

课程论文题目

吴之越

摘要

本文主要研究如何利用先验信息提高高光谱图像的空间分辨率。近年来，基于深度学习的单高光谱图像超分辨率方法取得了较好的效果。然而，它们大多只是简单地将一般的图像超分辨率深度网络应用于高光谱数据，从而忽略了高光谱数据本身的一些具体特征。为了充分利用高光谱图像的光谱信息，我们通过带有自编码器的空洞投影校正网络将高光谱图像的空间超分问题从图像域转换到丰度域，我们将该网络称为 aeDPCN。特别是，我们首先将低分辨率的高光谱图像编码成丰度表达，并在解码网络中保留光谱信息，这可以大大降低计算复杂度。然后，为了提高丰度嵌入的空间分辨率，我们用空洞投影校正网络对丰度表达进行了由粗到精的超分辨，并引入反向投影策略进一步消除了光谱失真。最后，预测图像是由相同的解码器得到的，这增加了我们方法的稳定性，即使在很大的放大倍数的情况下也是如此。在真实高光谱图像场景上的大量实验表明，我们的方法在精度和效率方面都优于最先进的方法。

关键词：高光谱图像；超分辨；深度学习；自编码器

1 引言

高光谱成像在从紫外线到红外波长的大量连续和狭窄的光谱波段上收集同一场景的光谱信息。因此，高光谱图像以高光谱分辨率获取不同目标的空间关系和反射特性。作为一个三维数据立方体，高光谱图像的每个像素代表了相关物质的一条光谱曲线，可以用来区分图像场景中的物体，特别是地球物体。由于这一特性，高光谱图像已被证明在许多情况下都很有用，例如陆地表面分类^{[1][2]}，异常检测^{[3][4]}，环境监测^[5]等等。然而，由于在空间分辨率，光谱分辨率和信噪比之间不可避免的权衡取舍。高光谱图像通常有低空间分辨率，因此限制了应用范围^[6]。因此，探索基于软件技术在不丢失光谱信息的情况下提高高光谱的方法变得非常必要^{[7][8]}。这种技术通常被称为超分辨技术，探索超分辨技术在计算机视觉领域是一个热门的话题。超分辨技术是一种通过一个或多个对应的低分辨率图像重建高分辨率图像的经典方法。这是一种在不修改硬件的情况下提高空间分辨率的后处理方法。现有的用于高光谱图像的超分辨方法大多是基于融合的方法^{[9][10][11][12][13][14]}，该方法通过将低空间分辨率高光谱图像与同一场景的相应辅助图像进行融合来恢复高空间分辨率高光谱图像。这些附加观测通常具有高空间分辨率和低光谱分辨率，包括多光谱图像、全色图像和 RGB 图像。然而，由于高光谱成像场景的特殊性，这些与高光谱图像在同一场景拍摄的辅助图像往往稀缺甚至无法获得。因此，辅助图形的确实阻碍了基于融合获得高空间分辨率高光谱图像的实际应用。基于融合的方法的另一个障碍是辅助观测已经完全注册到高光谱图像的前提，这也是一个具有挑战性的任务。因此，探索基于单输入的高光谱图像超分辨方法是很有价值的。无需任何附加信息，单幅图像超分辨方法由单幅低空间分辨率高光谱图像生成高空间分辨率高光谱图像。在过去的几十年里，许多基于单幅图像的方法被提出来解决高光谱图像超分问题^{[15][16]}。一些传统方法通常使用自相似性、稀疏性或高光谱图像低秩等先验(手工复杂设计)来正则化超分辨率重建过程^{[17][18]}。近年来，由于卷积神经网络具有较强的图像特征提取能

力,已成功应用于端到端重建高空间分辨率高光谱图像。虽然这些方法都被证明是有效的,但仍存在一些不足。具体来说,前者的浅层启发式模型表达能力有限。因此,他们通常不能适当地恢复复杂的图像细节。而基于卷积神经网络的方法以数百个光谱波段的整个数据作为输入,要么将高光谱图像视为普通的三维数据立方体,忽略了高光谱图像的光谱特征,要么计算复杂度较高。为了解决上述挑战,我们将高光谱图像的基本频谱作为先验信息,以解决高光谱图像在丰度域中的超分辨率问题,这是由自编码器的空洞投影校正网络实现的,称为 aeDPCN。在该方法中,我们利用基于矩阵分解的自编码器,将高光谱图像的超分辨问题转化从图像域转变为丰度域。根据线性谱混合模型的假设,非负矩阵分解作为一种有用的解混方法,可以识别端元谱并估计相应的高光谱图像丰度。特别是,在描述相同的成像场景时,低空间分辨率高光谱图像和高空间分辨率高光谱图像应该具有相同的端元。基于这一观察结果,我们通过自编码器网络对非负高光谱数据进行了转换,该网络起到了分解和重构的作用。具体地说,我们在编码器中嵌入非负约束和一到一约束来生成丰度表示,而解码器则保留基本的光谱信息。为了提高空间分辨率,采用空洞投影校正网络对比原始高光谱图像小得多的丰度嵌入进行超分辨。在该网络中,我们逐步重构高空间分辨率的嵌入,并使用混合空洞卷积提取具有大接收域的信息特征。共享相同的解码器,生成的高空间分辨率表示需要遵循与低空间分辨率表示类似的模式。因此,我们设计了反向投影模块来约束生成的丰度来校正光谱失真。与其他基于卷积神经网络的方法相比,我们在丰度域进行超分辨,大大降低了计算复杂度。此外,解码器保留了谱先验,使我们的方法在重构时更精确,即使在大尺度放大倍数下也不会有很大的谱损失。如图 1 所示,光谱分解也是基于融合的方法的一般程序^{[19][20]}。在此过程中,通过基于模型的方法或基于学习的方法,分别从低空间分辨率高光谱图像和高空间分辨率辅助观测中提取端元谱 S 和高空间分辨率丰度 A 。相反,我们提出的 aeDPCN 重新定义了丰度域的超分问题。特别地,将丰度表示作为嵌入来提高空间分辨率。由于缺乏对单个高光谱图像的超分的辅助观测,aeDPCN 试图学习低空间分辨率丰度 a 与相应高空间分辨率丰度 A 之间的映射函数,这在高光谱图像的超分问题中研究很少。总之,我们的主要贡献有三方面。首先,我们解决了丰度域的超分问题,这是在单高光谱图像的超分任务中的首次尝试。具体来说,我们将单个高光谱图像的超分问题从图像域转化为丰度域,并将矩阵变换和超分端到端的统一起来。这样可以大大降低高光谱图像超分问题的关键计算开销,更好地保存光谱信息。其次,我们在有限层中部署混合空洞卷积来获取深度空间特征,以超分辨丰度表示。因此,我们的 aeDPCN 是一个具有相当性能的轻量级模型,这在广泛的实验中得到了证明。最后,在提高空间分辨率时,为了确保高空间分辨率图像和低空间分辨率图像丰度表示具有相似的模式,集成了反向投影校正策略以降低光谱失真。

2 相关工作

2.1 基于深度神经网络的超分辨技术

随着图像数据库的扩大和计算技术的进步,基于学习的方法在图像分类、目标检测、场景分割等高级计算机视觉任务中取得了很大的成就。同样,对于超分辨率等低层次问题,基于学习的方法主要通过深度神经网络在自然图像中也取得了非常优秀的结果。由于深度神经网络具有强大的学习能力,它被设计为端到端的学习低分辨率图像和高分辨率图像图像对之间的映射函数。^[21]首次应用三卷积层成功地解决了超分辨问题,与传统的超分辨方法相比,显示出极大的优越性。随后,通过跳过连接,在该

任务中引入残差学习^[22]，以简化深度网络的训练过程，例如超分辨率超深网络 (VDSR)^[23]，深度递归卷积网络 (DRCN)^[24]深度递归残差网络 (DRRN)^[25]。此外，增强型深度超分辨率网络 (EDSR)^[26]去掉了剩余模块中不必要的部件，赢得了 NTIRE2017 超分辨率挑战赛。为了最大限度地重用特性，在^[27]和^[28]的工作中，卷积层的输出被紧密地连接起来，以更好地传播信息。在此之后，研究人员研究了更深层次的网络或更复杂的结构来重建高分辨率图像。例如，生成式对抗网络 (GAN)^[29]被使用在^[30]和^[31]中，对于真实感超分辨的工作，其可以在一定程度上缓解模糊和过度平滑的伪影。为了生成多尺度预测，拉普拉斯金字塔 SR 网络 (LapSRN) 采用了一种渐进重建方法。为了生成多尺度预测，拉普拉斯金字塔 SR 网络 (LapSRN)^[32]采用了一种渐进重建方法，以一种由粗到细的方式进行超分。相似的，^[33]利用带有误差反馈机制的上下迭代下采样层来构建深度反投影网络 (DBPN)，为大尺度因子建立了最新的结果。首分类任务的启发，残差通道注意力网络 (RCAN)^[28]以残差的方式利用注意机制，堆叠多达 400 层，达到了最佳的性能。虽然自然图像超分辨问题已经得到了广泛的研究，上述方法在 RGB 图像上也取得了很好的效果，但大多数方法都不能直接应用于高光谱图像超分辨。一方面，这些方法通常被训练为处理三通道或单通道图像。由于高光谱图像数据有数百个通道，这些在 RGB 图像上预训练的基于卷积神经网络的方法应该以一个带一个带的方式扩展到超分辨率的单带图像。但是，由于高光谱图像数据的每个像素都反映了特定材料的光谱信息，这种方法可能会导致光谱失真。另一方面，如果将这些模型的输入和输出维调整为相应的高光谱图像的数据维，则需要从头开始训练网络。为了避免过度拟合，大多数深度网络都需要大量数据来驱动。没有足够的高光谱数据来满足这一要求。因此，有必要专门为高光谱图像设计轻量级方法。

2.2 基于卷积神经网络的超分辨技术

近年来，研究人员提出了几种基于神经网络的高光谱图像的光谱反演方法。为了缓解光谱畸变，提出了一种结合空间误差修正 (SEC) 模型的谱差卷积神经网络 (SECNN)^{[34][35]}，实现了空间信息增强和光谱信息保存。同样的，^[36]将空间约束 (SCT) 策略与 SDCNN 模型相结合，使通过低空间分辨率高光谱图像重构高空间分辨率高光谱图像空间维度上接近于输入的低空间分辨率高光谱图像。为了进一步利用空间光谱信息，^[37]提出了一种基于深度信息蒸馏网络的内部融合操作。随后，^[38]利用协作非负矩阵分解 (CNMF) 的迁移学习来加强低空间分辨率高光谱图像和高空间分辨率高光谱图像之间的协作。与矩阵分解的思想相同，深度特征矩阵分解 (DFMF)^[39]将深度神经网络提取的特征矩阵与 NMF 策略相结合，用于超分辨率真实场景高空间分辨率高光谱图像。一般而言，这些方法中的随机共振问题都是采用多步优化策略来解决的。首先，针对高光谱波段的冗余性，对原始高光谱图像数据进行波段选择。然后，在某种程度上，关键波段将由卷积神经网络进行超分辨，以提高其空间分辨率。最后，后处理策略利用高分辨率关键频段恢复高空间分辨率高光谱图像并对其光谱信息进行校正。虽然这些方法取得了很好的效果，但它们严重依赖人工处理，耗费大量时间。因此，已经提出了一些端到端方式实现高空间分辨率高光谱图像的方法。^[40]设计了一种三维全细胞神经网络 (3D-FCNN) 方法，既利用了相邻像素的空间背景信息，又利用了相邻波段的光谱相关性。虽然三维卷积可以同时捕捉空间和光谱相关性，但计算复杂度相对较高，特别是在大输入时。针对常规 2D 卷积引起的频谱混乱，^[41]提出的分组深递归残差网络 (GDDRN) 在全局残差结构中嵌入分组递归模块，并加入联合损失，以减少数值误差和频谱失真。同样，可分离光谱和初始网络 (SSIN)^[42]通过可分离光谱卷积独立地提取每个波段图像

的特征，并以从粗到细的方式进行融合。最近，为了挖掘高光谱图像的空间特征，设计了一种新的混合卷积网络 (MCNet)^[43]，通过 2D/3D 卷积提取潜在特征。然而，这些端到端的方法都是直接对原始高光谱图像的数据进行处理，将其作为普通的数据立方体进行分辨率提升，忽略了高光谱图像的频谱知识。在这种情况下，它们不能很好地保存光谱信息。因此，根据自动编码器进行的变换，我们收集解码器中的基本光谱作为先验信息，并使用扩展的反投影校正网络来学习低空间分辨率和高空间分辨率丰度图之间的映射函数。

3 本文方法

3.1 本文方法概述

本文所使用方法与传统融合方法的对比如图 1 所示：

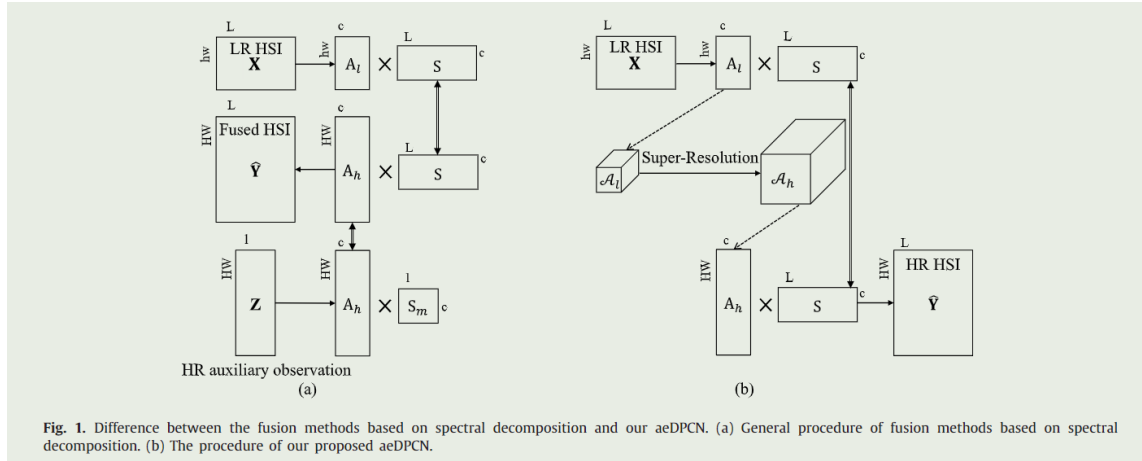


Fig. 1. Difference between the fusion methods based on spectral decomposition and our aeDPCN. (a) General procedure of fusion methods based on spectral decomposition. (b) The procedure of our proposed aeDPCN.

图 1: 基于谱分解的融合方法与我们的 aeDPCN 的区别。(a) 基于谱分解的融合方法的一般程序。(b) 我们提出的 aeDPCN 的程序。

问题给定一个低空间分辨率高光谱图像， $X \in \mathbb{R}^{h \times w \times L}$ 其中图像有 L 个通道， h 和 w 分别代表每个通道中的高和宽，用于高光谱的单图像的超分辨率的目标是估计相应的高空间分辨率高光谱图像， $\hat{Y} \in \mathbb{R}^{H \times W \times L}$ 其中 \hat{Y} 代表 X 经过模型后得到提升的高空间分辨率高光谱图像。其中提升的倍数为 s ， $H = s \times h$ ， $W = s \times w$ 。使得 \hat{Y} 尽可能的接近真实的高空间分辨率高光谱图像 $Y \in \mathbb{R}^{H \times W \times L}$ 。为了解决这个问题，我们将端元谱作为先验信息，在丰度域中重新定义超分辨率问题。为了便于矩阵分解，我们将 3D 高光谱图像立方体展开为二维矩阵，即二维矩阵的每一行表示给定像素的光谱反射率。低空间分辨率高光谱图像数据可改写为 $X \in \mathbb{R}^{hw \times L}$ 。线性光谱混合模型由于其物理有效性和数学简单性，被广泛应用于高光谱图像的解混问题中。因此，我们假设展开高光谱图像数据的每一行都是 c 个端元谱的线性组合，低空间分辨率高光谱图像非负矩阵可以分解为：

$$X = A_l S \quad (1)$$

其中， $A_l \in \mathbb{R}^{hw \times c}$ 是丰度矩阵，每个行向量表示该像素处所有端元的丰度系数， $S \in \mathbb{R}^{c \times L}$ 表示端元谱， c 表示有 c 的端元数量。由于高分辨率高光谱图像和低分辨率高光谱图像的观测描述了相同的场景，因此底层材料 (即端元) 应该是相同的。因此，如果我们提高丰度矩阵的空间分辨率，共享相同的端元谱，则对应的高分辨率高光谱图像 $Y \in \mathbb{R}^{HW \times L}$ 可重构为：

$$\hat{y} = A_h S \quad (2)$$

其中 $A_h \in \mathbb{R}^{HW \times c}$ 为高分辨率高光谱图像的预测丰度矩阵。为此，我们的目标是学习低分辨率高光谱图

像和高分辨率高光谱图像丰度矩阵之间的映射函数。由于丰度矩阵表明了光谱基在特定空间位置的混合情况，因此它们仍然保留了原始高光谱图像的空间结构。为了获取其空间关系，我们将丰度矩阵折叠为 $A_l \in \mathbb{R}^{h \times w \times c}$ 和 $A_{limits_h} \in \mathbb{R}^{H \times W \times c}$ 。因此，映射函数为：

$$A_h = M(A_l; \theta) \quad (3)$$

通过这种方式，我们将图像域的超分问题转化为丰度域，从而降低了模型的复杂性。利用端元光谱作为先验信息，在提高空间分辨率的同时，尽可能多地保留光谱细节。

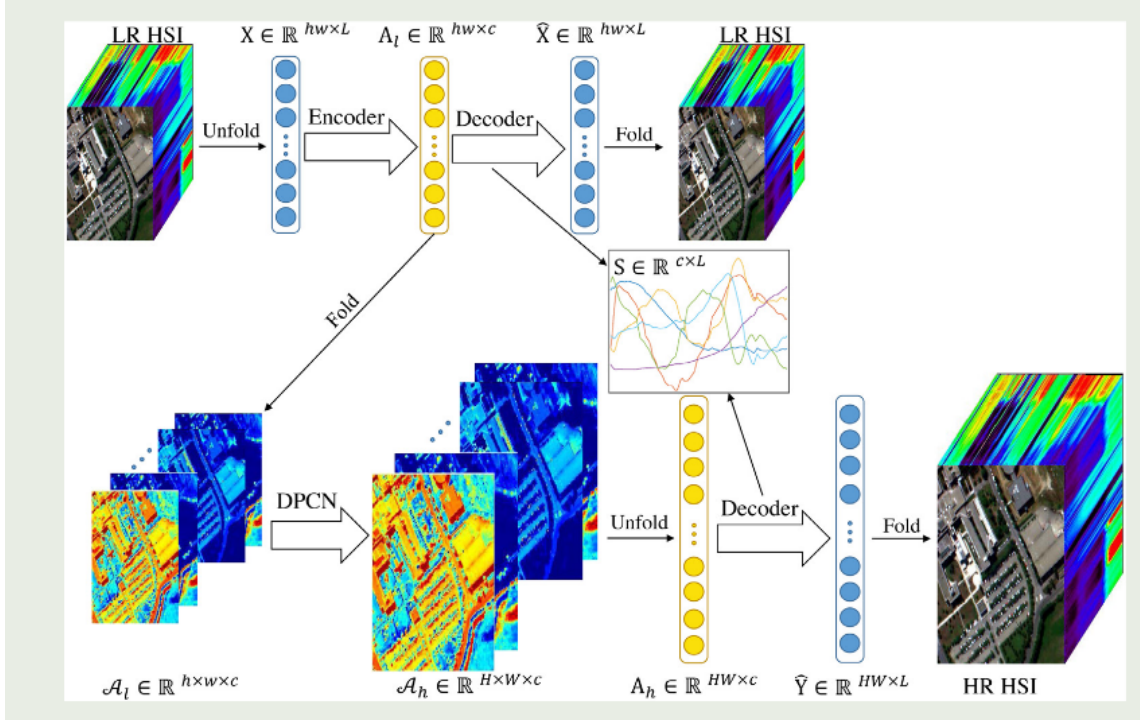


图 2: 我们提出的 aeDPCN 的整个示意图。其中 DPCN 表示空洞的反向投影校正网络。

图 2 描述了我们提出的 aeDPCN 的示意图。首先将三维的低空间分辨率高光谱图像转换成二维矩阵进行谱分解，该分解由自动编码器网络实现。为了将图像域问题转换到丰度域，编码器试图生成满足和为 1 且非负约束的丰度矩阵。因此，我们可以将数据的维数从 $\mathbb{R}^{hw \times L}$ 降至 $\mathbb{R}^{hw \times c}$ 其中 $c \ll l$ 。此外，受非负矩阵分解的基本思想启发，利用解码器网络表示端元矩阵以重构高光谱图像，同时保留高光谱图像中包含的光谱信息。然后，我们将低空间分辨率高光谱图像的丰度表示作为高光谱图像的嵌入。由于高光谱图像的空间结构在丰度域中仍然保持不变，因此将高光谱图像嵌入折叠并输入到空洞反向投影校正网络中以提高空间分辨率。超分辨嵌入给出了 HR 丰度系数的估计。由于高空间分辨率高光谱图像和低空间分辨率高光谱图像是从同一场景中采样的，因此通过共享相同的解码器，在丰度表示上乘以相同的端元矩阵来生成高空间分辨率高光谱图像。

3.2 基于自编码器的谱分解

基于自动编码器的网络被广泛用于降维和表示学习，试图逼近身份映射，从而使输出接近输入。该体系结构通常包括用于将输入数据映射到低维表示的编码器和用于重构的解码器，配置的过程如同 1，该体系适用于并且常用与高光谱解混问题^{[44][45][46]}。通常，丰度矩阵和端元矩阵可由该结构产生或由其参数表示。为了解决 3 中的 SR 问题，我们将原始输入编码为低维丰度表示：

$$A_l = f_e(X; \theta_e) \quad (4)$$

其中 f_e 表示参数为 θ_e 的编码器的操作。编码器网络由几个完全连接的层组成。激活函数必须是非线性的，否则编码器将简单地执行降维。基于线性光谱混合模型，丰度代表光谱基的相对分量。因此，它们必须满足非负约束和矩阵只和为 1 的约束，即

$$a_{ij} \geq 0, \forall i, j \quad (5)$$

$$A_l = f_e(X; \theta_e) \quad (6)$$

对于非负约束，一些著作引入了阈值或非负自编码器^[45]强制向量为非负。此外，正则化^[45]，或标准化^[45]使用算子来保证和到一的约束。为了避免网络的复杂设计，我们在编码器的末端使用 ReLU 激活函数来强制编码器的输出为非负的，并且这个非负向量由其条目的和归一化，即

$$A_{ij} = \frac{A_{ij}}{\sum_{j=1}^c A_{ij}} \quad (7)$$

在获得低维嵌入后，解码器试图通过

$$\hat{X} = f_d(A_l; \theta_d) \quad (8)$$

其中 f_d 表示参数为 θ_d 的解码器的操作。与式 2 相比，很明显解码器必须是线性变换 $f_{limits_d} : \mathbb{R}^{hw \times c} \rightarrow \mathbb{R}^{hw \times L}$ 。因此，解码器的权重有望成为携带光谱信息的端元矩阵 S 。因此我们有

$$S = \theta_d \quad (9)$$

为此，以丰度矩阵 A 为输入，解码器网络重构高光谱数据，模拟方程 1 和 2。通过基于自编码器的网络，我们将单个高光谱图像超分问题从图像域转化为丰度域，并在超分重构之前保留光谱信息。由于高空间分辨率高光谱图像和低空间分辨率高光谱图像都捕获了相同的成像场景，因此共享相同的解码器是合理的。

3.3 超分的空洞投影校正网络

该网络的目标是将低空间分辨率高光谱图像丰度表示映射到高空间分辨率高光谱图像空间中。由于原始数据的空间关系仍然保留在丰度矩阵中，我们将其折叠成三维数据到二维，以方便卷积层进行特征提取。2 所示，在比例因子为 4 中描述超分网络，我们的模型有两个基本部分：用于提高空间分辨率的扩展特征重建模块和用于光谱校正的反向投影校正模块。

3.3.1 空洞特征重建

受^[32]等人的启发。我们以从粗到细的方式对低空间分辨率高光谱图像嵌入进行超分。具体来说，高空间分辨率高光谱图像丰度是在 $\log_2 s$ 级别逐步预测的，其中 s 是比例放大倍数。在第 k 层，对于图 2(b) 中的扩张特征重建模块，我们首先通过简单地堆叠卷积层来提取深度特征。给定 itA_{k-1} 作为第 k 层的输入，我们有

$$F_k = H(\cdots (H(A_{k-1}))) \quad (10)$$

其中 F_k 表示低空间分辨率高光谱图像空间在该层次上的深度特征， $H(\cdot)$ 表示卷积运算后的 ReLU 激活函数。在高光谱图像中，像素中物质的丰度与相邻像素的丰度密切相关^[47]。由于超分网络处理的是丰度系数而不是像素值，因此需要考虑大尺度依赖性。因此，为了在有限的层中获得更大的感受野，我们利用空洞卷积来获得更多信息的特征。空洞卷积以其在保留传统卷积优点的同时扩大接受野容量

而闻名。具体来说，二维空洞卷积是通过在卷积核中插入“孔”(零)而不添加参数来实现的。对于大小为 3×3 的卷积核，得到的膨胀滤波器大小为 $(2d+1) \times (2d+1)$ ，膨胀率为 d 。在标准卷积中， $d=1$ 。当 $d>1$ 时，在相同的参数下，空洞卷积会有更大的感受野。然而，对所有层使用相同的膨胀率可能会导致网格化，丢失大量信息。为了避免这种恶化，采用混合空洞卷积^[48]覆盖一个没有任何孔或缺失边的正方形区域。实际上，七层叠加来提取深层特征，这遵循锯齿波般的重复。因此，每层的等效接受场分别为 3、5、7、9、7、5 和 3。因此，可以很容易地得到各层的感受场为 33×33 。如果使用传统的 3×3 卷积层，网络的接受场大小为 15×15 且网络深度相同(即 7)，或者接受场深度为 16 且接受场相同(即 33×33)。这样，我们的模型可以在充分利用丰度图中的空间信息的同时，在接受野的大小和网络深度之间进行权衡。为了提高空间分辨率，本文根据高阶层次，使用了一个标度为 2 的上采样模块，将深度表示映射到高空间分辨率空间中。因此，上阶残差由

$$R_k = U(F_k) \quad (11)$$

其中 $U(\cdot)$ 对应基本因子为 2 的上采样操作。我们利用 PixelShuffle^[49]操作进行上采样，然后使用卷积层进行上采样过程。以残差方式重建中间丰度图 A_k ，即

$$A_k = A_{k-1} \uparrow + R_k \quad (12)$$

其中 $A_{k-1} \uparrow$ 指输入的双三次线性插值的上采样图像。该模块的输出可视为具有上采样倍数 2 的中间结果，并馈送到下一层进行进一步重构。与^[32]不同，我们的模型没有以监督的方式约束中间结果。因此，这种渐进式结构试图简化学习过程，特别是在一个大尺度的放大倍数下。

3.3.2 反向投影模块

在大多数情况下，超分辨方法通常将上采样模块部署为端到端网络的最后一部分。一方面，我们将超分辨问题转化为丰度域，但生成的丰度表示没有直接监督；另一方面，保留端元谱的解码器重构低空间分辨率和高空间分辨率的高光谱图像。因此，高空间分辨率高光谱图像丰度图应与低空间分辨率高光谱图像具有相似的光谱模式。在这种情况下，我们设计了背投影校正模块，以隐式约束丰度表示，其灵感来自经典的迭代背投影方法^[50]。在该校正模块中，将下采样粗丰度与原始低空间分辨率高光谱图像之间的差值反投影到高空间分辨率高光谱图像空间中，以残差的方式校正失真。同时，我们学习每个位置上所有分量的线性组合，以利用邻近丰度值之间的谱相关性。因此，低空间分辨率高光谱图像误差将从频谱方面指导失真校正。利用可学习校正策略，生成的高空间分辨率高光谱图像丰度表示可以调整到与低空间分辨率高光谱图像对应的相似模式，从而减少同一频谱重构引起的失真。反向投影校正模块的详细结构如图 2(c) 所示。该模块高空间分辨率高光谱图像以粗略的表示 A_h 为输入，首先将超分辨丰度投影到低空间分辨率高光谱图像空间，即

$$S = H(D(\hat{A}_h)) \quad (13)$$

其中 S 为高空间分辨率高光谱图像空间中粗丰度的下采样对应物， D 为下采样操作， H 为带有 ReLU 的卷积层。在本文中，我们使用了 PixelShuffle 的反向操作，然后使用卷积层来进行下采样过程。为了校正畸变，观测到的低空间分辨率高光谱图像映射的 S 与原始 \hat{A}_l 之间的差值通过逐元素相减得到

如下:

$$e_l = A_l - S \quad (14)$$

理想情况下, 投影低空间分辨率高光谱图像映射的 S 与原始 A_l 之间的差异应该尽可能小。在这种情况下, 粗略丰度的下采样版本可能等于低空间分辨率高光谱图像丰度。然后, 我们将差值映射到高空间分辨率高光谱图像空间, 获取残差进行反向校正:

$$e_h = U(H(e_l)) \quad (15)$$

其中 u 对应上采样操作, h 表示一个卷积层与 ReLU。最后, 粗略的预测可以通过以下方式进行修正:

$$A_h = \hat{A}_h + e_h \quad (16)$$

通过反投影校正策略, 低空间分辨率高光谱图像和高空间分辨率高光谱图像嵌入可以具有相似的模式, 通过相同的端元频谱来预测像素值。类似地, 我们在高空间分辨率空间中对生成的丰度应用 RELU 函数和归一化来满足非负和一的约束。

3.4 损失函数的设计

由于高光谱图像数据同时具有空间和光谱特征, 我们采用了一个联合函数来约束与地面真实情况相似的预测图像, 该函数被定义为重建损失项:

$$L_{rec} = L_{spatial} + \lambda L_{spectral} \quad (17)$$

其中, λ 是平衡空间损耗和频谱差异之间的权衡的参数。由于 L1 损失已被证明比 L2 损失更有效^[26], 空间损失由 L1 范数构成, 其定义如下:

$$L_{spatial} = \frac{1}{N} \sum_{k=1}^N \|I^n - \hat{I}^n\| \quad (18)$$

其中, \hat{I}^n 和 I^n 分别是重建的高空间分辨率高光谱图像和地面真实高空间分辨率高光谱图像, N 表示一个训练批次中的图像数量。我们利用基于光谱角度映射器 (SAM)^[51]标准的测量^[41]来评估光谱相似性, 即:

$$L_{spectral} = \frac{1}{\pi} \frac{1}{N} \sum_{k=1}^N \arccos\left(\frac{I^n \cdot \hat{I}^n}{\|I^n\|_2 \|\hat{I}^n\|_2}\right) \quad (19)$$

其中 \cdot 是张量积, $\|\cdot\|_2$ 是张量范数。通过对这两个方面的约束, 我们可以同时显著地减小数值误差和频谱失真。

4 复现细节

4.1 与已有开源代码对比

本次实现并没有参考任何他人的代码

4.2 训练过程

Procedure 1 training procedure of aeDPCN

Input: Low-resolution HSI X

Output: High-resolution HSI Y

for k_1 epochs **do**

 Evaluate loss Eqs17

 Update parameters of auto-encoder by AdamOptimizer

end

Freeze the parameters of auto-encoder and use it to guide the training of DPCN

for K_2 epochs **do**

 Evaluate loss Eqs18

 Update parameters of DPCN by AdamOptimizer

end

for K_3 epochs **do**

 Evaluate loss Eqs19

 Update parameters of aeDPCN by AdamOptimizer

end

Return $= M_{aedpcn}(X)$;

4.3 参数设置以及数据集的制作

根据数据集，该网络的输入和输出节点都是高光谱数据中相同数目的谱带，记为 L 。我们将除反投影校正模块外的所有卷积层的滤波大小设置为 3×3 ，其中核大小设置为 1×1 。为了确保特征图的大小不变，对核大小为 3×3 的卷积层采用 padding=1 策略。对于扩展重建模块，除最后一层外，用于特征提取的卷积层具有 64 个滤波器，即该层的输出通道与用于剩余学习的端元谱 c 的数目相同。对于损失函数，我们设置 $\lambda=0.1$ 。在训练阶段，我们经验地选择了 16 个小批次，并使用了权值衰减为 $1E-4$ 的 ADAM 优化器。初始学习率设置为 0.0001，2000 个 epoch 后衰减 10 倍，而总历元为 3000 个 epoch。我们的模型是在 Nvidia Tesla P100 上用 PyTorch 实现的。提出的方法在数据集上进行了评估：Chikusei^[6]数据集。Chikusei 数据集是由 Headwall Hyperspecc-VNIR-C 成像传感器在日本茨城 Chikusei 的农业和城市地区拍摄的航空高光谱数据集。高光谱场景由 2517×2335 个像素组成，128 个波段，光谱范围从 363 nm 到 1018 nm。我们首先提取这幅图像的顶部区域，形成测试数据，其中有四幅不重叠的高光谱图像，像素为 $400 \times 400 \times 128$ 。这幅图像的剩余区域被剪裁成 400×400 个子图像，重叠 200 个像素。10% 的子图被随机划分为验证数据集。为了丰富训练数据集，通过以 0.5 的概率随机旋转和翻转来扩大其余的子图像。为了在有限的层内获取更多的信息，我们采用空洞卷积进行特征提取。由于空洞卷积，我们的方法可以在有限的层中获取深层空间信息，这有助于获得具有相当性能的轻量级模型。考虑到共享保存光谱信息的同一解码器，超分辨丰度应该具有与低空间分辨率丰度相似的图案。因此，我们设计了反投影校正模块来校正光谱失真。

4.4 评价指标

为了进行定量比较，我们使用四个质量度量来评估预测的高空间分辨率高光谱图像 \hat{Y} 与其对应的地面真实图像 Y ，包括峰值信噪比 (PSNR)、误差相对全局尺度合成 (ERGAS)^[52]、光谱角度映射器 (SAM)^[51] 和结构相似性 (SSIM)^[53]。广泛使用的 PSNR 指标是图像的最大功率与所有光谱波段的残差

功率之比的平均值。对于第 i 个频段，峰值信噪比通过以下公式计算：

$$PNSR(y, \hat{y}) = 10 \cdot \log_{10} \left(\frac{\max(y^2)}{\text{mean} \|\hat{y}_i - y_i\|_2^2} \right) \quad (20)$$

较高的 PSNR 值表示重建 HSI 的图像质量较好。ERGAS 测量参考的高空间分辨率高光谱图像 Y 和重建的高空间分辨率高光谱图像 \hat{Y} 之间的带状归一化均方误差根，最佳值为 0。它被定义为：

$$ERGAS(y, \hat{y}) = \frac{100}{s} \sqrt{\frac{1}{L} \sum_{i=1}^n \frac{\text{mean}(\|y_i - \hat{y}_i\|_2^2)}{\text{mean}(y_i^2)}} \quad (21)$$

其中 s 是低空间分辨率高光谱图像和高空间分辨率高光谱图像之间的比例因子，而 L 是高光谱图像数据中的谱带数量。SAM 通常用于量化每个像素的光谱信息保留量。更准确地说，该指数计算估计光谱和参考光谱的两个向量之间的夹角，以表示光谱相似性。接近零的 SAM 值表明光谱相似度高，没有光谱失真。SAM 定义为：

$$SAM(y, \hat{y}) = \arccos \left(\frac{y \cdot \hat{y}}{\|y\|_2 \|\hat{y}\|_2} \right) \quad (22)$$

另一个广为人知的指数是 SSIM。对于第 i 个频段，其定义为：

$$SSIM(y, \hat{y}) = \frac{(2\mu_{y_i} \mu_{\hat{y}_i} + c_1)(2\sigma_{y_i, \hat{y}_i} + c_2)}{(\mu_{y_i}^2 + \mu_{\hat{y}_i}^2 + c_1)(\sigma_{y_i}^2 + \sigma_{\hat{y}_i}^2 + c_2)} \quad (23)$$

其中 μ_{y_i} 和 $\mu_{\hat{y}_i}$ 分别是 y_i 和 \hat{y}_i 的均值， σ_{y_i} 和 $\sigma_{\hat{y}_i}$ 分别是 y_i 和 \hat{y}_i 的方差， σ_{y_i, \hat{y}_i} 是 y_i 和 \hat{y}_i 的协方差， c_1 和 c_2 分别是设置为 0.0001 和 0.0009 的常量。SSIM 的最佳值为 1。通过平均 HSI 数据中所有频带的 SSIM 值来估计平均 SSIM。

4.5 一些改进方案

论文浮现结果如图 3，左边左上角图片为原论文复现后的结果。模型收敛之后的效果非常差，主要原因我认为应该有两点。第一：DPCN 中的卷积层过于简单，表达能力有限。第二：线性解混层的复杂度不足以支撑解混的精度。因此我对卷积层进行了更改，将两层逐步上采样的空洞卷积替换成逐步上采样的 resnet50，将线性解混层的全连接替换成卷积，并添加 Relu 函数保证了他的 $x > 0$ 的范围是线性的，最后结果采用 sum-to-one 的归一化。重新训练相同数据集后结果如图 3 左边中间所示，其在测试集上结果如图 3 左边左下角所示。虽然还是和原论文有差距，但是证明解混超分是可以做的，并且更复杂的丰度域的表达更有利于最后的结果。

5 实验结果分析

实验分析见 4.5

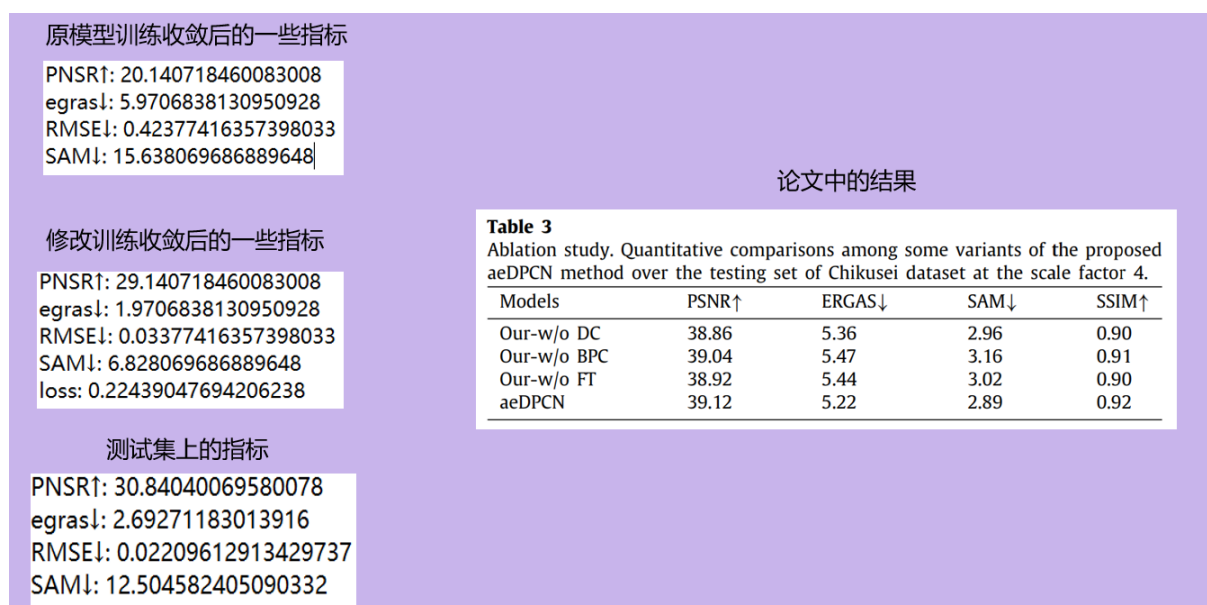


图 3: 实验结果示意

6 总结与展望

丰度域高光谱超分是一个值得研究的课题。基于丰度域的超分是基于解混的。因此解混的好坏很大程度上决定了超分的结果。因此使用非线性解混生成的丰度应当比线性解混生成的丰度更加具有可信度。在丰度超分的过程中,我认为有监督的丰度超分一定比无监督的丰度超分更加具有说服力。收敛的也更快。多源遥感是一个非常热门的课题。我们可以通过 RGB, Pansharping, multispectral image 等多源遥感数据的融合的角度来进行超分。

参考文献

- [1] LIU T, GU Y, CHANUSSOT J, et al. Multimorphological Superpixel Model for Hyperspectral Image Classification[J]. IEEE Transactions on Geoscience and Remote Sensing, 2017.
- [2] WU H, PRASAD S. Semi-Supervised Deep Learning Using Pseudo Labels for Hyperspectral Image Classification[J]. IEEE Transactions on Image Processing, 2018.
- [3] DU B, ZHANG L. A Discriminative Metric Learning Based Anomaly Detection Method[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014.
- [4] XIE W, LI Y, LEI J, et al. Unsupervised spectral mapping and feature selection for hyperspectral anomaly detection.[J]. Neural Networks, 2020.
- [5] PLAZA A, DU Q, BIOUCAS-DIAS J M, et al. Foreword to the Special Issue on Spectral Unmixing of Remotely Sensed Data[J]. IEEE Transactions on Geoscience and Remote Sensing, 2011.
- [6] YOKOYA N, GROHNFELDT C, CHANUSSOT J. Hyperspectral and Multispectral Data Fusion: A comparative review of the recent literature[J]. IEEE Geoscience and Remote Sensing Magazine, 2017.
- [7] LIU D, WEN B, FAN Y, et al. Non-local recurrent network for image restoration[J]. neural information processing systems, 2018.

- [8] WEN B, KAMILOV U S, LIU D, et al. Deepcasd: An End-to-End Approach for Multi-Spectral Image Super-Resolution[J]. international conference on acoustics, speech, and signal processing, 2018.
- [9] AKHTAR N, SHAFAIT F, MIAN A. Bayesian sparse representation for hyperspectral image super resolution[J]. computer vision and pattern recognition, 2015.
- [10] DONG W, FU F, SHI G, et al. Hyperspectral Image Super-Resolution via Non-Negative Structured Sparse Representation[J]. IEEE Transactions on Image Processing, 2016.
- [11] QU Y, QI H, KWAN C. Unsupervised Sparse Dirichlet-Net for Hyperspectral Image Super-Resolution [J]. computer vision and pattern recognition, 2018.
- [12] SIMOES M, BIOUCAS-DIAS J M, ALMEIDA L B, et al. A convex formulation for hyperspectral image superresolution via subspace-based regularization[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014.
- [13] VEGANZONES M A, SIMOES M, LICCIARDI G, et al. Hyperspectral Super-Resolution of Locally Low Rank Images From Complementary Multisource Data[J]. IEEE Transactions on Image Processing, 2016.
- [14] WEI Q, BIOUCAS-DIAS J M, DOBIGEON N, et al. Hyperspectral and Multispectral Image Fusion based on a Sparse Representation[J]. IEEE Transactions on Geoscience and Remote Sensing, 2014.
- [15] JIANG J, SUN H, LIU X, et al. Learning Spatial-Spectral Prior for Super-Resolution of Hyperspectral Imagery[J]. IEEE Transactions on Computational Imaging, 2020.
- [16] WANG X, MA J, JIANG J. Hyperspectral Image Super-Resolution via Recurrent Feedback Embedding and Spatial-Spectral Consistency Regularization[J]. IEEE Transactions on Geoscience and Remote Sensing, 2021.
- [17] HE S, ZHOU H, WANG Y, et al. Super-resolution reconstruction of hyperspectral images via low rank tensor modeling and total variation regularization[J]. international geoscience and remote sensing symposium, 2016.
- [18] HUANG H, YU J, SUN W. Super-resolution mapping via multi-dictionary based sparse representation [J]. international conference on acoustics, speech, and signal processing, 2014.
- [19] LANARAS C, BALTSAVIAS E P, SCHINDLER K. Hyperspectral Super-Resolution by Coupled Spectral Unmixing[J]. international conference on computer vision, 2015.
- [20] YOKOYA N, YAIRI T, IWASAKI A. Coupled Nonnegative Matrix Factorization Unmixing for Hyperspectral and Multispectral Data Fusion[J]. IEEE Transactions on Geoscience and Remote Sensing, 2012.
- [21] DONG C, LOY C C, HE K, et al. Image Super-Resolution Using Deep Convolutional Networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2014.

- [22] HE K, ZHANG X, REN S, et al. Deep Residual Learning for Image Recognition[J]. computer vision and pattern recognition, 2015.
- [23] KIM J, LEE J K, LEE K M. Accurate Image Super-Resolution Using Very Deep Convolutional Networks [J]. computer vision and pattern recognition, 2015.
- [24] KIM J, LEE J K, LEE K M. Deeply-Recursive Convolutional Network for Image Super-Resolution[J]. computer vision and pattern recognition, 2015.
- [25] TAI Y, YANG J, LIU X. Image Super-Resolution via Deep Recursive Residual Network[J]. computer vision and pattern recognition, 2017.
- [26] LIM B O, SON S, KIM H, et al. Enhanced Deep Residual Networks for Single Image Super-Resolution [J]. computer vision and pattern recognition, 2017.
- [27] TAI Y, YANG J, LIU X, et al. MemNet: A Persistent Memory Network for Image Restoration[J]. international conference on computer vision, 2017.
- [28] ZHANG Y, LI K, LI K, et al. Image Super-Resolution Using Very Deep Residual Channel Attention Networks[J]. european conference on computer vision, 2018.
- [29] GOODFELLOW I, POUGET-ABADIE J, MIRZA M, et al. Generative Adversarial Nets[J]. neural information processing systems, 2014.
- [30] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-Realistic Single Image Super-Resolution Using a Generative Adversarial Network[J]. computer vision and pattern recognition, 2022.
- [31] WANG X, YU K, WU S, et al. ESRGAN: Enhanced Super-Resolution Generative Adversarial Networks [J]. european conference on computer vision, 2018.
- [32] LAI W S, HUANG J B, AHUJA N, et al. Deep Laplacian Pyramid Networks for Fast and Accurate Super-Resolution[J]. computer vision and pattern recognition, 2017.
- [33] HARIS M, SHAKHNAROVICH G, UKITA N. Deep Back-Projection Networks For Super-Resolution [J]. computer vision and pattern recognition, 2018.
- [34] HU J, LI Y, XIE W. Hyperspectral Image Super-Resolution by Spectral Difference Learning and Spatial Error Correction[J]. IEEE Geoscience and Remote Sensing Letters, 2017.
- [35] XIE W, LI Y, HU J, et al. Trainable spectral difference learning with spatial starting for hyperspectral image denoising.[J]. Neural Networks, 2018.
- [36] LI Y, HU J, ZHAO X, et al. Hyperspectral image super-resolution using deep convolutional neural network[J]. Neurocomputing, 2017.
- [37] HU J, ZHAO M, LI Y. Hyperspectral Image Super-Resolution by Deep Spatial-Spectral Exploitation [J]. Remote Sensing, 2019.

- [38] YUAN Y, ZHENG X, LU X. Hyperspectral Image Superresolution by Transfer Learning[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2017.
- [39] XIE W, JIA X, LI Y, et al. Hyperspectral Image Super-Resolution Using Deep Feature Matrix Factorization[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019.
- [40] MEI S, YUAN X, JI J, et al. Hyperspectral Image Spatial Super-Resolution via 3D Full Convolutional Neural Network[J]. Remote Sensing, 2017.
- [41] LI Y, ZHANG L, DING L C, et al. Single Hyperspectral Image Super-Resolution with Grouped Deep Recursive Residual Network[J]. IEEE International Conference on Multimedia Big Data, 2018.
- [42] ZHENG K, GAO L, RAN Q, et al. Separable-spectral convolution and inception network for hyperspectral image super-resolution[J]. International Journal of Machine Learning and Cybernetics, 2019.
- [43] LI Q, WANG Q, LI X. Mixed 2D/3D Convolutional Network for Hyperspectral Image Super-Resolution [J]. Remote Sensing, 2020.
- [44] PALSSON B, SIGURDSSON J, SVEINSSON J R, et al. Hyperspectral Unmixing Using a Neural Network Autoencoder[J]. IEEE Access, 2018.
- [45] SU Y, LI J, PLAZA A, et al. DAEN: Deep Autoencoder Networks for Hyperspectral Unmixing[J]. IEEE Transactions on Geoscience and Remote Sensing, 2019.
- [46] XIE W, LEI J, LIU B, et al. Spectral constraint adversarial autoencoders approach to feature representation in hyperspectral anomaly detection.[J]. Neural Networks, 2019.
- [47] IRMAK H, AKAR G B, YUKSEL S E. A MAP-Based Approach for Hyperspectral Imagery Super-Resolution[J]. IEEE Transactions on Image Processing, 2018.
- [48] WANG P, CHEN P, YUAN Y, et al. Understanding Convolution for Semantic Segmentation[J]. workshop on applications of computer vision, 2018.
- [49] SHI W, CABALLERO J, HUSZAR F, et al. Real-Time Single Image and Video Super-Resolution Using an Efficient Sub-Pixel Convolutional Neural Network[J]. computer vision and pattern recognition, 2016.
- [50] IRANI M, PELEG S. Improving resolution by image registration[J]. CVGIP: Graphical Models and Image Processing, 1991.
- [51] KRUSE F A, LEFKOFF A B, BOARDMAN J W, et al. The spectral image processing system (SIPS) interactive visualization and analysis of imaging spectrometer data[J]. Remote Sensing of Environment, 1993.
- [52] ZENG Y, HUANG W, LIU M, et al. Fusion of satellite images in urban area: Assessing the quality of resulting images[J]. international conference on geoinformatics, 2010.
- [53] WANG Z, BOVIK A C, SHEIKH H R, et al. Image quality assessment: from error visibility to structural similarity[J]. IEEE Transactions on Image Processing, 2004.