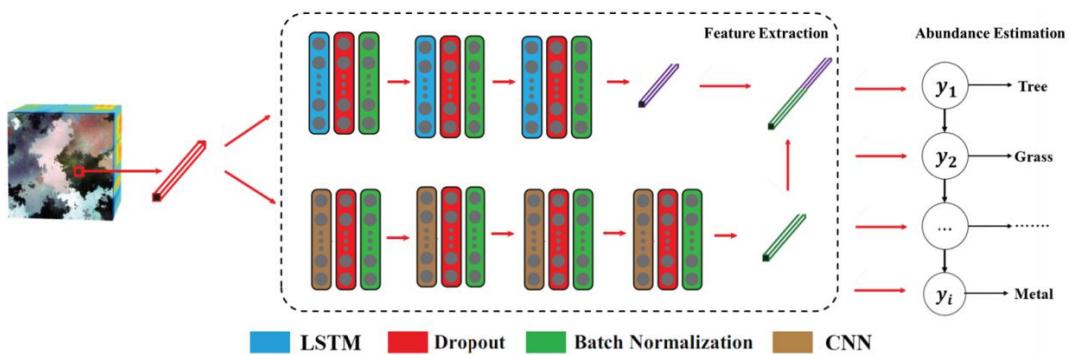


图 2-12 FNNC 高光谱解混示意图

Lei M^[68]提出一种带链式分类器的融合网络 (Fusion Network with Chain Classifier, FNNC) 的新网络结构用于 HU。如图 2-12 所示，FNNC 框架主要包含两个阶段：基于 CNN 和 Bi-LSTM 的特征提取阶段和基于链式分类器的丰度估计阶段。其中，Bi-LSTM 可以更好地捕捉光谱带之间的长距离信息，链式分类器更适合于处理丰度估计过程中存在的多标签任务。

考虑到单像素光谱带的序列性质，FNNC 利用 Bi-LSTM 来处理光谱带的序列数据，并提取其全局和长途特征。并与 CNN 从光谱带中提取局部特征相融合，然后将融合后的特征输入到链式分类器进行丰度估计。FNNC 可以充分利用 HSI 中丰富的光谱信息，获得更明显的特征，从而提高解混性能。

从提取的特征中估计像素丰度的过程类似于多标签分类，因此提出了一种新的基于多标签问题的丰度估计方法，通过引入链式分类器的概念来代替传统的丰度估计方法。链式分类器可以将多标签估计问题分解成几个有序的二分类子问题，显著提高了解混性能。具体来说，训练一个二分类器，将其预测结果传递给具有共享特征的下一个二分类器，得到下一个预测，然后继续这个过程，直到训练完最后一个分类器。链式分类器的先进之处在于，结构中除了第一个子分类器外的每一个子分类器都可以从之前的分类器中获得额外的信息，从而提高其预测

的精度。然后根据预测的指标对分类器进行排序，并重新训练链式分类器以获得更好的性能。链分类器中的每个分类器可以表述为：

$$y_i = LSTM(\{h_{global} || h_{local}, y_{i-1}\}), i=1, 2, \dots, M \quad (2-4)$$

式中， y_i 表示第 i 个分类器的预测值， M 为像素对应的分类器个数。 h_{global} 和 h_{local} 分别是通过 RNN 和 CNN 提取的特征， $||$ 表示级联操作。

2.5 本章小结

本章主要阐述了高光谱混合模型的基本概念以及深度学习中的相关概念，并在章节末尾介绍了三种基于深度学习的解混算法。

3.基于双注意力卷积网络的高光谱解混

3.1 引言

近年来，深度学习在 HU 领域广泛应用，而目前的解混算法都存在对光谱-空间信息运用不充分的问题。为此，本文提出一种基于双注意力卷积网络 (Dual Attention Convolutional Network, DACN) 的端到端解混算法，该方法在 CNN 提取端元特征的基础上增加了两种类型的注意力模块，对光谱维度和空间维度上的特征信息进行建模，以自适应地融合局部特征与全局特征。此外，DACN 使用层标准化和池化来避免过拟合。在 Jsaper Ridge 和 Urban 两个高光谱数据集上进行了完整的性能评估，与现有的 HU 算法相比，DACN 能更有效地提取光谱-空间特征信息，使得解混精度有较为显著的提高¹。

3.2 双注意力卷积网络

近几年，注意力机制已被广泛应用于深度学习领域。基于注意力机制的模型不但能够记录信息之间的位置关系，并且可以根据信息权重的不同度量信息特征的重要性。为了增强关键信息，弱化无用信息，在相关信息特征和不相关信息特征之间进行选择，建立动态权重参数，从而提高深度学习算法的效率，改善传统深度学习的一些缺陷。考虑到 HSI 可获得的真实地物信息有限，解混算法需要特别关注网络中的关键特征信息，为此提出利用 DACN 将 HSI 的特征学习和解混集成到一个网络中。在该方法中，利用 DACN 提取特征，并将这些特征作为网络最后一层的多层感知机的输入，用于基于像素的模糊分类过程，该过程通过归一化结果来获得相应像素的丰度。

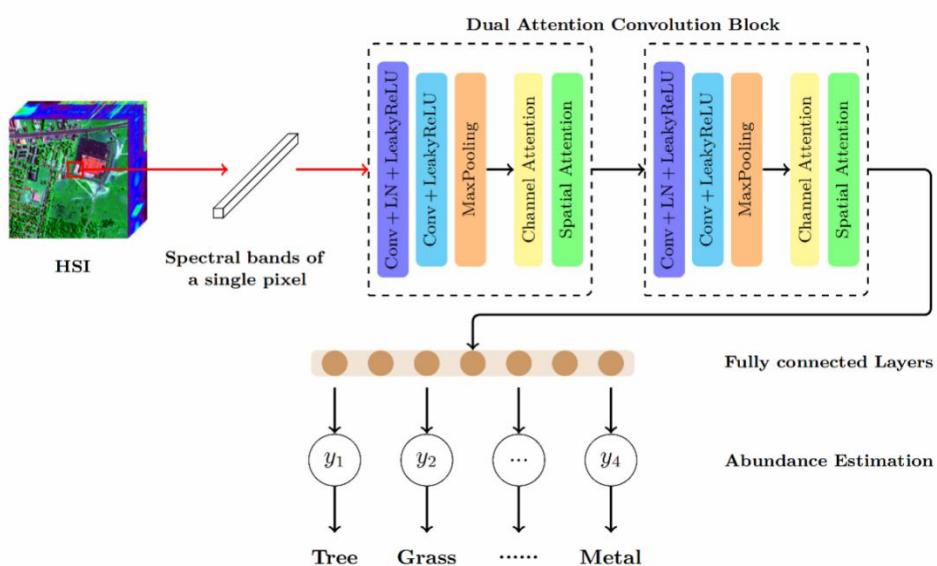


图 3-1 基于双注意力卷积网络 (DACN) 的高光谱解混示意图

¹ DACN 算法已开源：<https://github.com/laugh12321/DACN>

3.2.1 网络结构

如图 3-1 所示 DACN 包含两个双注意力模块。其中，每个模块采用 2 个连续的卷积层，卷积层之间不设置池化层，以实现特征信息的保留与传递。在卷积层之后设置 1 个最大池化层 (MaxPooling)，用以减少参数数量，利于模型收敛。在池化层之后利用卷积块注意力模块 (CBAM) 学习要加强或弱化哪些信息，有效地提取光谱-空间特征信息。考虑到 HSI 中单个像素光谱波段的序列性，同时为了避免过拟合现象并提高模型训练速度与精度，网络采用层标准化，在每一层对单个样本的所有维度特征做标准化。由于激活函数 ReLU 的输入值为负时，其输出与一阶导数始终为零，导致神经元参数不能正常更新。为此，采用 Leaky ReLU 来避免神经元失活。表 3-1 为 DACN 网络结构具体细节。

表 3-1 DACN 网络结构

Layers	K	Network Structure
Input	-	-
	1×8	Conv-LN-LeakyReLU
	1×16	Conv-LN-LeakyReLU
Dual Attention Block 1	1×2	MaxPooling
	1×7	Channel attention
	-	Spatial attention
	1×32	Conv-LN-LeakyReLU
	1×64	Conv-LN-LeakyReLU
Dual Attention Block 2	1×2	MaxPooling
	1×7	Channel attention
	-	Spatial attention
Flatten	-	Flatten
FC Layer 1	192	FC-sigmoid
FC Layer 2	150	FC-sigmoid
FC Layer 3	6	FC-softmax

设 $S \in \Re^{H \times W \times C}$ 为网络输入的 HSI 中一个像元的光谱向量，其中 H 、 W 和 C 分别表示三个维度。输入 S 或特征映射 F^S 与卷积核卷积，得到特征映射如下：

$$F_l^S = F_{l-1}^S * W_l + b_l \quad (3-1)$$

式中， F_{l-1}^S 和 F_l^S 分别表示第 l 层卷积的输入和输出。当 $l=0$ 时， $F_0^S=S$ 。此外， $*$ 代表卷积运算符。 W_l 和 b_l 是第 l 层卷积的权重和偏差。

考虑到高光谱图像可获得的地面对真实信息有限，算法需要特别关注网络中的关键特征信息，以提高训练精度。因此，引入了注意力机制来解决这个问题。利用卷积块注意力模块 (CBAM) 对光谱和空间维度中的语义信息进行建模，以自适应地融合局部和全局特征。图 3-2 描述了每个注意力模块的计算过程，下面介绍每个注意力模块的详细信息。

F_{l-1}^S 将被重塑为一个适当的维度 $F^{S1} \in \Re^{H \times W \times C}$ ，有利于通道注意力网络的

输入。一维通道注意力图被定义为 $M_{Ch} \in \Re^{1 \times 1 \times c}$ ，而二维空间注意力图被定义为 $M_{Sp} \in \Re^{H \times W \times 1}$ 。考虑到特征图 F_{l-1}^S ，CBAM 模块可以计算出两个独立维度(通道和空间)的注意力权重，然后将其乘以 F_{l-1}^S 来实现特征图的细致化。通道注意力是沿着通道轴寻找“什么”是有意义的输入，而空间注意力是用来寻找空间维度上“哪里”是有用的信息，是对通道注意力的补充。

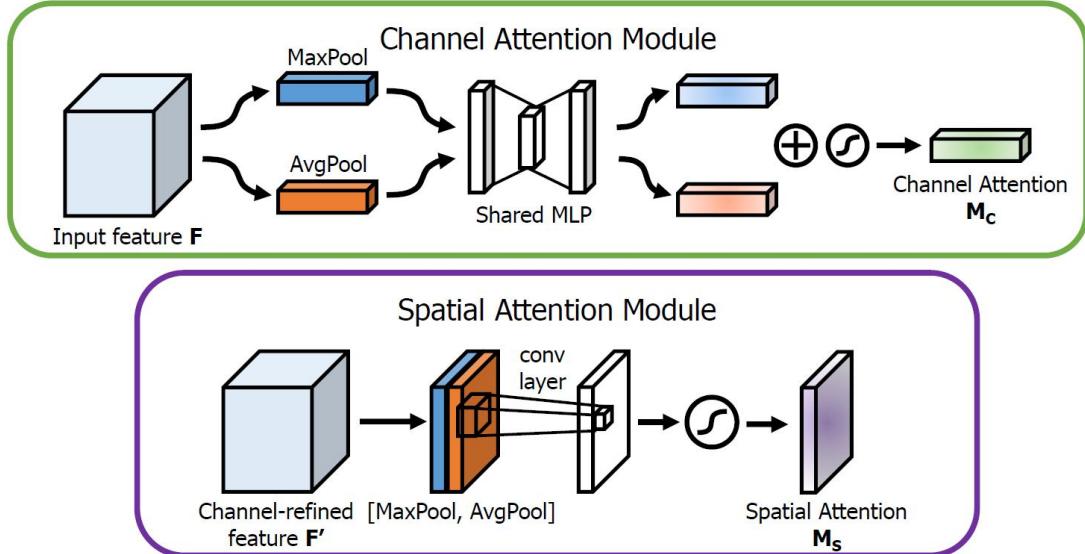


图 3-2 各注意力模块示意图

由[64]可知，最大池化层与平均池化层都与由具有一个隐藏层的 MLP 组成的共享网络一起应用于通道注意力模块。两个输出特征向量使用元素求和法进行合并，然后利用 Sigmoid 激活函数来获得通道注意力图，如图 3-2 所示：

$$M_{Ch}(F^{S1}) = \sigma(\omega^* F_{avg}^{S1} + \omega^* F_{max}^{S1}) \quad (3-2)$$

式中， ω 为 MLP 的共享权重。

因此，通道注意力模块的输出可描述为：

$$F^{S2} = M_{Ch}(F^{S1}) \otimes F^{S1} \quad (3-3)$$

式中， \otimes 表示元素相乘。

对于空间注意力模块，沿通道轴应用平均池化与最大池化运算，这些输出被串联起来，生成一个特征描述符，然后用卷积层获得空间注意力图 M_{Sp} 。其中，卷积层的滤波器 $f^{7 \times 7}$ ，其大小为 $f^{7 \times 7}$ ，Sigmoid 激活函数 σ 。

$$M_{Sp}(F^{S2}) = \sigma(f^{7 \times 7}([F_{avg}^{S2}, F_{max}^{S2}])) \quad (3-4)$$

式中， F_{avg}^{S2} 和 F_{max}^{S2} 分别表示平均池化运算符和最大池化运算符。

之后，空间注意力模块的输出可以描述为：

$$F^{S3} = M_{Sp}(F^{S2}) \otimes F^{S2} \quad (3-5)$$

最后，整个注意力模块的输出结果是：

$$F_l^S = F^{S3} \quad (3-6)$$

在执行最后一个特征提取层后，利用全连接层对提取的特征进行模糊分类，并获得每个像素所对应的端元丰度。

对于全连接层，假设输出单元与实际端元类别相同时为“1”，其余时为“0”。尽管网络使用输出向量中的二进制值进行训练，但处理层的激活函数为 Sigmoid 函数，导致输出值在[0,1]范围内。因此，可以认为这些值与模糊隶属度值相关，代表相应的端元丰度。第 i 类所对应的丰度 a_i 如下：

$$a_i = \frac{o_i}{\sum_{k=1}^M o_k} \quad (3-7)$$

式中， o_k 表示与第 k 个端元相关的 DACN 输出， M 为端元数。

Algorithm 1 Procedure of the proposed method.

Input: Hyperspectral datasets; Epoch size: \mathbf{Q} .

Output: Abundances: $\hat{\mathbf{A}} \in \mathbb{R}^{R \times N}$

```

1: Initialization:
2:   Initialize weights and bias of other layers randomly;
3:   Initialize  $\mathbf{q} = 1$ ;
4: while ( $\mathbf{q} \leq \mathbf{Q}$ ) do
5:   Update  $\mathbf{F}_1^S$  in (3-6) via back propagation;
6:   Update  $\mathbf{a}_i$  in (3-7) and other network parameters via gradient derivation;
7:   if ( $\mathbf{q} = \mathbf{Q}$ ) then
8:      $\hat{\mathbf{A}} \leftarrow \mathbf{a} = [\mathbf{a}_1, \mathbf{a}_2, \dots, \mathbf{a}_i]$ 
9:   else
10:     $\mathbf{q} = \mathbf{q} + 1$ 
11:   end if
12: end while

```

3.2.2 网络设置

Algorithm 1 介绍了 DACN 算法流程的具体细节。构建模型后，通过误差反向传播训练整个网络，并使用小批量策略进行参数更新。每次迭代之后保存模型，并在训练结束后保存最优模型。损失函数 L 采用预测真实丰度值之间的均方误差 (mean squared error, MSE)，计算公式如下：

$$L = \frac{1}{n} \sum_{i=1}^n (\hat{\mathbf{a}}_i - \mathbf{a}_i)^2 \quad (3-8)$$

式中， $\hat{\mathbf{a}}_i$ 表示输出的各组分的丰度向量； \mathbf{a}_i 表示真实的地物丰度向量； n 代表批处理数量。DACN 模型采用自适应动量估计 (adaptive moment estimation, Adam) 优化器进行参数学习，以自适应调整学习率。

3.3 本章小结

在本章中我们通过分析传统解混算法以及现有的基于深度学习解混算法的不足，提出了一种基于双注意力卷积神经网络(DACN)的端到端解混方法。详细介绍了 DACN 的网络结构与理论实现，并在章节末尾介绍了 DACN 的相关设置。

4.试验设计与结果分析

4.1 性能评价指标

本文通过模型在测试集上性能来评估模型精度，采用定性和定量两种方式评价解混精度。定性评价主要采用目视的方法观察解混效果。定量评价主要选取了用于 HU 中常用的两种评估函数来进行评价。由于混合像元中存在着多种地物类型，所以本文采用总体和分地物类型精度两种方法进行评价。其中，总体精度采用丰度角距离均方根 (Root Mean Square Of Abundance Angle Distance, rmsAAD) 进行评估，AAD 是用来衡量真实丰度 (a_{wi}) 和网络输出丰度 (a_{ri}) 之间的相似性，具体表述如下：

$$AAD_{a_{wi}} = \cos^{-1} \left(\frac{a_{wi}^T a_{ri}}{\|a_{wi}\| \|a_{ri}\|} \right) \quad (4-1)$$

rmsAAD 以 AAD 为基础计算的总体精度，值越小表示精度越高，公式如下：

$$\text{rmsAAD} = \sqrt{\frac{\sum_{i=1}^N (AAD_{a_{wi}})^2}{N}} \quad (4-2)$$

式中， N 表示测试样本的数量。

分地物类型精度采用均方根误差 (Root Mean Square Error, RMSE) 和进行评估，RMSE 计算的是每类地物的真实丰度 (a_{wi}) 和网络输出丰度 (a_{ri}) 之间的均方根误差。其数学表达式为：

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^N (a_{wi} - a_{ri})^2}{N}} \quad (4-3)$$

式中， N 表示测试样本的数量。一般来说，RMSE 的值越小，输出丰度与真实丰度两者越接近，表示丰度估计的精度越高。

4.2 试验数据及环境介绍

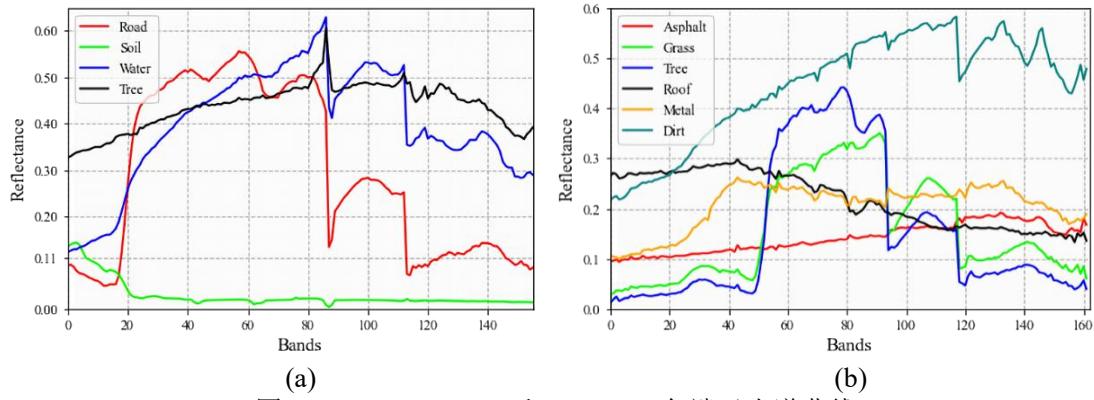


图 4-1 Jasper Ridge (a)和 Urban (b)各端元光谱曲线

4.2.1 数据集介绍

- (1) Jasper Ridge: Jasper Ridge 是经典的高光谱数据，数据由 AVIRIS 传感器捕获。该数据集包含 100×100 个像素和 224 个波段，光谱范围从 380~2500nm。在消除掉被水蒸汽和大气环境影响的波段(1~3, 108~112, 154~166, 220~224)后，得到 198 个波段的数据。图 4-1(a) 所示为 Jasper Ridge 各端元光谱曲线。数据集共有 4 类端元：树木、水、土壤和公路。
- (2) Urban: Urban 数据集是 HU 研究中使用最普遍的高光谱数据集之一。该数据集包含 307×307 个像素和 210 个波段，光谱范围从 400~2500nm。在消除掉被水蒸汽和大气环境影响的波段(1~4, 78, 87, 101~111, 136~153, 198~210)后，得到 162 个波段的数据。图 4-1(b) 所示为 Urban 各端元光谱曲线。数据集共有 6 类端元：沥青路、草地、树木、屋顶、金属和泥土。

4.2.2 实验环境及参数设置

(1) 实验环境

本文所有的实验都是通过 NVIDIA RTX2060 6G GPU、Tensorflow-gpu、Keras 与 Python 3.8 实现的。

(2) 参数设置

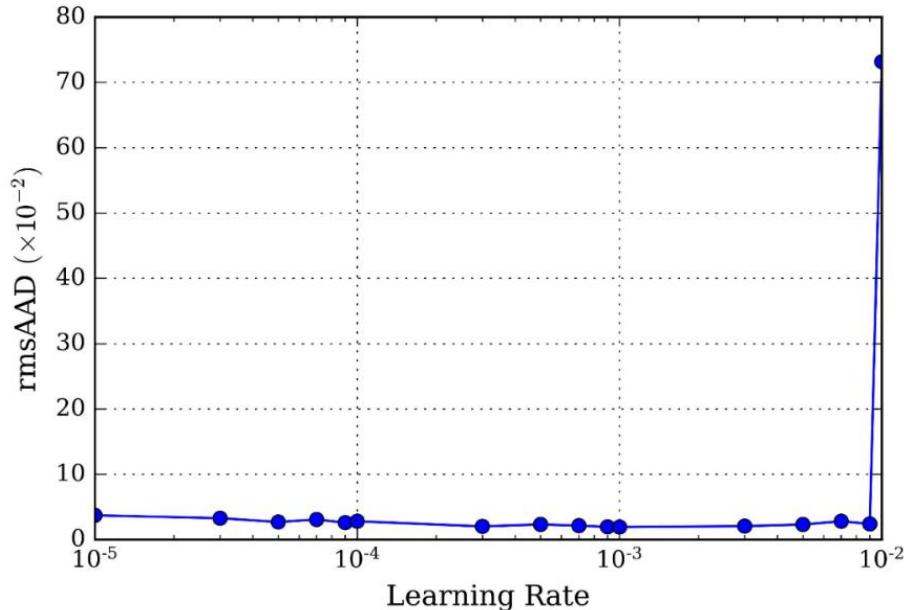


图 4-2 DACN 在不同学习率下的 rmsAAD

对于 AANN，使用经典的单隐层结构，迭代 500 次，学习率设置为 0.01。对于 RNN，四层 GRU 的神经元数被设定为 8、32、128 和 512，学习率设置为 0.001，批量处理值为 128。CNN 和 FNNC 参数引用了[67][68]中的设置，使用了批量处理值为 128 的 Adam 优化器。DACN 使用批量处理值为 128 的 Adam 优化器。在 Jasper Ridge 数据集上对 DACN 进行了学习率参数的敏感性分析，如图

4-2 所示。可以看出当学习率为 0.0009 时, rmsAAD 的值最低。因此, 将学习率 λ 设定为 0.0009。其中, 对 CNN、RNN、FNNC 以及 DACN 都进行 200 次迭代训练, 同时使用早停法(Early Stopping)以获得更好地泛化性能。

4.3 试验分析

为了评估本文设计的 DACN 的性能, 我们选择 AANN、eSVM、CNN、RNN 和 FNNC 方法在不同的真实高光谱数据集中进行比较测试, 以验证 DACN 的有效性。高光谱数据集按照传统的 7:2:1 比例被随机分为训练集、交叉验证集和测试集, 利用训练集及其相应的端元丰度向量来训练模型。

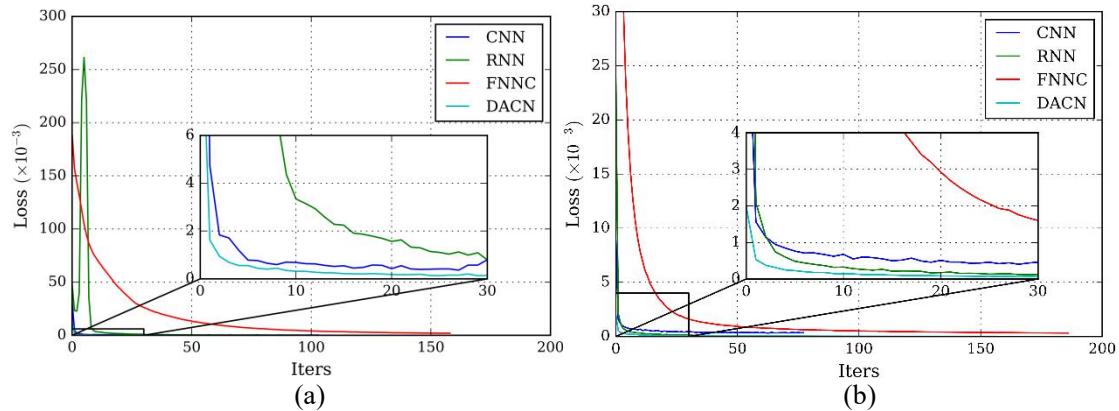


图 4-3 Jasper Ridge (a)、Urban (b) 数据集上不同方法的训练损失曲线

由图 4-3 可以明显看出 FNNC 的泛化性能最弱, DACN 的损失值最低、最先趋于稳定、泛化性能最强。由 4-3 (a) 图 RNN 存在明显的波动现象可以得出, RNN 在 Jasper Ridge 数据集下的泛化能力不稳定。

表 4-1 Jasper Ridge 数据集上不同方法的 RMSE

端元	AANN	eSVM	CNN	RNN	FNNC	DACN
树木	0.0249	0.0357	0.0119	0.0097	0.0188	0.0060
水	0.0267	0.0431	0.0113	0.0076	0.0180	0.0062
土壤	0.0386	0.0543	0.0157	0.0155	0.0283	0.0095
公路	0.0296	0.0379	0.0146	0.0162	0.0270	0.0093
总和	0.1198	0.1711	0.0536	0.0506	0.0956	0.0313

表 4-1 统计了在 Jasper Ridge 数据集上不同方法计算得到的每种地物丰度的 RMSE 值, 其中粗体代表每一行的精度最高值。对于 Jasper Ridge 数据集中每个端元的 RMSE, AANN 优于 eSVM, FNNC 优于 AANN, CNN 优于 FNNC, RNN 优于 CNN, DACN 优于前五种方法。从四个端元的 RMSE 的总和值来看, DACN 比 CNN、RNN 和 FNNC 精度分别提高了 41.6%、38.1% 和 67.2%。FNNC 方法得到的精度低于 CNN 和 RNN 的精度, 与文章不符, 据分析是由于没有有效地结合 RNN 与 CNN 所得到的特征信息而造成的。DACN 方法计算得到的精度最高, 说明该方法能够综合利用光谱-空间特征信息, 对进一步提高解混精度具有重要意义。

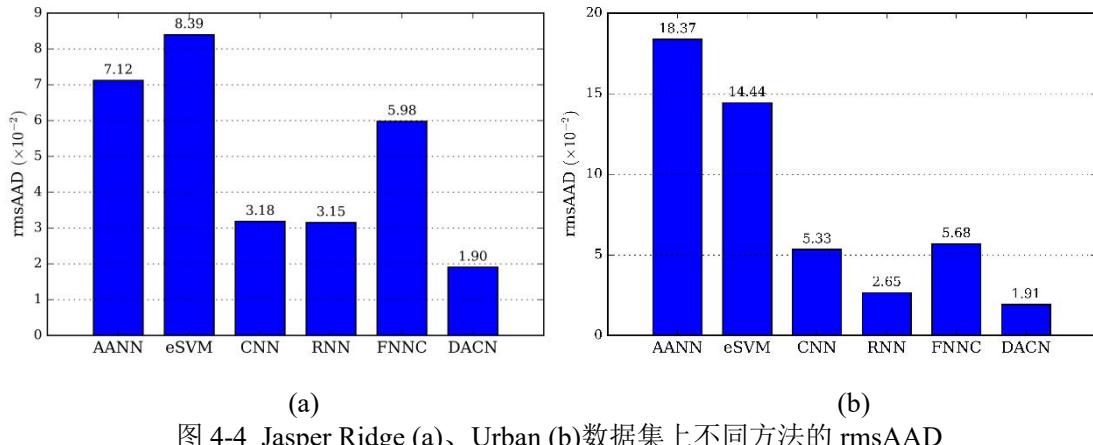


图 4-4 Jasper Ridge (a)、Urban (b) 数据集上不同方法的 rmsAAD

Jasper Ridge 数据试验 rmsAAD 统计值反映了各种方法的总体精度，从图 4-4(a) 中可以看出本文方法计算得到的 rmsAAD 最低，说明 DACN 的总体丰度估计精度最好。从图 4-4(a) 中的统计值为基础对不同方法进行比较，DACN 比其他 5 种方法精度分别提高了 73.3%、77.3%、40.2%、39.6% 和 68.2%。图 4-4(a) 计算的结果与表 4-1 中 RMSE 精度评价结果保持一致。

表 4-2 Urban 数据集上不同方法的 RMSE

端元	AANN	eSVM	CNN	RNN	FNNC	DACN
沥青路	0.0650	0.0467	0.0204	0.0094	0.0187	0.0071
草地	0.0568	0.0531	0.0160	0.0079	0.0169	0.0057
树木	0.0575	0.0381	0.0133	0.0073	0.0150	0.0050
屋顶	0.0397	0.0348	0.0116	0.0058	0.0150	0.0044
金属	0.0492	0.0224	0.0134	0.0066	0.0149	0.0050
泥土	0.0477	0.0528	0.0175	0.0085	0.0210	0.0057
总和	0.4954	0.2481	0.0924	0.0458	0.1018	0.0331

Urban 数据集上不同方法计算得到的每种地物丰度的 RMSE 值如表 4-2 所示，其中粗体代表各端元的精度最高值。可以看出以下解混算法中，CNN 能以较为理想的精度分解出地物丰度，而 AANN 精度稍低，本文提出的 DACN 性能仍然最优。以四个端元的 RMSE 的总和值为准，DACN 比 CNN、RNN 和 FNNC 精度分别提高了 64.1%、27.7% 和 67.4%。

图 4-4(b) 图为不同算法在 Urban 数据集下计算的 rmsAAD 统计值，可以看出本文方法计算得到的值最低，说明 DACN 的总体丰度估计精度最好。DACN 精度比其他 5 种方法精度分别提高了 89.6%、86.7%、64.1%、27.9% 和 66.3%。DACN 方法综合运用光谱-空间特征信息，同时引入层标准化防止过拟合，在整体上保持较为理想的解混效果。

为了更直观地表现出 DACN 算法的解混效果，将 Jasper Ridge 与 Urban 数据集下不同方法的 rmsAAD 值绘制在同一张图中，如图 4-5 所示。结合以上图表信息可以得出，FNNC 理论与实际存在很大出入，同时 RNN 在 Jasper Ridge 与 Urban 数据集上的解混性能仅此于 DACN。由于 DACN 具有解混精度高、特征信息丰

富、泛化能力强、训练周期短等优势，在实际应用中具有一定的应用前景。

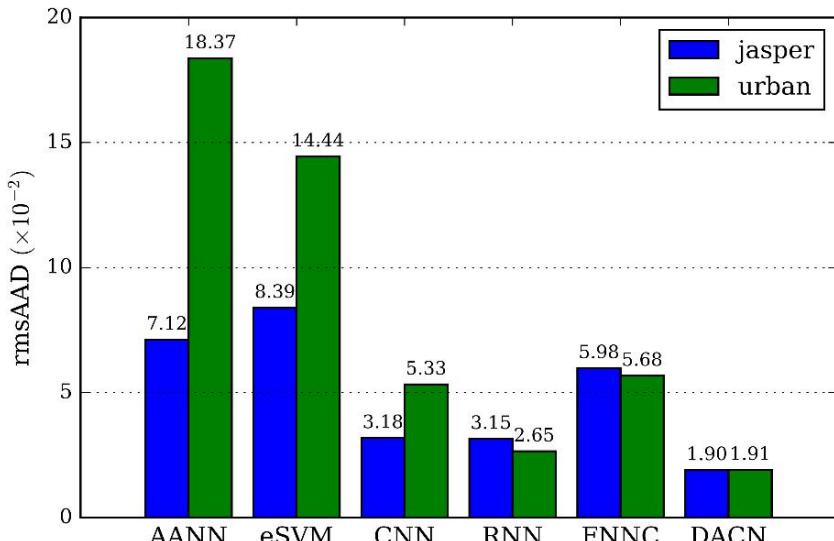


图 4-5 Jasper Ridge、Urban 数据集上不同方法的 rmsAAD

图 4-6 给出了 Urban 数据集的地表真实丰度值以及 CNN、RNN、FNNC 和 DACN 这 4 种方法对应的丰度分布图。可以看出，每个端元的分布都是相对离散的和复杂的。然而，我们提出方法的解混性能得到了显著的提高。从泥土的丰度，分布图中可以看出，DACN 的解混精度相比其他方法更为准确。

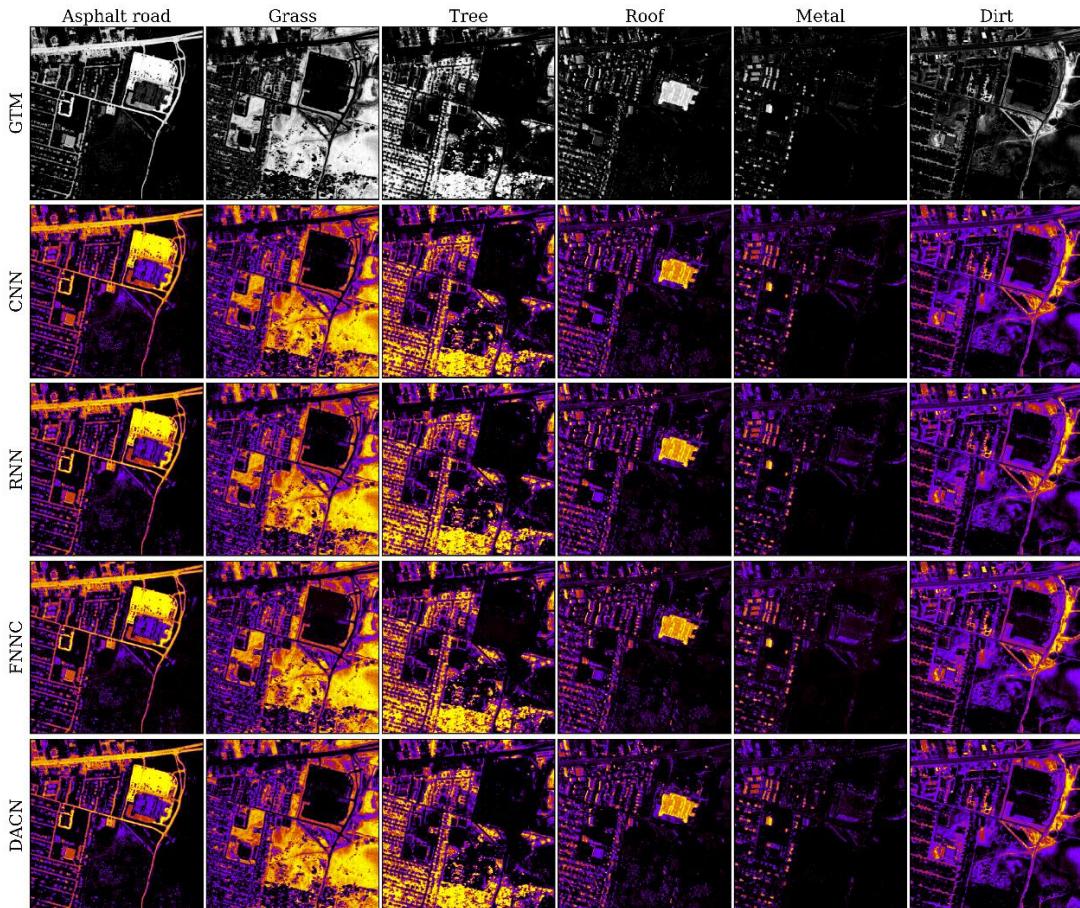


图 4-6 Urban 数据集地面真实与不同方法的估计丰度

4.4 本章小结

本章介绍了本文提出的 DACN 算法在试验中使用的评价指标、所使用的数据集、试验环境以及训练参数。通过对比试验，证明了本文提出的 DACN 方法能提取更丰富全面的光谱-空间特征信息，相较于 AANN、eSVM、CNN、RNN、FNNC 等几种有代表性的方法，DACN 能进一步改善解混精度。

5.总结与展望

5.1 总结

受限于高光谱成像仪有限的空间分辨率，再加上地表微观物质混合和多重散射现象，HSI 中通常会存在大量的混合像元，即单个像元中包含两种或两种以上的物质。混合像元的出现严重制约了 HSI 的分类、检测、融合等任务及其后续在环境监测、城市观测、矿物勘探等领域的应用。因此，用于推断 HSI 混合像元中的端元及其相应丰度分数的 HU 技术成为利用高光谱数据的一项重要基础技术。然而，现有 HU 算法对 HSI 中的空间信息、光谱信息存在利用不够充分的问题。针对上述问题，本文提出了基于双注意力卷积网络的 HU 算法，提取 HSI 的光谱-空间特征信息，从而提升 HSI 的解混精度。本文主要工作如下：

- (1) 提出了一种基于双注意力卷积网络的 HU 算法。该方法通过在卷积神经网络提取特征的基础上增加了两种类型的注意力模块，对光谱维度和空间维度上的语义信息进行建模，以自适应地融合局部特征与全局特征。此外，DACN 使用层标准化和池化来避免过拟合。
- (2) 在 Jasper Ridge 和 Urban 两个高光谱数据集上进行了完整的性能评估，与现有的解混算法相比，DACN 能更有效地提取光谱-空间特征信息，解混精度有较为显著的提高。

5.2 展望

本文的研究对于注意力机制在 HU 中的应用取得了一定的进展，在深度解混网络中引入注意力机制方面，其精度还有待提高，更高效的解混网络架构也亟待引入，本文方法仍有待改进的空间：

- (1) **数据降噪：**由于光谱成像设备的限制，使得 HSI 不可避免地受到各类噪声的污染，导致 HSI 视觉效果变差且限制了后续处理的精度。因此在 HSI 前对其进行降噪处理是很有必要的。
- (2) **深度残差网络：**基于深度神经网络模型的解混性能会随着模型层数的增加而下降。因此可在每两个卷积层之间添加映射来建立深度残差网络，防止解混精度下降。
- (3) **光谱-空间注意力网络：**未来考虑设计一种光谱-空间注意力网络来进一步提高解混精度。其中，带注意力的循环神经网络在连续频谱内学习内部光谱相关性，带注意力的卷积神经网络关注显著特征与相邻像素间的空间相关性。
- (4) **并行运算：**在实际应用中，当 HSI 过大时，训练时间可能较长。未来考虑采用 GPU 并行运算的方式缩短训练时间。

参考文献

- [1] 童庆禧, 张兵, 郑兰芬. 高光谱遥感——原理、技术与应用[J]. 中国科学院遥感应用研究所, 2006.
- [2] 张钧萍, 张晔, 周廷显. 成像光谱技术超谱图像分类研究现状与分析[J]. 中国空间科学技术, 2001, V21(001):37-44.
- [3] Li H , Liu J , Yu H . An Automatic Sparse Pruning Endmember Extraction Algorithm with a Combined Minimum Volume and Deviation Constraint[J]. Remote Sensing, 2018, 10(4):509.
- [4] Jiao C , Chen C , Mcgarvey R G , et al. Multiple instance hybrid estimator for hyperspectral target characterization and sub-pixel target detection[J]. ISPRS Journal of Photogrammetry and Remote Sensing, 2018, 146(DEC.):235-250.
- [5] 卢思佳. 高光谱遥感影像混合像元分解算法研究[D]. 电子科技大学.
- [6] Wang X , Zhong Y , Yao X , et al. Saliency-Based Endmember Detection for Hyperspectral Imagery[J]. IEEE Transactions on Geoscience & Remote Sensing, 2018, PP(7):1-14.
- [7] Xu X , Shi Z , Pan B . A supervised abundance estimation method for hyperspectral unmixing[J]. Remote Sensing Letters, 2018, 9(4-6):383-392.
- [8] 袁静, 章毓晋, 高方平. 线性高光谱解混模型综述[J]. 红外与毫米波学报, 2018, 37(05):43-61.
- [9] Chen S , Le W . Linear Spatial Spectral Mixture Model[J]. IEEE Transactions on Geoscience and Remote Sensing, 2016, 54(6):1-14.
- [10] 杨斌, 王斌. 高光谱遥感图像非线性解混研究综述[J]. 红外与毫米波学报, 2017(2).
- [11] Li C , Liu Y , Cheng J , et al. Hyperspectral Unmixing with Bandwise Generalized Bilinear Model[J]. Remote Sensing, 2018, 10(10).
- [12] Marinoni A , Plaza A , Gamba P . Harmonic Mixture Modeling for Efficient Nonlinear Hyperspectral Unmixing[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2016.
- [13] Boardman J W , Kruscl F A , Green R O . Mapping target signatures via partial unmixing of AVIRIS data. JPL Publication, 1995.
- [14] Winter M E . N-FINDR: An algorithm for fast autonomous spectral end-member determination in hyperspectral data[C]// Spies International Symposium on Optical Science. International Society for Optics and Photonics, 1999.
- [15] Nascimento J , Dias J . Vertex component analysis: a fast algorithm to unmix hyperspectral data[J]. IEEE Transactions on Geoscience & Remote Sensing, 2005, 43(2):898-910.
- [16] Moussaoui S , Carteret C , Brie D , et al. Bayesian analysis of spectral mixture data using Markov Chain Monte Carlo Methods[J]. CHEMOMETRICS AND INTELLIGENT LABORATORY SYSTEMS, 2006.
- [17] Wang J , Chang C I . Applications of Independent Component Analysis in Endmember Extraction and Abundance Quantification for Hyperspectral Imagery[J]. IEEE Transactions on Geoscience & Remote Sensing, 2006, 44(9):2601-2616.
- [18] Nascimento J , Lisboa E , Conselheiro R , et al. Hyperspectral unmixing algorithm via dependent component analysis[C]// IEEE International Geoscience & Remote Sensing Symposium. IEEE, 2007.
- [19] 孙乐. 空谱联合先验的高光谱图像解混与分类方法[D]. 南京理工大学, 2014.
- [20] Lee D D , Seung H S . Learning the parts of objects by non-negative matrix factorization.[J]. Nature, 1999, 401(6755):788-791.
- [21] 陈善学, 储成泉. 基于稀疏和正交约束非负矩阵分解的高光谱解混[J]. 计算机应用, 2019, 39(008):2276-2280.
- [22] Hoyer P O . Nonnegative matrix factorization with sparseness constraints[J]. Journal of Machine Learning Research, 2004, 5(9):1457-1469.
- [23] Qian Y , Member, IEEE, et al. Hyperspectral Unmixing Via L1/2 Sparsity-constrained Nonnegative Matrix Factorization[J]. IEEE Transactions on Geoscience & Remote Sensing, 2011, 49(11):4282-4297.

- [24] 祝伟, 王雪, 黄岩,等. 重加权稀疏和全变差约束下的深度非负矩阵分解高光谱解混[J]. 遥感学报, 2020(4):401-416.
- [25] 甘玉泉, 刘伟华, 冯向朋, et al. 一种稀疏约束的图正则化非负矩阵光谱解混方法[J]. 光谱学与光谱分析, 2019, 39(04):128-137.
- [26] 彭倩, 张兵, 孙旭,等. 结合空间光谱预处理和约束非负矩阵分解的高光谱图像混合像元分解[J]. 中国图象图形学报, 2017(4).
- [27] 袁博. 空间与谱间相关性分析的 NMF 高光谱解混[J]. 遥感学报, 2018.
- [28] Chandrasekhar, S. Radiative transfer[M]. New York : Dover, 1960.
- [29] Broadwater J , Chellappa R , Banerjee A , et al. Kernel fully constrained least squares abundance estimates[C]// Geoscience and Remote Sensing Symposium, 2007. IGARSS 2007. IEEE International. IEEE, 2007.
- [30] Broadwater J , Banerjee A . A comparison of kernel functions for intimate mixture models[C]// Hyperspectral Image and Signal Processing: Evolution in Remote Sensing, 2009. WHISPERS '09. First Workshop on. IEEE, 2009.
- [31] Broadwater J , Banerjee A . A generalized kernel for areal and intimate mixtures. IEEE, 2010.
- [32] Broadwater J , Ba Nerjee A . Mapping intimate mixtures using an adaptive kernel-based technique[J]. IEEE, 2011.
- [33] Raksunthorn N , D Qian. Nonlinear Spectral Mixture Analysis for Hyperspectral Imagery in an Unknown Environment[J]. IEEE Geoscience & Remote Sensing Letters, 2010, 7(4):836-840.
- [34] Somers B , Cools K , Delalieux S , et al. Nonlinear Hyperspectral Mixture Analysis for tree cover estimates in orchards[J]. Remote Sensing of Environment, 2009, 113(6):1183-1193.
- [35] Halimi A , Altmann Y , Dobigeon N , et al. Nonlinear Unmixing of Hyperspectral Images Using a Generalized Bilinear Model[J]. IEEE Transactions on Geoscience & Remote Sensing, 2011, 49(11):4153-4162.
- [36] Altmann Y , Dobigeon N , Tourneret J Y . Bilinear models for nonlinear unmixing of hyperspectral images[C]// Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS), 2011 3rd Workshop on. IEEE, 2011.
- [37] Liu W , Wu E Y . Comparison of non-linear mixture models: sub-pixel classification[J]. Remote Sensing of Environment, 2005, 94(2):145-154.
- [38] Guilfoyle K J , Althouse M L , Chang C I . A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks[J]. IEEE Transactions on Geoscience & Remote Sensing, 2001, 39(10):2314-2318.
- [39] Plaza J , Plaza A , Perez R , et al. On the use of small training sets for neural network-based characterization of mixed pixels in remotely sensed hyperspectral images[J]. Pattern Recognition, 2009, 42(11):3032-3045.
- [40] Plaza J , Plaza A . Spectral Mixture Analysis of Hyperspectral Scenes Using Intelligently Selected Training Samples[J]. IEEE Geoscience & Remote Sensing Letters, 2010, 7(2):371-375.
- [41] Licciardi, Giorgio, A, et al. Pixel Unmixing in Hyperspectral Data by Means of Neural Networks.[J]. IEEE Transactions on Geoscience & Remote Sensing, 2011.
- [42] Licciardi G , Marpu P R , Chanussot J , et al. Linear Versus Nonlinear PCA for the Classification of Hyperspectral Data Based on the Extended Morphological Profiles[J]. IEEE Geoscience & Remote Sensing Letters, 2012, 9(3):447-451.
- [43] Altmann Y . Supervised Nonlinear Spectral Unmixing Using a Postnonlinear Mixing Model for Hyperspectral Imagery[J]. IEEE Transactions on Image Processing, 2012, 21(6):3017-3025.
- [44] Heylen, R, Burazerovic, et al. Non-Linear Spectral Unmixing by Geodesic Simplex Volume Maximization[J]. Selected Topics in Signal Processing, IEEE Journal of, 2011.
- [45] Heylen R , Scheunders P . Calculation of Geodesic Distances in Nonlinear Mixing Models: Application to the Generalized Bilinear Model[J]. IEEE Geoscience & Remote Sensing Letters, 2012, 9(4):644-648.

- [46] Close R , Gader P , Wilson J , et al. Using physics-based macroscopic and microscopic mixture models for hyperspectral pixel unmixing[J]. Proceedings of SPIE - The International Society for Optical Engineering, 2012, 8390:83901L-1.
- [47] Close R , Gader P , Wilson J . Hyperspectral unmixing using macroscopic and microscopic mixture models[J]. Journal of Applied Remote Sensing, 2014, 8(1):083642.
- [48] Guilfoyle K J , Althouse M L , Chang C I . A quantitative and comparative analysis of linear and nonlinear spectral mixture models using radial basis function neural networks[J]. IEEE Transactions on Geoscience & Remote Sensing, 2001, 39(10):2314-2318.
- [49] Mustard J F , Pieters C M . Quantitative abundance estimates from bidirectional reflectance measurements[J]. Journal of Geophysical Research Solid Earth, 1987, 92(B4):E617-E626.
- [50] Keshava N , Mustard J F . Spectral unmixing[J]. IEEE Signal Processing Magazine, 2002, 19(1):44-57.
- [51] Singer R B , McCord T B . Mars - Large scale mixing of bright and dark surface materials and implications for analysis of spectral reflectance. Lunar and Planetary Science Conference Proceedings, 1979.
- [52] Borel C C , Gerstl S . Nonlinear spectral mixing models for vegetative and soil surfaces[J]. Remote Sensing of Environment, 1994, 47(3):403-416.
- [53] Hapke B . Theory of Reflectance and Emittance Spectroscopy: Contents. 1993.
- [54] Judith E , Deleo J M . Artificial neural networks[J]. Cancer, 2001, 91(S8):1615-1635.
- [55] Hubel D H , Wiesel T N . Receptive fields, binocular interaction and functional architecture in the cat's visual cortex.[J]. Journal of Physiology, 1962, 160(1):106-154.
- [56] Fukushima K . Neocognitron: A self-organizing neural network model for a mechanism of pattern recognition unaffected by shift in position[J]. Biological Cybernetics, 1980, 36(4):193-202.
- [57] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, L. D. Jackel; Backpropagation Applied to Handwritten Zip Code Recognition. Neural Comput 1989; 1 (4): 541–551.
- [58] Lecun Y , Bottou L . Gradient-based learning applied to document recognition[J]. Proceedings of the IEEE, 1998, 86(11):2278-2324.
- [59] Bahdanau D , Cho K , Bengio Y . Neural Machine Translation by Jointly Learning to Align and Translate[J]. Computer Science, 2014.
- [60] Xu K , Ba J , Kiros R , et al. Show, Attend and Tell: Neural Image Caption Generation with Visual Attention[J]. Computer Science, 2015:2048-2057.
- [61] Yang Z , Yang D , Dyer C , et al. Hierarchical Attention Networks for Document Classification[C]// Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies. 2016.
- [62] Jaderberg M, Simonyan K, Zisserman A. Spatial transformer networks[C]//Advances in Neural Information Processing Systems. 2015: 2017-2025.
- [63] Fe I W , Jiang M , Chen Q , et al. Residual Attention Network for Image Classification[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). IEEE, 2017.
- [64] Woo S , Park J , Lee J Y , et al. CBAM: Convolutional Block Attention Module[C]// European Conference on Computer Vision. Springer, Cham, 2018.
- [65] Jie H , Li S , Gang S , et al. Squeeze-and-Excitation Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2017, PP(99).
- [66] Fu J , Liu J , Tian H , et al. Dual Attention Network for Scene Segmentation[J]. 2018.
- [67] Zhang X , Sun Y , Zhang J , et al. Hyperspectral Unmixing via Deep Convolutional Neural Networks[J]. IEEE Geoence & Remote Sensing Letters, 2018, PP:1-5.
- [68] Lei M , Li J , Qi L , et al. Hyperspectral Unmixing via Recurrent Neural Network With Chain Classifier[C]// IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium. IEEE, 2020.
- [69] FOODY, G. M . Approaches for the production and evaluation of fuzzy land cover classifications from remotely-sensed data[J]. International Journal of Remote Sensing, 1996, 17(7):1317-1340.
- [70] Wang L , Jia X . Integration of Soft and Hard Classifications Using Extended Support Vector Machines[J]. IEEE Geoscience & Remote Sensing Letters, 2009, 6(3):543-547.