

# 基于深度学习的像素级全色图像锐化研究综述

杨勇<sup>1</sup>, 苏昭<sup>1</sup>, 黄淑英<sup>2</sup>, 万伟国<sup>3</sup>, 涂伟<sup>1</sup>, 卢航远<sup>1</sup>

1. 江西财经大学 信息管理学院, 南昌 330032;
2. 天津工业大学 软件学院, 天津 300387;
3. 江西财经大学 软件与物联网工程学院, 南昌 330032

**摘要:** 全色图像锐化是遥感数据处理领域的一个基础性问题, 在地物分类、目标识别等方面具有重要的研究意义和应用价值。近年来, 深度学习在自然语言处理、计算机视觉等领域取得了巨大进展, 也推动了像素级全色图像锐化技术的发展。本文提出从经典方式和协同方式两个方面对深度学习在全色图像锐化中的研究进行系统的综述, 并在此基础上进行前景展望。首先, 给出全色图像锐化常用的数据集和全色图像锐化的质量评价指标; 接着, 从经典方式与协同方式两个方面对基于深度学习的全色图像锐化最新研究成果进行分门别类的介绍, 并进行算法性能的对比、分析和归纳; 然后, 对全色图像锐化的3个主要应用领域如地物分类、目标识别和地表变化检测进行分析; 最后, 本文探讨了基于深度学习的全色图像锐化的5个未来研究方向。

**关键词:** 全色图像锐化, 深度学习, 经典方式, 协同方式, 质量评价, 遥感图像融合

**中图分类号:** P2

**引用格式:** 杨勇, 苏昭, 黄淑英, 万伟国, 涂伟, 卢航远. 2022. 基于深度学习的像素级全色图像锐化研究综述. 遥感学报, 26(12): 2411-2432

Yang Y, Su Z, Huang S Y, Wan W G, Tu W and Lu H Y. 2022. Survey of deep-learning approaches for pixel-level pansharpening. National Remote Sensing Bulletin, 26(12): 2411-2432 [DOI: 10.11834/jrs.20211325]

## 1 引言

由于受遥感卫星传感器设计、遥感成像机理等因素的影响, 单源遥感图像在空间、光谱分辨率等方面相互制约, 一般遥感卫星只能获得单幅低空间分辨率的MS (Multispectral) 图像或高空间分辨率的PAN (Panchromatic) 图像 (Vivone 等, 2015; Li 等, 2020a); 然而, 在实际遥感应用中, 如地物分类 (Aguilar 等, 2013; Gianinetto 等, 2017)、目标识别 (Chen 等, 2019; Mohammadzadeh 等, 2006) 和地表变化检测 (Chen 等, 2018; Souza 等, 2003) 等任务往往需要高空间分辨率多光谱HRMS (High Spatial Resolution Multispectral) 的图像。全色图像锐化又称Pansharpening是将PAN图像的高分辨率空间细节信息与MS图像的丰富光谱信息进行融合, 得到高质量、理想的HRMS图像 (陈毛毛 等, 2021)。因此, 如何有效提取MS图像

和PAN图像的相关信息, 并通过融合得到HRMS图像, 一直都是遥感图像处理领域亟待解决的重要问题 (Yang 等, 2018b)。

遥感图像融合起源于20世纪70年代, 按照不同的处理级别, 可以划分为3个层次: 像素级融合、特征级融合和决策级融合。其中, 像素级融合是直接原始遥感图像各像素上的直接融合处理, 其目的是为了获得质量更高的融合图像, 如提升观测图像的分辨率、增强原始图像的清晰度等。像素级融合在3个层次的融合中应用范围最为广阔、研究成果最为丰富, 也是目前研究的热点之一 (李树涛 等, 2021; Zhou 等, 2022)。因此, 本文将立足于像素级层次的全色图像锐化展开综述。

像素级全色图像锐化方法通常分为3类 (Yang 等, 2018a): (1) 成分替换CS (Component Substitution)法 (Yang 等, 2018c; Liu 等, 2018a;

收稿日期: 2021-05-21; 预印本: 2022-07-07

基金项目: 国家自然科学基金 (编号: 62072218, 61862030); 江西省自然科学基金 (编号: 20192ACB20002, 20192ACBL21008)

第一作者简介: 杨勇, 研究方向为遥感图像处理、计算机视觉与机器学习。E-mail: greatyangy@126.com

通信作者简介: 黄淑英, 研究方向为图像处理、人工智能。E-mail: shuyinghuang2010@126.com

Shahdoosti 和 Javaheri, 2017; Li 等, 2018)。CS 方法先将 MS 图像转换到一个新的空间, 然后在新的映射空间用 PAN 图像对转换后的 MS 图像空间信息成份进行替换 (Alparone 等, 2015), 其不足在于光谱失真较为明显; (2) 多分辨率分析 MRA (Multi-Resolution Analysis) 法 (Khan 等, 2008; Lu 等, 2017; Xing 等, 2018b)。MRA 方法首先利用多尺度变换方法, 如小波变换或者金字塔变换等, 将源图像分解获得高、低频成份, 再运用适当的融合规则对不同尺度的高、低频成份进行融合, 最后将融合后的高、低频成份反变换获得融合图像 (陈毛毛 等, 2021), 其不足在于空间细节失真较为严重; (3) 基于模型 MB (Model-Based) 的方法 (Kang 等, 2014; Palsson 等, 2015)。MB 方法根据理想的融合图像与 PAN 图像、MS 图像之间的关系建立能量函数, 并通过最优化求解获得 HRMS 融合图像, 但其计算较为复杂 (Lee 等, 2021)。此外, 也有不少学者 (Yang 等, 2018c; He 等, 2019) 将 CS、MRA 两种方法进行结合来构建注入模型融合框架。

近年来, 由于神经网络具有能够自动发现数据特征表示的能力, 它促进了自然语言处理、计算机视觉等研究取得巨大进展 (Bowman 等, 2015; He 等, 2016; Huang 等, 2017)。因此, 近 6 年来 (2015 年以来), 许多研究者对基于深度学习的像素级全色图像锐化进行了深入的研究, 并取得了丰硕的研究成果 (Yang 等, 2017; 王芬等, 2021; Lee 等, 2021; Zhou 等, 2022); 同时, 已有几位学者对深度学习在全色图像锐化领域中的研究进行了部分的总结和评述 (Liu 等, 2018c; Ma 等, 2019; Tsagkatakis 等, 2019; Javan 等, 2021)。Liu 等 (2018c) 综述了基于深度学习的各种像素级图像融合的研究进展, 包括多聚焦、医学及遥感图像融合等, 文中对 7 个代表性的基于深度学习的全色图像锐化模型进行了较为详尽的评述, 但对方法缺乏定量化的比较及分析; Ma 等 (2019) 从元分析 (Meta-analysis) 的角度对深度学习在全色图像锐化等多个应用领域进行了综述, 但对全色图像锐化方法也缺少具体的算法性能分析; Tsagkatakis 等 (2019) 从自编码器 AE (Autoencoders) 和卷积神经网络 CNN (Convolutional Neural Networks) 两个方面对基于深度学习的全色图像锐化算法进行了阐述, 但论文未涉及到深度

学习与传统方法相结合的融合模型; Javan 等 (2021) 主要从 CS、MRA、Hybrid 和 VO (Variational Optimization) 等 4 类方法对传统全色图像锐化算法进行详尽阐述, 然而对基于深度学习的全色图像锐化算法仅做了简要介绍, 缺乏对算法的定量比较及分析。总体而言, 以上文献 (Liu 等, 2018c; Ma 等, 2019; Tsagkatakis 等, 2019; Javan 等, 2021) 只是部分地或从某一角度对基于深度学习的像素级全色图像锐化研究成果进行阐述, 在系统性和定量分析方面存在一定的不足。为此, 本文将对基于深度学习的全色图像锐化研究成果进行全面和系统的归纳与总结, 建立一个较完整的领域研究视图、给出该领域的发展脉络, 并提出未来 5 个重要的研究方向。本文的创新工作包括: (1) 提出从经典方式和协同方式两个方面对基于深度学习的像素级全色图像锐化研究成果进行综述; (2) 对一些代表性的融合算法运用有参考和无参考图像质量评价指标进行了性能提升比的对比、分析和归纳, 并以此给出其发展脉络; (3) 结合深度学习和全色图像锐化的发展态势, 给出基于深度学习的像素级全色图像锐化的 5 个主要研究方向。

本文给出了遥感卫星图像常用数据集, 并对全色图像锐化质量评价进行介绍, 以经典方式与协同方式两个方面分门别类地介绍了基于深度学习的全色图像锐化方法, 同时对代表性的算法进行对比、分析和归纳, 最后对全色图像锐化的 3 个常见应用领域进行分析以及对该领域的发展前景进行展望。

## 2 遥感卫星图像数据集

全色图像锐化所采用的数据为 PAN 图像和 MS 图像, 数据集多由遥感光学卫星商业公司提供。常用的遥感图像数据集主要来自国外 Digital Globe 卫星公司的 WorldView 系列、IKONOS、QuickBird 和 GEOEYE 等卫星数据集。WorldView 系列中的 WorldView-2 卫星 (<https://gbdxdocs.digitalglobe.com/docs/worldview-2>[2021-06-18]) 2009 年发射, 能够提供 0.46 m 全色和 1.85 m 多光谱影像, 其星载多光谱传感器不仅有 4 个标准波段 (红、蓝、绿和近红外), 还含有 4 个额外波段 (黄、红边、海岸和近红外 2); Pléiades 卫星 (<https://www.geoimage.com.au/satellite/Pleiades>[2021-06-18]) 发

射于2011年，能够提供0.5 m全色和2.0 m多光谱影像，多光谱传感器有4个标准波段。其他卫星数据集还有Landsat等。部分卫星图像参数见表1 (Yang等, 2020b); <https://gbdxdocs.digitalglobe.com/docs/worldview-3> [2021-06-18]; <https://gbdxdocs.digitalglobe.com/docs/quickbird> [2021-06-18]; <https://gbdxdocs.digitalglobe.com/docs/ikonos> [2021-06-18])。

com/docs/worldview-3 [2021-06-18]; <https://gbdxdocs.digitalglobe.com/docs/quickbird> [2021-06-18]; <https://gbdxdocs.digitalglobe.com/docs/ikonos> [2021-06-18])。

表1 常用国外卫星基本参数

Table 1 Basic parameters of common foreign satellites

卫星种类	波段参数/nm	分辨率参数/m
WorldView-3	PAN: 450—800 MS: Coastal: 397—454 Red: 626—696 Blue: 445—517 Red Edge: 698—749 Green: 507—586 Near-IR1: 765—899 Yellow: 580—629 Near-IR2: 857—1039	PAN nadir: 0.31, 20° off-nadir: 0.34 MS nadir: 1.24, 20° off-nadir: 1.38
WorldView-2	PAN: 450—800 MS: Coastal: 400—450 Green: 510—580 Blue: 450—510 Yellow: 585—625 Red: 630—690 Red Edge: 705—745 Near-IR1: 770—895 Near-IR2: 860—1040	PAN nadir: 0.46, 20° off-nadir: 0.52 MS nadir: 1.85, 20° off-nadir: 2.07
QuickBird	PAN: 405—1053 MS: Blue: 430—545 Green: 466—620 Red: 590—710 Near-IR: 715—918	PAN nadir: 0.55 MS nadir: 2.16
IKONOS	PAN: 526—929 MS: Blue: 445—516 Green: 506—595 Red: 632—698 Near-IR: 757—853	PAN: 0.82 MS: 3.2
Pléiades	PAN: 470—830 MS: Blue: 430—550 Green: 500—620 Red: 590—710 Near-IR: 740—940	PAN: 0.5 MS: 2.0

国内遥感图像数据集多来自高分卫星家族，如“GF-1”与“GF-2”，但“GF-2”的PAN图像和MS图像分辨率比“GF-1”都提高1倍，分别达到了1 m和4 m的空间分辨率，其技术参数如表2 (<http://www.kosmos-imagemall.com/> [2021-06-18])所示。GF-1技术参数在光谱范围与GF-2一致，全色图像分辨率是2 m，多光谱图像分辨率为8 m。高分卫星家族在环保、农业等领域都发挥了重大的经济价值与社会价值。

表2 GF-2卫星基本参数

Table 2 Basic parameters of GF-2 satellite

参数	1 m分辨率全色/4 m	分辨率多光谱相机
		全色
光谱范围		0.45—0.52 μm
	多光谱	0.52—0.59 μm
		0.63—0.69 μm
		0.77—0.89 μm
空间分辨率	全色	1 m
	多光谱	4 m
覆盖周期(不侧摆)	69 d	

### 3 全色图像锐化质量评价

全色图像锐化质量评价，通常包括主观评价和

客观评价两部分。主观评价是从视觉的角度判断融合图像的清晰度与原始MS图像之间的光谱相似度 (Alparone等, 2015); 客观评价是一种对融合图像进行量化的评价方法，该方法通常分为两类，即有参考和无参考质量评价 (Alparone等, 2015)。下面就常用的客观评价指标给出其定义和说明。

#### 3.1 有参考质量评价指标

(1) 相关系数 CC (Correlation Coefficient) (Alparone等, 2015)。计算公式如下：

$$CC = \frac{1}{M \times N} \times \frac{\sum_{i=1}^M \sum_{j=1}^N (R(i,j) - \bar{R}) \cdot (F(i,j) - \bar{F})}{\sqrt{\left( \sum_{i=1}^M \sum_{j=1}^N (R(i,j) - \bar{R})^2 \right) \cdot \left( \sum_{i=1}^M \sum_{j=1}^N (F(i,j) - \bar{F})^2 \right)}} \quad (1)$$

式中， $M \times N$ 为图像的大小， $R$ 为参考图像， $F$ 为融合图像。 $R(i, j)$ 是 $R$ 中位置坐标 $(i, j)$ 的像素值， $\bar{R}$ 为 $R$ 的平均像素值； $F(i, j)$ 是 $F$ 中位置坐标 $(i, j)$ 的像素值， $\bar{F}$ 为 $F$ 的平均像素值。 $CC$ 用于对融合图像在空间上的扭曲度进行定量评价。 $CC$ 的

值域区间是  $[0, 1]$ ,  $CC$  越趋近于理想值 1, 表明融合图像的质量越趋近于参考图像。

(2) 光谱角映射 SAM (Spectral Angle Mapper) (Alparone 等, 2007)。计算公式如下:

$$SAM = \arccos\left(\frac{\langle u_R, u_F \rangle}{\|u_R\|_2 \cdot \|u_F\|_2}\right) \quad (2)$$

式中,  $u_R$ 、 $u_F$  分别为标准图像和融合图像的谱向量。SAM 用于定量评价标准图像与融合图像光谱信息的相似度。SAM 的值越趋近于理想值 0, 融合图像与参考图像的光谱信息相似度越高, 表明融合图像的光谱失真越小。

(3) 全局相对光谱损失 ERGAS (Erreur Relative Global Adimensionnelle de Synthèse) (Wald, 2000)。计算公式如下:

$$ERGAS = 100 \frac{h}{l} \sqrt{\frac{1}{k} \sum_{k=1}^K \left(\frac{RMSE}{\bar{F}_k}\right)^2} \quad (3)$$

式中,  $h$  表示 PAN 图像的空间分辨率,  $l$  表示 MS 图像的空间分辨率,  $\bar{F}_k$  为各个通道的总体像素平均值。ERGAS 用来对融合图像的光谱与空间质量进行定量评估, ERGAS 的值越趋近于最优值 0, 表明融合图像与参考图像的相似度越高。

(4) 全局融合指标 ( $Q_4$ ) (李红 等, 2016)。对图像进行分块 ( $D \times D$ ), 每块计算  $Q_4$ 。

$$Q_4 = E\left(\|Q_{4(D \times D)}\|\right) \quad (4)$$

$Q_{4(D \times D)}$ ,  $x$  和  $y$  由式 (5) 计算得到:

$$\left\{ \begin{aligned} Q_{4(D \times D)} &= \frac{4(E(x \cdot y^*) - \bar{x} \cdot \bar{y}^*)}{E(\|x\|^2) - \|\bar{x}\|^2 + E(\|y\|^2) - \|\bar{y}\|^2} \\ &\quad \frac{\|\bar{x}\| \cdot \|\bar{y}\|}{\|\bar{x}\|^2 + \|\bar{y}\|^2} \\ x &= X_1(i, j) + \alpha X_2(i, j) + \beta X_3(i, j) + \gamma X_4(i, j) \\ y &= F_1(i, j) + \alpha F_2(i, j) + \beta F_3(i, j) + \gamma F_4(i, j) \end{aligned} \right. \quad (5)$$

式中,  $X(i, j)$  为参考图像,  $F(i, j)$  为融合图像,  $Q_4$  反映了参考图像与融合图像在空间和光谱间的整体相似度, 融合结果越好,  $Q_4$  的值越接近于 1。

### 3.2 无参考质量评价指标

$QNR$  (杨勇 等, 2020) 是遥感图像融合领域常用的无参考图像质量定量评价指标, 它是基于  $UIQI$  的质量评价指标, 用来表征对比图像与融合

图像间的亮度、对比度以及局部相关性。令  $Q(\bullet)$  是计算  $UIQI$  的函数, 则  $QNR$  可用如下公式定义。

$$QNR = (1 - D_s) \cdot (1 - D_\lambda) \quad (6)$$

式 (6) 中的空间失真指数  $D_s$  和光谱失真指数  $D_\lambda$  计算公式如下:

$$D_s = \sqrt[q]{\frac{1}{K} \sum_{t=1}^K |Q(F_t, PN_L) - Q(M_t, PN)|^q} \quad (7)$$

式中,  $PN$  为 PAN 图像,  $PN_L$  为低空间分辨率的  $PN$ ,  $q$  代表范数。 $D_s$  的值越小, 代表空间失真度低 (杨勇 等, 2020), 即图像融合算法的实际效果越理想。

$$D_\lambda = \sqrt[p]{\frac{1}{K(K-1)} \sum_{t=1}^K \sum_{d=1, t \neq d}^K |Q(M_t, M_d) - Q(F_t, F_d)|^p} \quad (8)$$

式中,  $p$  是一个正整数, 用来表征光谱差,  $t$  和  $d$  为通道的数量。 $M_t$  为 MS 图像的第  $t$  个通道,  $F_t$  为融合图像的第  $t$  个通道。 $D_\lambda$  的值越小, 即融合算法的效果越好 (杨勇 等, 2020)。

$D_s$  和  $D_\lambda$  的最优值均为 0,  $D_s$  和  $D_\lambda$  值越小, 则  $QNR$  值越大, 融合图像的质量越高,  $QNR$  的理想值为 1。

## 4 基于深度学习的全色图像锐化

深度学习 DL (Deep Learning) 的概念由 Hinton 和 Salakhutdinov (2006) 以及 Hinton 等 (2006) 提出, 属于机器学习的范畴, 其目标是构建模拟人类大脑自动分析并能够理解数据的神经网络, 并于 2018 年获得图灵奖。

本节与像素级全色图像锐化相关的 5 种重要的 DL 模型, 分别是 AE、CNN、深度残差网络 DRN (Deep Residual Network)、生成对抗网络 GAN (Generative Adversarial Nets) 和循环神经网络 RNN (Recurrent Neural Network)。

深度学习因提取特征能力强, 识别精度高等优点被广泛应用在计算机视觉领域, 许多研究人员将深度学习应用到像素级全色图像锐化中, 在最近 6 年中取得了巨大的进步。通过对现有诸多文献的分析挖掘, 本文给出了基于深度学习的全色图像锐化方法的发展历程及其分类, 如图 1 所示。下面将根据深度学习模型与传统融合算法, 从经典方式与协同方式两个方面来分门别类地综述这些研究成果。

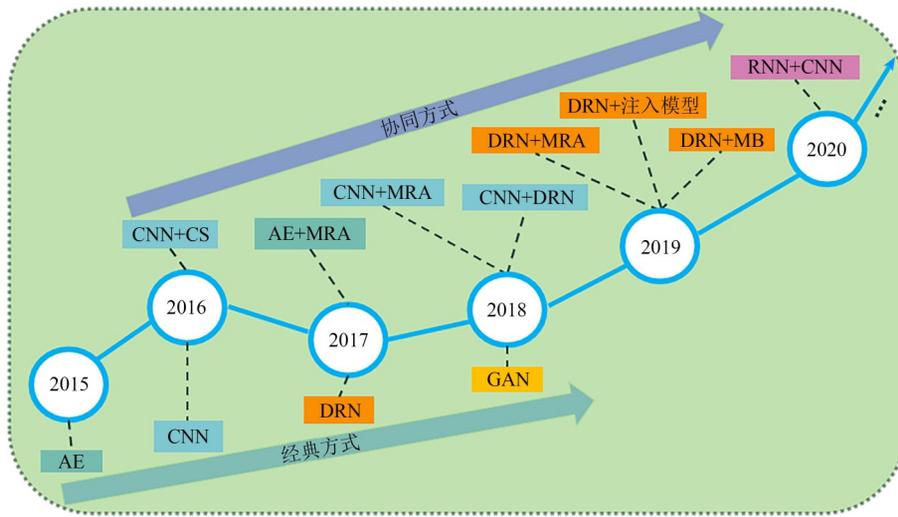


图1 基于深度学习的像素级全色图像锐化发展历程及分类

Fig. 1 Deep learning based pixel-level pansharpening development and classification

### 4.1 经典方式

经典方式主要是直接把深度学习的一种模型单独应用于全色图像锐化。当前, 常见的是基于AE、CNN、DRN和GAN等4种深度学习模型的全色图像锐化方法的研究。

#### 4.1.1 基于AE的全色图像锐化

AE是由输入层、隐含层(编码层)和输出层构成的3层深度神经网络, 其中, 输入层X和输出层Z具有同样大小。输入层X到隐含层Y是一个编码过程; 隐含层Y到输出层Z是一个解码过程。AE通过BP反向传播算法训练, 不断调整编码器和解码器的参数(权重), 重构损失最小化, 使输入层X等于输出层Z, 最后得到编码层Y的结果(低维向量的形式)即是对输入层X的特征表达。

Huang等(2015)首次把DNN(Deep Neural Network)应用于遥感图像融合, 假定HR/LR PAN图像之间的映射关系与HR/LR MS图像的映射关系是一致的, 提出了一个主体包含MSDA(Modified Sparse Denoising Autoencoder)的深度神经网络融合模型。MSDA能够学习获得LR PAN图像与HR PAN图像之间的映射关系, 然后把训练后的MSDA通过叠加构造一个DNN, 最后通过DNN完成从LR MS图像到HR MS图像的重构, 从而达到融合目的。

为训练网络中的参数, Huang等(2015)定义了以下损失函数:

$$L_2\left(\{x_p^i, y_p^i\}_{i=1}^N; \Theta\right) = \frac{1}{N} \sum_{i=1}^N \|x_p^i - \hat{x}(y_p^i)\|_2^2 + \frac{\lambda}{2} \sum_{l=1}^L \|W_l\|_F^2 \quad (9)$$

式中,  $x_p^i$ 表示PAN图像块,  $y_p^i$ 表示MS图像块,  $N$ 表示训练样本数,  $L$ 表示MSDA的数目,  $W_l$ 表示网络中第 $l$ 层的参数。

Huang等(2015)提出的方法与数个传统融合算法进行了比较, 在QuickBird和IKONOS两个卫星数据集上, 前者能得到更优的主客观融合图像, 验证了基于DL的全色图像锐化的有效性。

Cai等(2018)提出了一个CSDA(Coupled Sparse Denoising Autoencoder)的融合框架。CSDA由4个子网络组成, 分别是LM-HP网络、HR-MS网络、特征映射网络FMN(Feature Mapping Network)和微调网络(Fine-Tuning Network)。其中LM-HP网络和HR-MS网络是SAE(Sparse AutoEncoder)网络, 分别用来学习MS、PAN图像和HRMS图像的特征表达信息; 特征映射网络SDAE(Sparse Denoising Autoencoder)用来对由SAE提取的两个特征建立映射关系; 微调网络通过BP算法对CSDA的参数进行调整。

LM-HP网络的损失函数定义如下:

$$l_1 = \sum \|LR_i - \widehat{LR}_i\|^2 + KL(\hat{\rho}_1 \| \rho_1) \quad (10)$$

HR-MS网络的损失函数定义如下:

$$l_2 = \sum \|Pan_i - \widehat{Pan}_i\|^2 + KL(\hat{\rho}_2 \| \rho_2) \quad (11)$$

FMN的损失函数定义如下:

$$l_3 = \sum_i \|h(Pan_i) - h(\widehat{Pan}_i)\|^2 + KL(\hat{\rho}_3 \| \rho_3) \quad (12)$$

优化网络的损失函数定义如下:

$$l = \sum_i \| \text{Pan}_i - \text{HR}_i \|^2 \quad (13)$$

式(10) — (13)中,  $\text{LR}_i$ 表示LM-HP图像向量,  $\text{Pan}_i$ 表示HR-MS图像向量真实值,  $\text{HR}_i$ 表示CSDA网络的输出,  $\text{KL}(\hat{\rho} \parallel \rho)$ 表示稀疏正则项。

该融合方法工作过程是:首先,对上采样的MS图像和PAN图像向量化,然后,利用LM-HP网络、HR-MS网络分别提取MS、PAN图像的特征,接着,利用SDAE网络建立MS、PAN图像特征的映射关系,得到HRMS图像特征,最后,重构获得HRMS融合图像。

作为早期的基于深度学习的像素级遥感图像融合模型,虽然较传统融合算法在效果上有较大提升,但由于AE结构的限制、不能直接处理图像以及较少考虑到遥感图像融合领域的先验知识等,相较其他基于深度学习的遥感图像融合模型存在一些不足。另外,该类方法提升融合效果往往使得网络模型结构、损失函数等较为繁杂,不具简明性。

#### 4.1.2 基于CNN的全色图像锐化

CNN是一种多层感知机,主要用来处理二维图像数据(张洪群等,2017)。相比传统的多层感知机,CNN的特点在于:网络的输入直接采用原始信号,从而使得传统识别算法中复杂的特征提取得以避免;卷积阶段通过权值共享结构减少了网络参数量,使得网络的训练效率显著提高;池化阶段通过对特征图下采样操作,在主要特征信息得以保存的同时,使得特征空间的大小明显减小。

受Dong等(2016)的启发,Masi等(2016和2017)把SRCNN(Super-Resolution CNN)中的CNN应用到遥感图像融合领域,提出了首个基于CNN的全色图像锐化模型PNN(CNN-based Pansharpening),其架构如图2所示。PNN模型以PAN和上采样后的MS图像为输入,然后经过3层CNN,最后得到HRMS融合图像。PNN模型在WorldView-2、IKONOS和GeoEye-1这3个卫星数据集上与十几个传统融合算法做了大量的对比实验,验证了PNN模型应用于遥感图像融合的有效性。

He等(2018)通过实验发现:当训练网络的图像太小时,网络对图像边缘、纹理等细节信息

的学习能力会变弱,导致融合图像空间细节信息失真;当训练网络的图像太大时,容易导致网络陷入过拟合。基于此,He等(2018)设计了一个三通道的CNN框架,其中每个通道对应不同分辨率大小的训练MS图像和PAN图像,最后经过统一池化操作得到融合图像;Ye等(2018)设计了一个多尺度的CNN模型,以上采样的MS图像和增强PAN图像作为训练后融合网络的输入,最后得到融合图像,其中增强PAN图像由低频的MS图像注入到经NSLP(Non-Subsampled Laplacian Pyramid)变换的PAN图像而得到;李红等(2016)在CNN的基础上,提出了一个DSVL Nets(Deep Support Value Learning Networks)的融合框架,DSVL Nets在传统CNN的基础上,移除特征层的下采样操作,提升了图像融合的效果。

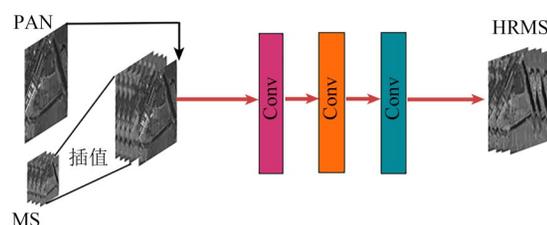


图2 PNN模型结构

Fig.2 The network structure of PNN

Masi等(2016)提出的PNN模型具有开创性的意义,后续许多学者沿着此路径开展研究,无论是从网络类别、网络宽度、网络深度、目标函数等方面都受PNN模型的影响。然而,由于PNN模型网络深度较浅,只是3层的CNN,其信息表征(特征提取)能力受到制约。同时,该网络模型是将PAN图像和上采样的MS图像直接作为PNN模型的输入,没有充分考虑到遥感图像融合领域的特殊性,导致部分空间、光谱信息的丢失。后续的研究(He等,2018)在CNN的网络宽度上均有增加,实验结果表明,在主观视觉和客观指标上均达到比PNN模型更高质量的融合效果。

#### 4.1.3 基于DRN的全色图像锐化

在深度学习中,图像特征随着网络深度的加深而变得更为抽象,增加网络深度可以有效地提取更多高层的特征(Zeiler和Fergus,2014)。但是随着网络模型深度的逐步增加,参数更新过程中可能会出现“梯度消失”,使得网络模型难以优化

(He 和 Sun, 2015)。为了解决该问题, He 等 (2016) 创新性地把“Shortcut Connections”概念 (Ripley, 1996) 引入到 CNN 中, 设计了深度残差网络模型 ResNet, 也即 DRN。

通过对 PNN 等模型不足之处的分析以及结合遥感图像融合领域的先验知识, Yang 等 (2017) 以遥感图像融合的两大主要目标: 光谱信息、空间信息的理想保留, 创新性地提出了一个主体包含 ResNet 结构的 PanNet (Deep Network for Pan-Sharpening) 模型, 其结构框架如图 3 所示。该模型为显著地保留 MS 图像和 PAN 图像中的光谱和空间信息, 引入了两个损失函数。其中, 针对空间信息保持, 其定义如下损失函数:

$$L = \left\| f_w(G(P), \uparrow G(M)) + \uparrow M - X \right\|_F^2 \quad (14)$$

式中,  $f_w$  代表图 5 中的 ResNet,  $G$  表示高通滤波操作,  $\uparrow$  是上采样操作,  $P$ 、 $M$ 、 $X$  分别表示 PAN 图像、MS 图像和 HRMS 融合图像。

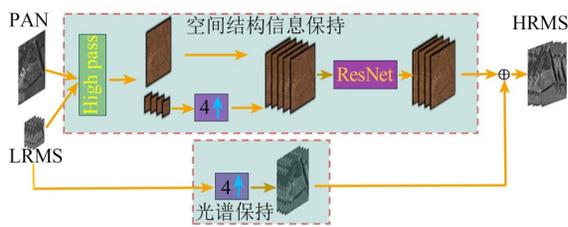


图 3 PanNet 模型结构

Fig.3 The network structure of PanNet

PanNet 的另一显著特点在于, 在输入到训练网络前, 对 MS 图像和 PAN 图像进行高通滤波操作, 使得 PanNet 模型的泛化性能得到提升, 在 WorldView-2 和 WorldView-3 卫星数据集上表现出显著的融合效果。

Wei 和 Yuan (2017)、Wei 等 (2017) 受 Kim 等 (2016) 使用深度残差网络的启发, 提出 DRPNN (Deep Residual Pan-sharpening Neural Network) 融合模型, 取得了比 PNN 模型以及数个经典融合算法更高质量的融合图像; 受传统 CS、MRA 方法以及 ResNet 模型的启发, 为学到 HRMS 图像与 MS 图像之间的空间“残差”信息, Rao 等 (2017) 设计了一个具有残差学习功能的 CNN 以完成高质量的图像融合; 受 Kim 等 (2016) 的启发, 并假定 LRPAN、HRPAN 图像与 MS、HRMS 图像具有相同的映射关系, Li 等 (2017) 训练了一个 DRN 模型, 该模型通过训练获得 LR PAN 图像与

HR PAN 图像之间的映射函数, 然后把该训练模型用于完成从 MS 图像到 HRMS 图像的重构, 从而达到融合的目的; Jiang 等 (2019) 设计了一个 DRN 框架, 该模型能够将 HRPAN 图像与 MS 图像之间的差异信息映射到 HRMS 图像与 HRPAN 图像之间的差异信息, 从而达到融合的目的。

Yang 等 (2020a) 做了进一步的研究, 提出了 PCDRN (Progressive Cascade Deep Residual Network) 融合框架, 其以串联的形式包含两个 ResNet 子网络, 分两次学习 LR MS 图像和 PAN 图像到 HR MS 图像的非线性的特征映射, 使得融合图像可以保留 PAN 图像中更丰富的空间信息。另外, 为解决融合过程中过拟合的现象, Yang 等 (2020a) 还设计了一个新的多任务损失函数:

$$L^{\text{Fusion}} = \alpha \cdot L_{\text{NMSE}}^{\text{Fusion}} + (1 - \alpha) \cdot L_{\text{UIQI}}^{\text{Fusion}} \quad (15)$$

式中,  $L_{\text{NMSE}}^{\text{Fusion}}$  表示归一化的 MSE 损失,  $L_{\text{UIQI}}^{\text{Fusion}}$  表示归一化的 UIQI 损失,  $\alpha$  表示权重系数。

在遥感图像不同的区域上, 光谱、空间分辨率存在差异, 利用显著性分析对光谱、空间信息进行划分具有合理性。受此启发, Zhang 等 (2019a) 将注意力机制中的显著性分析原理与 DRN 相结合, 提出基于显著性引导的 DRN 融合模型, 同时, 通过设计了一个新的损失函数, 使得融合图像获得更丰富的细节信息, 进一步提升了融合图像的质量; Yao 等 (2018) 设计了一个主体结构是 AE 的 PUNet (U-Net for Pansharpening) 融合框架, 在 IKONOS 与 WorldView-2 两个卫星数据集上的实验表明, PUNet 获得了比 GIHS (Tu 等, 2001)、AWLP (Otazu 等, 2005) 等传统算法更好的融合效果。

Scarpa 等 (2018) 从损失函数、网络深度、是否加入残差学习模块等 3 个方面对遥感图像融合 PNN 模型进行全面的实验分析、改进, 提出了一个目标自适应的带有残差学习模块的 3 层 CNN 融合框架, 在 WorldView-2、WorldView-3、IKONOS 和 GeoEye-1 这 4 个卫星数据集上得到高质量的融合图像; 随后, Scarpa 等 (2018) 对该模型框架又开展了进一步的研究 (Vitale 等, 2018; Vitale, 2019); Vinothini 和 Bama (2019) 使用 RDN (Residual Dense Network) 模型 (Zhang 等, 2018) 设计了一个 RDMPsnet (Residual Dense Model for Pan-Sharpening Network) 的融合框架, 在 WorldView-2 和 IKONOS-2 两个卫星数据集上的实验表明, RDMPsnet 获得了比

PNN、PanNet 等模型更好的融合效果；Xiang 等 (2019) 设计了一个多尺度密集深度学习的融合框架，框架的特色在于包含一个多尺度的卷积层、一个密集深度学习模块 (Dense Learning Block) 和一个全局的特征连接结构。其中多尺度卷积层用来提取 PAN 图像的多尺度特征；密集深度学习模块用来提高网络对图像的光谱、空间表达能力；全局的特征连接结构则保留了上采样后的 MS 图像的光谱结构信息。

Guo 等 (2019) 设计了一个主体包含 DMB (Dilated Multilevel Block) 的融合框架。Li 和 Cheng (2019) 在 Landsat 8 卫星数据集上设计了一个由特征提取、非线性映射关系构造、融合图像重构 3 部分构成的 DRN 融合框架；以改进 PNN 模型为出发点，提升对图像各尺度的信息表达能力，Li 等 (2020b) 设计了一个基于多尺度通道注意力机制的 DRN 融合框架，该框架的一个核心模块是 MSRBCA (Multi-scale Residual Block With Channel Attention)，其主要作用是实现对源图像多尺度的特征提取。

基于 DRN 的全色图像锐化方法充分利用其网络中的快捷/跳转连接 (Shortcut Connections)，使得用于学习图像特征的 DRN 兼顾到较为浅层的图像特征的同时，网络更深，学习效率更高。在实验效果上，无论从光谱信息、空间细节的保留上均达到比传统 CNN 模型更好的主客观融合效果；另外，Zhang 等 (2019a) 和 Li 等 (2020b) 把注意力机制加入到 DRN 模型，使得基于 DRN 的全色图像锐化方法更为丰富和更具灵活性。

#### 4.1.4 基于 GAN 的全色图像锐化

Goodfellow 等 (2014) 提出生成对抗网络 GAN，其重要特征在于其包含两个网络模型：生成器 G (Generator) 和判别器 D (Discriminator)。生成器 G 尽可能使得生成的图像更为“真实”，导致判别器 D 难辨真伪；判别器 D 判定输入样例是来自数据集还是生成器 G 生成的图像。两个网络模型相互对抗并不断提升各自的算法能力，直到判别器无法分辨数据集图像与生成图像。

Liu 等 (2018b) 首次将 GAN 用于遥感图像融合，提出了 PSGAN 模型，其中生成器用来得到 HRMS 融合图像，其损失函数定义为

$$L(G) = \sum_{n=1}^N [-\alpha \log D_{\theta_D}(X, G_{\theta_G}(X, Y)) + \beta \|P - G_{\theta_G}(X, Y)\|_1] \quad (16)$$

判别器用来判断生成器生成的图像是真实的还是通过融合得到的图像，损失函数定义为：

$$L(D) = \sum_{n=1}^N [1 - \log D_{\theta_D}(X, G_{\theta_G}(X, Y)) + \log D_{\theta_D}(X, P)] \quad (17)$$

式 (16) 和 (17) 中， $X$ 、 $Y$  分别表示 MS 图像和 PAN 图像， $P$  表示 HRMS 融合图像， $N$  为训练图像样本数， $\alpha$ 、 $\beta$  为 1—100 范围的参数，G、D 分别是生成器和判别器。

PSGAN 模型在 Quickbird 和 GaoFen-1 两个卫星数据集上获得比 PNN 等数个融合算法更高质量的融合图像。

然而，PSGAN 在生成器的训练效率以及细节信息保持上仍有很大的提升空间。Shao 等 (2020) 把 GAN 与具有残差学习能力的自编码器思想相结合，设计了 RED-cGAN (Residual Encoder-Decoder Conditional Generative Adversarial Network) 的融合框架。RED-cGAN 中的生成器采用了双分支子网络结构，分别用它们提取 PAN 图像和 MS 图像的特征信息，接着通过 RED (Residual Encoder-Decoder Module) 结构生成 HRMS 融合图像，与 PSGAN 模型的不同在于：RED-cGAN 的判别器使用 PAN 图像与生成器产生的图像代替上采样的 MS 图像与生成器产生的图像作为其输入，以保留更多的空间细节信息。在 WorldView-2 和 WorldView-3 两个卫星数据集上 RED-cGAN 得到比 DRPNN、PanNet 和 PSGAN 等模型更高质量的融合图像。Ozcelik 等 (2020) 设计了一个 PanColorGAN 的融合框架，在 Pléiades 和 WorldView-2/3 卫星数据集上，PanColorGAN 获得了比 PanNet 等多个融合算法信息更丰富的融合图像；Ma 等 (2020) 设计了一个基于 GAN 模型的无监督学习融合框架 Pan-GAN，其包含两个判别器，分别是光谱判别器和空间判别器，以更充分地保留源图像的光谱和空间信息。在 WorldView-2 和 GaoFen-2 两个卫星数据集上 Pan-GAN 得到了比 PNN、PanNet 和 PSGAN 等模型更高质量的融合图像；另外，Zhang 等 (2019c)、Chen 等 (2020) 以及 Zhang 等 (2020) 也将 GAN 应用于遥感图像融合的研究。

基于 GAN 的全色图像锐化方法与其他模型

(如基于CNN、DRN的全色图像锐化)最大的不同在于GAN由两个网络构成,分别是生成器与判别器,目前生成器所采用的网络模型多是自编码器。GAN结构的优点在于可以并行地生成图像,与自回归模型相比,GAN生成图像的速度比较快,而且生成的图像比较清晰(陈佛计等,2021);另外,GAN结构中的判别器可以进一步提升生成器所生成的图像的质量。由于GAN结构具有以上优点,因此基于GAN的全色图像锐化成为当前的研究热点。

#### 4.1.5 对比小结

上述4节介绍了经典方式下,AE、CNN、DRN和GAN等4种常见的深度学习模型单独应用于遥感图像融合的研究,本节对这4类方法各自的特点和异同点进行小结,并在表3中将4类方法中的典型算法运用有参考和无参考图像质量评价指标进行了性能提升比的对比、分析和归纳(选择性能提升比作为定量分析的主要原因在于各融合算法采用的数据集存在差异)。综合考虑到算法对比的广泛性和代表性,选取BDS(DGarzelli等,2008)、GS(Laben和Brower,2000)和PNN(Masi等,2016)这3种算法作为比较的基准算法;考

虑到评价的全面性,选取有参考评价指标SAM(Alparone等,2007)和无参考评价指标QNR(杨勇等,2020)作为比较指标,得到以下结论:

(1) 相比于PNN模型,融合效果提升的方法中,大多采用了DRN网络结构或者网络宽度更宽的CNN网络,而极少使用AE网络;

(2) 残差模块以及GAN的使用,可以有效提升融合效果,如Liu等(2018b)提出的方法,GAN也表现比较强的潜力;

(3) 即使采用简单的浅层CNN,如果在损失函数、残差模块等方面优化融合模型,仍可得到高质量的融合图像,如Scarpa等(2018)提出的方法;

(4) 损失函数中加入评价指标因子(如UIQI等),可增加融合网络模型的可解释性,如Yang等(2020a)提出的方法。采用GAN的融合算法,损失函数中包含生成器损失 $L(G)$ 和判别损失 $L(D)$ ,相比于其他模型的损失函数较为复杂,如Liu等(2018b)和Ozcelik等(2020)提出的方法;

(5) 各算法在SAM指标上提升的幅度大于(或等于)在QNR指标上的提升幅度,少数算法在某些卫星数据集上QNR指标没有得到提升,如Ozcelik等(2020)提出的方法。

表3 经典方式算法比较(基准算法分为传统方法BDS、GS,基于DL的PNN,选择SAM和QNR为比较指标,↑表示相对当前指标的提升比,—表示空值)

Table 3 Comparison of classical algorithms (the benchmark algorithm is divided into traditional methods BDS, GS, and DL-based PNN, SAM and QNR are selected as the comparison indicators, ↑ represents the improvement ratio relative to the current indicator, — represents the null value)

算法	体系结构	残差模块	损失函数	卫星数据集	SAM			QNR		
					BDS ↑	GS ↑	PNN ↑	BDS ↑	GS ↑	PNN ↑
S-MSDA (Huang等,2015)	AE	无	$L_2(\{x_p^i, y_b^i\}_{i=1}^N; \Theta) = \frac{1}{N} \sum_{i=1}^N \ x_p^i - \hat{x}(y_b^i)\ _2^2 + \frac{\lambda}{2} \sum_{l=1}^L \ W_l\ _F^2$	IKONOS QuickBird	—	34%	—	—	—	—
CSDA (Cai等,2018)	AE	无	$l_1 = \sum \ LR_i - \widehat{LR}_i\ ^2 + KL(\hat{\rho}_1 \  \rho_1)$ $l_2 = \sum \ Pan_i - \widehat{Pan}_i\ ^2 + KL(\hat{\rho}_2 \  \rho_2)$ $l_3 = \sum \ h(Pan_i) - h(\widehat{Pan}_i)\ ^2 + KL(\hat{\rho}_3 \  \rho_3)$ $l = \sum \ Pan_i - HR_i\ ^2$	QuickBird	—	47%	—	—	—	—
PNN (Masi等,2016)	三层CNN	无	$L(W) = \frac{1}{128} \sum_{n=1}^{128} \ X_n - \widehat{X}_n(W)\ ^2$	WorldView-2 IKONOS	31%	—	—	2%	—	—
PanNet (Yang等,2017)	DRN	有	$L = \ f_w(G(P), \uparrow G(M)) + \uparrow M - X\ _F^2$	WorldView-3	42%	—	13%	13%	—	3%
DRPNN(Wei和Yuan,2017)	DRN	有	$LOSS = \ f_{MS} - F\ _2^2$	QuickBird WorldView-2	—	45%	17%	—	—	—

续表

算法	体系架构	残差模块	损失函数	卫星数据集	SAM			QNR		
					BDS ↑	GS ↑	PNN ↑	BDS ↑	GS ↑	PNN ↑
PCDRN (Yang等,2020a)	串联 双DRN	有	$L^{\text{Fusion}} = \alpha \cdot L_{\text{NMSE}}^{\text{Fusion}} + (1 - \alpha) \cdot L_{\text{NUIQI}}^{\text{Fusion}}$	Pléiades WorldView-3	— —	33% 19%	— —	21% 10%	— —	
L1-RL-FT (Scarpa等,2018)	DRN	有	$L(\bar{x}, \Phi) \propto \ f(x, \Phi) - \Delta x_{\text{ref}}\ _1$	WorldView-2 WorldView-3	42% 31%	— —	14% 38%	11% 4%	2% 5%	
(Li等,2020b)	DRN	有	$L(\Theta) = \frac{1}{n_i} \sum_{i=1}^{n_i} \ F(I_{\text{PAN}_i}, I_{\text{MS}_i}; \Theta) - I_{\text{REF}_i}\ _2^2$	GF-1 GF-2	— —	56% 53%	11% 2%	— —	26% 22%	0.4% 2%
PSGAN (Liu等,2018b)	GAN	有	$L(G) = \sum_{n=1}^N \left( -\alpha \log D_{\theta D}(X, G_{\theta G}(X, Y)) + \beta \ P - G_{\theta G}(X, Y)\ _1 \right)$ $L(D) = \sum_{n=1}^N \left( 1 - \log D_{\theta D}(X, G_{\theta G}(X, Y)) + \log D_{\theta D}(X, P) \right)$	QuickBird GF-1	— —	— —	19% 33%	— —	— —	
PanColorGAN (Ozcelik等,2020)	GAN	有	$L_D = L_{\text{RaGAN}}(Y_{\text{MS}}, \hat{Y}_G)$ $L_G = L_{\text{Rec}} + \alpha L_{\text{RaGAN}}(\hat{Y}_G, Y_{\text{MS}})$ $L_{\text{Rec}} = \ Y_{\text{MS}} - \hat{Y}_G\ _1$	Pléiades WorldView-2/3	49% 7%	— —	— —	-1% 0.5%	— —	

## 4.2 协同方式

在深度学习技术应用于全色图像锐化之前, 基于传统方法的全色图像锐化已经开展了数十年的研究, 在这期间, 产生了大量的融合算法, 不乏有些算法已经应用于专业遥感软件系统, 因而出现诸多把传统方法与DL技术结合开展全色图像锐化的研究 (Zhong等, 2016; Azarang和Ghassemian, 2017; He等, 2019)。由于CNN、DRN在全色图像锐化的显著效果, 也出现了CNN与DRN结合的全色图像锐化等方法。基于此, 本文将传统方法+DL技术的全色图像锐化方法以及不同DL模型相结合的全色图像锐化方法概括为协同方式。

### 4.2.1 基于DL+CS的全色图像锐化

为了在全色图像锐化中保留MS图像中的空间信息, Zhong等(2016)把DL技术与传统GS(Gram-Schmidt)方法结合构建一个新的融合框架。其中DL技术采用的是SRCNN模型, 用来增强MS图像, 然后将PAN图像与增强MS图像作为GS融合框架的输入, 最后得到HRMS图像; Azarang等(2019)使用训练后的CAE(Convolutional AutoEncoder)网络以提升LRMS图像的空间信息, 然后把CAE纳入CS的融合框架, 最后完成遥感图像融合。

上述研究可看作DL与CS方法结合完成遥感图像的融合。其共同点是DL技术用来增强MS图像

的空间信息, 然后再按照传统CS融合方法得到融合图像; 其不同点在于Zhong等(2016)采用SRCNN模型增强MS图像信息, 而Azarang等(2019)采用的是CAE网络。

### 4.2.2 基于DL+MRA的全色图像锐化

Azarang和Ghassemian(2017)设计了一个MRA和DNN(Deep Neural Networks)结合的融合框架, 其中DNN由基本的AE构成。Eghbalian和Ghassemian(2018)设计了一个包含两个DNN的融合框架, 并分两大阶段对遥感图像进行融合, 第一阶段完成图像融合, 第二阶段进行光谱补偿。其中第一阶段的深度卷积神经网络DCNN(Deep Convolution Neural Network)与MRA融合方法相结合完成图像融合, DCNN的作用是用来提取MS图像需要注入的空间信息; 第二阶段的光谱补偿网络SCN(Spectral Compensation Network)实质上是在DCNN网络的基础上, 增加了新的损失函数, SCN网络的目的是减少融合图像的光谱失真, 同时尽可能地保存空间信息, 在WorldView-2、GeoEye-1和Pléiades这3个卫星数据集上的实验结果表明该方法有效地提高了遥感图像融合效果。Huang等(2018)设计了一个主体由NSCT(Nonsubsampled Contourlet Transform)与DNN构成的融合框架。该方法通过训练和融合两个阶段完

成融合任务, 在训练阶段, 使用 NSCT 分别对 HR/LR PAN 图像在不同分辨率和方向上进行分解, 得到高、低频子带图像并作为 DNN 的输入来训练网络, 以提取更丰富的空间结构信息; 在融合阶段, 由 A-PCA (Adaptive PCA) 与 NSCT 联合把 MS 图像分解为高频子带图像和低频子带图像, 使用其高频子带图像作为训练后 DNN 的输入, 再把输出与 MS 图像的低频子带图像联合, 经过 NSCT 逆变换等操作得到 HRMS 融合图像。

Zhang 等 (2019b) 设计了 BDPN (Bidirectional Pyramid Network) 的融合框架, BDPN 包含两个分支子结构, 分别是重构分支和细节提取分支。细节提取分支结构是一个 DRN 网络, 用来提取 PAN 图像的空间细节信息, 再注入到上采样的 MS 图像中, 然后利用重构分支得到 HRMS 融合图像。Benzenati 等 (2020) 设计了一个基于 CNN 的广义拉普拉斯金字塔融合框架, 其中 CNN 用来提高 GLP (Generalized Laplacian Pyramid) 的空间细节注入信息。

通常情况下, CS 融合方法空间信息保持较好, 光谱失真大; MRA 融合方法光谱细节保持较好,

空间细节失真大。为客观评价基于 DL+MRA 的融合方法在空间信息、光谱细节保持上的表现, 表 4 给出了算法性能提升比分析, 与之比较的方法有两类, 包括 CS 方法: IHS (Carper 等, 1990)、文献 (Shah 等, 2008) 和 GFP (Liu 和 Liang, 2016); MRA 方法: Indusion (Khan 等, 2008)、文献 (Da 等, 2006) 和 Seg\_GLP (Aiazzi 等, 2007)。得到如下结论:

(1) 在空间信息保持上, 基于 DL+MRA 的融合方法优于单独使用 CS 的融合方法;

(2) 基于 DL+MRA 的融合方法, 在光谱细节保持上, 能获得比单独 MRA 方法更好的结果, 均达到超过 5% 的提升;

(3) 基于 DL+MRA 的融合方法, 在 QNR 指标上, 同样获得了比单独 MRA 方法更好的结果。

基于以上 3 点, 得出基于 DL+MRA 的融合方法, 起关键作用的是 DL 技术的运用, 其不仅能够增强该方法在光谱信息保持方面的能力, 而且能克服 MRA 方法在空间信息保持方面的不足, 甚至超过 CS 方法在空间信息保持方面的优势。

表 4 WorldView-3 数据集上 DL+MRA 算法比较 (选择 CC、SAM 和 QNR 为比较指标)

Table 4 Comparison of DL+MRA algorithms on WorldView-3 dataset (CC, SAM and QNR were selected as comparison indexes)

算法	体系架构	CS 方法			MRA 方法		
		对比算法	CC 指标 ↑	QNR 指标 ↑	对比算法	SAM 指标 ↑	QNR 指标 ↑
Azarang 和 Ghassemian(2017)	AE+MRA	IHS	2%	—	Indusion	43%	—
Huang 等(2018)	DNN+MRA	Shah (2008)	2%	—	Da 等(2006)	6%	—
BDPN(Zhang 等, 2019b)	DRN+MRA	GFP	4%	13%	Seg_GLP	8%	1%

### 4.2.3 基于 DL+MB 的全色图像锐化

Xing 等 (2018a) 提出了一个 DML-GMME (Deep Metric Learning based Geometric Multi-Manifold Embedding) 融合框架。DML-GMME 分两个阶段完成图像融合, 分别是训练阶段和重构阶段。训练阶段的网络模块是由数个 SSAE (Stacked Sparse AutoEncoder) 构成, 用来提取由聚类分组的 PAN 图像块和下采样 PAN 图像块的流形 (Manifold), 即分别是 HR 流形和 LR 流形; 在重构阶段, 假定使用 LR 流形重构 MS 图像块的系数与使用 HR 流形重构 HRMS 图像块的系数相同, 最后通过 HR 流形, 以及由 MS 图像块与 LR 流形计算出的系数共同重构出 HRMS 图像。

Ye 等 (2019) 设计了一个 DL 与优化模型相结

合的融合框架。该融合框架由预处理、DRN 和重构这 3 个部分构成, 其核心模块是 DRN, 主要用来提取 PAN 图像的空间信息。不同于大多数基于 CNN 的融合方法在像素域训练网络, DRN 是在图像梯度域训练网络; 重构模块是 ADMM (Alternating Direction Method of Multipliers), ADMM 把 DRN 输出的梯度信息和 MS 图像的上采样信息作为输入, 最后重构出融合结果 HRMS 图像。由于该融合框架的核心模块 DRN 是在图像梯度域训练网络, 使得该方法泛化性能得到提高, DRN 在 WorldView-2 和 QuickBird 两个卫星数据集上训练, 能够使得融合框架在 WorldView-3 和 IKONOS 卫星数据集上直接得到融合图像, 较 PNN 等得到更高质量的图像。

#### 4.2.4 基于DL+注入模型的全色图像锐化

He等(2019)设计了两个基于CNN的细节注入框架,分别是DiCNN1(Detail injection based CNN)和DiCNN2。DiCNN1通过CNN从MS图像和PAN图像中提取细节,然后注入到MS图像,得到HRMS图像;DiCNN2通过CNN从PAN图像中提取细节,然后注入到MS图像,从而得到HRMS图像。DiCNN1和DiCNN2共同之处是DL技术与细节注入模型结合完成图像融合,不同之处主要是网络结构、损失函数和应用目标的差异。DiCNN1和DiCNN2在World View-2, IKONOS和Quickbird这3个卫星数据集上较PNN与DRPNN等模型得到信息更丰富的融合图像;He等(2020)设计了SA-PNN的融合框架,在上述3个常见卫星数据集上获得了比PNN与DRPNN等模型更高质量的融合结果。

Deng等(2020)把传统融合方法与深度学习方法相结合,设计了Fusion-Net的融合框架。Liu等(2020)提出了基于浅-深卷积网络与细节注入模型结合的融合框架MIPSM(Multispectral Image Pan-Sharpening Method),包含SDCN(Shallow-Deep Convolutional Network)和SDDI(Spectral Discrimination-based Detail Injection)两个子模型。其中SDCN中的浅层CNN分支网络用来提取PAN图像中层的空间特征,另一个深层DRN分支网络用来提取PAN图像高层的空间特征,然后将提取的空间细节注入到上采样的MS图像得到融合图像。

特别地,MIPSM定义如下的损失函数:

$$L = \sum_{j=1}^S \|G_{MS}^j + k^j g_D^j - f_{MS}^j\|_2^2 \quad (18)$$

式中, $G_{MS}^j$ 表示上采样MS图像第j个通道的光谱信息, $g_D$ 表示SDCN输出的空间信息, $f_{MS}$ 表示“真实图像”, $S$ 表示上采样MS图像的通道数。

整体上,MIPSM在IKONOS和QuickBird两个卫星图像数据集上的融合实验达到比PNN等模型更好的结果。但是,MIPSM在某些指标上(如SAM)泛化性能表现不足。

从结构上分析,He等(2019)提出的DiCNN融合模型与Liu等(2020)提出的MIPSM融合模型共同点在于都是基于DL+注入模型的全色图像锐化方法,不同点在于MIPSM融合模型利用高通滤波提取PAN图像高频信息,有效提高了该方法的

泛化能力。

#### 4.2.5 基于CNN+DRN的全色图像锐化

受Masi等(2016)工作的启发,为提升基于DL的全色图像锐化的融合质量,Yuan等(2018)设计了一个包含PNN和多尺度深度神经网络的两个子网络结构的融合框架MSDCNN(Multiscale and Multidepth Convolutional Neural Network),其中PNN用来提取简单特征;多尺度深度神经网络用来提取多尺度的特征,以增强对高分辨率图像中丰富的空间信息的提取、提高各种复杂地面场景中特征提取的鲁棒性。另外,为了让多尺度深度神经网络深度更深以及提升训练效率,该子网络也加入了跳转连接,MSDCNN在QuickBird和WorldView-2两个卫星数据集上获得比PNN等更好的融合效果,Shao和Cai(2018)设计了一个包含双分支CNN的融合框架RSIFNN(CNN-based remote sensing image fusion)。两个分支CNN分别提取源图像的特征,最后经过进一步处理来得到融合图像。特别地,Shao和Cai(2018)定义了以下的损失函数:

$$L = \frac{1}{n} \sum_{i=1}^n \|r^{(i)} - g(x_1^{(i)}, x_2^{(i)})\|^2 \quad (19)$$

式中, $r^{(i)}$ 表示MS图像与HRMS图像的实际残差信息, $g(x_1^{(i)}, x_2^{(i)})$ 表示MS图像与HRMS图像的预测残差信息, $x_1^{(i)}$ 、 $x_2^{(i)}$ 分别表示第*i*个下采样MS图像和PAN图像训练数据, $n$ 表示训练样本数。

Hu等(2019)设计了一个DSL(Deep Self-Learning model)用于自适应遥感图像融合。DSL融合框架由3部分构成,分别是点扩展函数PSF(Point Spread Function)计算模块、图像配准模块和图像融合模块。其中PSF计算模块实质上是一个CNN网络,用来完成对MS图像的PSF计算;图像配准模块完成MS图像与PAN图像的配准;图像融合模块实质上是一个DRN网络,使用PSF对配准后的MS图像与PAN图像进行降采样得到训练样本,对DRN网络进行训练,最后通过训练后的DRN网络自适应地完成图像的融合。DSL在GF-2、GF-1和JL-1A这3个卫星数据集的实验表明,较数个传统模型,DSL得到更高质量的融合图像。

上述基于CNN+DRN的全色图像锐化的共同点是融合框架里都包含两个网络,分别是CNN和DRN,融合框架较其他方法在网络的深度、宽度上

都有较大的突破, 这样就使得此类模型对 PAN 图像、MS 图像信息提取更为充分; 不同之处是 PAN 图像与 MS 图像共同作为 MSDCNN 的输入和 DSL 融合框架中融合子模块 DRN 的输入, 而 RSIFNN 的两分支网络单独对 PAN 图像和 MS 图像分别提取信息, 即 PAN 图像作为 RSIFNN 的一个分支网络的输入, MS 图像作为其另一分支网络的输入。

#### 4.2.6 基于 RNN+CNN 的全色图像锐化

RNN 可以将神经元输出再输入到该神经元, 网络结构的串联性便于处理时间序列数据, 从而使得对数据中的依赖关系能有效保持。然而, 传统的 RNN 对信息的长期保存往往难以实现。为解决这个难题, Hochreiter 和 Schmidhuber (1997) 设计了长短期记忆 LSTM (Long-Short Time Memory) 网络, LSTM 较一般的 RNN 具有更为灵活的内部结构。Shi 等 (2015) 把 2D 的卷积操作引入到 LSTM 设计出 CLSTM, 能够处理 2 维特征图信息并自动捕捉时间依赖性。同时, CLSTM 也能够用于处理三维数据 (Wang 等, 2020)。

除上述 5 类主流的协同方式全色图像锐化外, 也有学者将其它的 DL 模型运用于该研究领域, 如 RNN 模型。Wang 等 (2020) 设计了一个包含 CLSTM (Convolutional Long-Short Term Memory) 的融合框架 DCNet (dual-channel network, DCNet), 首次将基于 CLSTM 的 RNN 模型应用于全色图像锐化领域。DCNet 包含两个分别用于提取空间特征信息和光谱特征信息的 CNN 分支网络, 其中提取空间特征信息的子网络采用 2D CNN 实现, 提取光谱特征信息的子网络采用 3D CNN 实现。为了更充分地融

合提取的空间、光谱特征信息, DCNet 借助于  $S^2$ -CLSTM (Spatial-Spectral CLSTM) 模块并采用分层的融合策略, 最后通过重构模块得到融合图像 HRMS。DCNet 在 GaoFen-2 和 WorldView-2 两个卫星数据集上得到了比 PNN 等模型更好的融合结果。

#### 4.2.7 对比小结

上述介绍了协同方式下, DL+CS、DL+MRA、DL+MB、DL+注入模型、CNN+DRN 和 RNN+CNN 应用于解决遥感图像融合问题。本节对这 6 类方法各自的特点和异同点进行小结, 并在表 5 中将 6 类方法中典型算法进行性能提升对比分析, 可以得到以下结论:

- (1) 较 PNN 模型, 融合效果有所提升的方法中, 大多数采用了 DRN 网络结构或者 CNN+DRN 网络;
- (2) 较 PNN 模型提升融合效果最好的几个方法中, DL+注入模型等 DL 技术与传统融合方法相结合的方法表现出良好性能, 如 He 等 (2019)、Deng 等 (2020) 和 Liu 等 (2020) 提出的方法;
- (3) 相对 PNN 模型提升 10% 以上的方法中, 使用自适应学习机制, 可使得模型在不同的卫星数据集上的融合质量偏差限制在较小范围内 ( $\pm 5\%$  以内), 表现出较强的泛化性能, 如 Hu 等 (2019) 提出的方法;
- (4) 大多数算法在 SAM 指标上提升的幅度大于 (或等于) 在 QNR 指标上的提升幅度, 少数算法在 QNR 指标上得到较高提升, 如 Liu 等 (2020) 提出的方法。

表 5 协同方式算法比较 (基准算法分为传统方法 BSDS、GS, 基于 DL 的 PNN, 选择 SAM 和 QNR 为比较指标)

Table 5 Comparison of collaborative methods algorithms (the benchmark algorithm is divided into traditional methods BSDS and GS, and DL-based PNN, and SAM and QNR are selected as comparison indexes)

算法	体系架构	残差模块	损失函数	卫星数据集	SAM			QNR		
					BSDS $\uparrow$	GS $\uparrow$	PNN $\uparrow$	BSDS $\uparrow$	GS $\uparrow$	PNN $\uparrow$
Azarang 和 Ghassemian (2017)	AE+MRA	无	$F_2\left(\left\{u_p^i, z_p^i\right\}_{i=1}^N; T\right) = \frac{1}{N} \sum_{i=1}^N \left\  u_p^i - \hat{z}(x_p^i) \right\ ^2 + \frac{\lambda}{2} \sum_{j=1}^M \left\  W_j \right\ ^2$	QuickBird	—	44%	—	—	—	—
				WorldView-3	—	4%	—	—	—	—
Eghbalian 和 Ghassemian (2018)	CNN+MRA	无	$E = \frac{1}{n} \sum_{\omega \in \Omega, \hat{\omega} \in \hat{\Omega}} L(\omega, \hat{\omega})$	GeoEye-1	27%	—	1%	3%	—	0.2%
				WorldView-2	69%	—	20%	15%	—	6%
BDPN (Zhang 等, 2019b)	DRN+MRA	有	$\text{LOSS} = \lambda \text{loss}_1 + (1 - \lambda) \text{loss}_2$ $\text{loss}_i = \sqrt{\frac{1}{B} \sum_{b=1}^B \left( \text{RMSE}_i(b) \times e^{-u_i(b)} \right)^2}$	GaoFen-2	—	28%	4%	—	5%	-2%
				WorldView3	—	13%	11%	—	0.4%	1%

续表

算法	体系架构	残差模块	损失函数	卫星数据集	SAM			QNR		
					BDS ↑	GS ↑	PNN ↑	BDS ↑	GS ↑	PNN ↑
Di CNN1 (He等, 2019)	DRN+ 注入模型	有	$l(\theta) = \frac{1}{N_p} \sum_{i=1}^{N_p} \left\  \widehat{D}^{(i)}(G^{(i)}; \theta) + \widetilde{M}^{(i)} - Y^{(i)} \right\ _F^2$	Worldview-2 IKONOS	23% 20%	— —	18% 11%	— —	— —	
Fusion-Net (Deng等, 2020)	DRN+ 注入模型	有	$Loss(\Theta_{FS}) = \frac{1}{n} \sum_{k=1}^n \left\  \widetilde{MS}_{\{k\}} + f_{\Theta_{FS}}(P_{\{k\}}^D - \widetilde{MS}_{\{k\}}) - GT_{\{k\}} \right\ _F^2$	WorldView3	38%	40%	16%	9%	8% 3%	
MIPSM (Liu等, 2020)	DL+ 注入模型	有	$L = \sum_{j=1}^S \left\  G_{MS}^j + k^j g_D^j - f_{MS}^j \right\ _2^2$	IKONOS QuickBird	— —	23% -49%	82% 25%	— —	283% 139%	28% 23%
Ye等 (2019)	DRN+模型 (优化)	有	$L = \arg \min_x \frac{\alpha}{2} \sum_i \ y_i - k \cdot x_i\ _2^2 + \frac{\gamma}{2} \sum_{i,j} \ u_{i,j}\ _1 + \frac{\beta}{2} \sum_{i,j} \ D_j x_i - G_{i,j}\ _2^2 + \frac{\lambda}{2} \sum_{i,j} \ D_j x_i - u_{i,j} + v_{i,j}\ _2^2$	WorldView-2	29%	—	22%	11%	— 5%	
MSDCNN (Yuan等, 2018)	CNN+DRN	有	$LOSS' = \sum_{i=1}^S \left\  f_{MS(i)} - F_{MS(i)}' \right\ _2^2$ $\overline{LOSS}' = \frac{1}{Batchsize} \sum_{b=1}^{Batchsize} \sum_{i=1}^S \left\  f_{MS_b(i)} - F_{MS_b(i)}' \right\ _2^2$	QuickBird WorldView-2	— —	39% 31%	7% 5%	— —	— 6% 2%	
RSIFNN (Shao和 Cai, 2018)	CNN+DRN	有	$L = \frac{1}{n} \sum_{i=1}^n \left\  r^{(i)} - g(x_1^{(i)}, x_2^{(i)}) \right\ _2^2$	GF-1 QuickBird	— —	— —	11% 8%	— —	9% 2%	
DSL (Hu等, 2019)	CNN+DRN	有	$Loss_{MSE}(\widehat{X}, X) = \frac{1}{u} \sum_{i=1}^u (\widehat{X}_i - X_i)^2$	GF-2 GF-1	— —	28% 38%	18% 13%	— —	— —	
DCNet (Wang等, 2020)	RNN+CNN	有	$\theta \stackrel{\Delta}{=} \arg \min_{\theta} \frac{1}{N} \sum_{y_{M,n} \in X_{max}} L(\hat{y}_{M,n}, y_{M,n})$ $L(\hat{y}_M, y_M) = \left\  \Phi(x_p, x_M, \theta) - y_M \right\ _1 + \lambda \ \theta\ _2^2$	GF-2 WorldView-2	— —	— —	51% 33%	— —	5% 1%	

协同方式下, 上述6节中采用的方法共同点在于网络结构较为复杂。为便于直观理解, 下面给出一个典型的基于DL+注入模型的遥感图像融合算法框架图4 (Liu等, 2020)。图4中DL技术使用了CNN

网络和DRN网络, 其中CNN采用浅层网络, DRN为深层网络, 两者联合起来提取PAN图像的空间细节信息; 采用注入模型, 把获得的空间信息注入到上采样的MS图像, 从而得到HRMS融合图像。

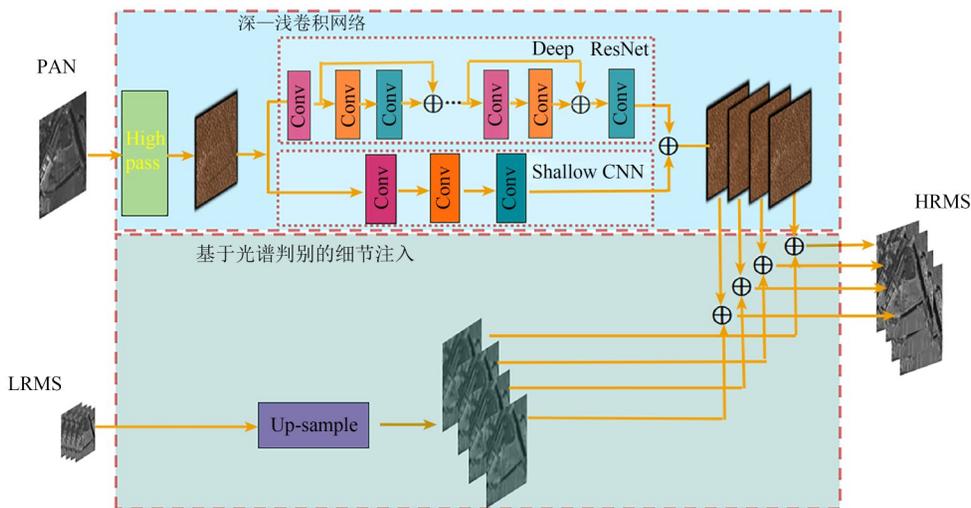


图4 典型的基于DL+注入模型的遥感图像融合框架

Fig.4 Typical remote sensing image fusion framework based on DL+ injection model

### 4.3 经典方式与协同方式对比小结

第4.1和4.2节分别从经典方式与协同方式两个视角对融合算法进行了比较和分析；本节对两种融合方式进一步归纳，可得出如下结论：

经典方式与协同方式的共同点是：都是利用DL技术能够自动学习复杂数据特征的优势，提取MS图像或PAN图像的特征信息，而这些信息正是HRMS融合图像需要保留的。

两种方式的不同点是：从融合模型结构上分析，一般经典方式结构较为简明，而协同方式由于是多个方法、框架的结合，往往结构较为复杂。此外，从表3、表5可看出，无论是经典方式，还是协同方式，融合模型结构的复杂性与损失函数复杂度是正相关的，如Cai等(2018)提出的CSDA方法、Ye等(2019)、Yuan等(2018)提出的MSDCNN方法以及基于GAN融合方法的损失函数。

从融合效果上看，从表4和表5可以看出，一般情况下，协同方式能够获得比传统融合算法以及经典方式更高质量的融合效果，如Zhang等(2019b)提出的BDPN方法。

为更好地理解经典方式与协同方式融合算法的发展脉络，我们选取了一些代表性的算法进行对比说明，如表6所示。从表6可以看出有大部分融合算法体系结构中包含DRN模块，损失函数多数采用L2 loss，采用GAN网络的融合算法的损失函数为Adversarial loss。

表6 经典方式与协同方式融合算法基本发展脉络  
Table 6 Development of classical and collaborative pansharpening

类别	代表算法	发表年份	体系架构	损失函数(类别)
经典方式	S-MSDA	2015	AE	L2 loss+ Frobenius regularization
	PNN	2016	CNN	L2 loss
	PanNet	2017	DRN	Frobenius loss (Frobenius norm)
	PSGAN	2018	GAN	Adversarial loss
	Vitale(2019)	2019	DRN	Perceptual loss
协同方式	MSDCNN	2018	CNN+DRN	L2 loss
	BDPN	2019	DRN+MRA	L2 loss
	Fusion-Net	2020	DL+注入模型	Frobenius loss
	DCNet	2020	RNN+CNN	L1 loss+L2 regularization

此外，从基于深度学习的全色图像锐化自身研究领域来看，早期的基于深度学习的融合方法，

多数利用了深度学习模型强大的数据拟合能力，较少地关注全色图像锐化的领域问题(Huang等, 2015)。随着研究的逐步深入，如通过深度学习方法和传统融合方法相结合等，这样设计出的融合模型就考虑到了光谱失真、空间失真等领域问题，并提出相应的解决方案，从而使得深度学习方法可以进一步提升全色图像锐化的效果，如PanNet、BDPN等融合模型。

## 5 全色图像锐化的应用

随着遥感技术的发展，全色图像锐化的应用领域愈加广泛，具体包括地物分类、目标识别以及变化检等。下面对此简要介绍：

(1) 地物分类。Gianinnetto等(2017)利用遥感融合图像，并结合基于目标的图像分析技术，设计了一个对复杂地貌进行层次分类的模型并应用于意大利威尼斯地区地物分类，达到了较高的分类准确率；Aguilar等(2013)利用GeoEye-1卫星和WorldView-2卫星的融合图像，对西班牙南部阿尔梅里亚省的遥感图像进行地物分类，并对两个卫星的融合图像分类结果加以对比分析，得出融合图像中的光谱信息对分类结果有较大影响；Pálsson等(2012)在IKONOS和QuickBird两个卫星图像数据集上，运用6种全色图像锐化方法分别得到融合图像，在此基础上对冰岛雷克雅未克和意大利罗马两个城市的遥感图像进行地物分类，并对其分类结果进行定量分析，得出如下结论：融合图像的光谱信息在可接受的范围内，空间信息保持越好越有利于地物分类。

(2) 目标识别。使用高质量的遥感融合图像对桥梁进行识别和损伤检测，在紧急救援和生命线安全评估中具有重要意义(Chen等, 2019)。Chen等(2019)设计了基于领域知识的方法对广东省惠阳市(区)遥感融合图像中的桥梁进行识别和损伤检测；Garzelli等(2009)使用遥感融合图像对意大利马尔彭萨机场的小型目标进行识别，并分析了不同融合方法对目标识别效果的差异。在地理空间目标识别中，道路信息的提取是一项重要的研究任务，Mohammadzadeh等(2006)把模糊逻辑和形态学应用于融合后的IKONOS卫星图像，对位于埃及金字塔周围城市、郊区和乡村等地区的道路中心线信息进行提取和识别。

(3) 地表变化检测。地表变化检测是一种评

估和预测灾害、监测地表覆盖变化的关键技术(李胜, 2018)。为了检测和评估自然灾害对水上桥梁的破坏程度, Chen等(2018)提出了一种利用遥感融合图像进行水上桥梁损伤检测与评估的方法, 并应用于浙江省舟山市水上桥梁的损伤检测和评估; Bovolo等(2010)利用意大利特伦托市的卫星图像, 从定性和定量方面分析了5种全色图像锐化方法所生成的融合图像在地表变化检测任务上的差异; 森林退化图的绘制能够改进地区或全球生态模型以及碳预算估计, Souza等(2003)利用遥感融合图像研究了亚马逊东部地区的森林退化情况, 并绘制了亚马逊东部地区的森林退化图。

## 6 展望

自2006年以来深度学习技术得到日新月异的发展, 并在多个领域取得突破性进展, 激励着许多学者将深度学习技术应用于全色图像锐化领域, 近6年来取得一系列研究成果, 但仍然有一些关键性问题需要深入探索, 通过分析、思考, 本文凝练出以下五个研究方向。

(1) 与遥感领域知识的结合。根据遥感图像的特点, 可利用遥感领域大量现有的特定领域知识有针对性地对深度学习模型进一步改进, 充分挖掘深度学习的潜力, 从而得到细节更为丰富, 光谱更加自然的融合图像(Deng等, 2020)。例如, 可以根据遥感领域特定的知识构建新的目标函数和度量方法, 找到MS图像、PAN图像以及融合图像之间的潜在联系; 根据MS图像多波段光谱信息以及PAN图像空间信息丰富的特点, 设计多层次、多尺度的全色图像锐化模型。

(2) 自监督的全色图像锐化方法。大部分基于深度学习的遥感图像融合方法都是通过使用大量的根据Wald协议(Wald等, 1997)所获得的模拟数据进行监督学习。然而, 所获得的模型受限于监督学习中所使用的样本数据, 不具备良好的泛化能力。当训练数据和测试数据的约束条件一致时, 可以获得显著的效果。然而, 当训练数据和测试数据的约束条件不一致时, 往往会导致光谱失真和空间细节丢失。自监督学习可以从未标记的数据中学习有效的特征表示而无需人工标注的标签信息, 这样使得自监督学习得到的模型不会受限于样本数据, 具有良好的泛化性能。所以, 基于自监督学习的遥感图像融合方法是一个

值得研究的方向。

(3) 全色图像锐化标准数据集。对于深度学习模型, 不断增加的高质量训练数据能够提升其性能。然而, 在当前研究中, 训练数据集大多从现有几个卫星数据集中由研究者自主选取, 不可避免地存在一定程度的主观性和局限性。标准数据集构建已经在图像处理其它领域取得成功, 如图像分类领域的ImageNet等公开标准数据集。然而, 在遥感图像融合领域还没有被深入广泛地探索。研究遥感图像融合的标准数据集, 构建大规模的标准数据集以及基于标准数据集的新型融合算法设计是未来可以努力的重要方向, 如Transformer在标准数据集上的应用、各类全色图像锐化模型在标准数据集上的比较研究等。

(4) 面向全色图像锐化应用的评价。全色图像锐化在诸如地物分类、目标识别等任务中有着重要的应用价值。然而从现有的文献来看, 评价各算法模型主要注重提升融合图像的质量, 对模型的应用评价却并未涉及。融合算法在具体应用中的性能如何, 不仅与高质量的融合图像有关, 更与具体应用领域的知识相关。因而, 面向应用的遥感图像融合算法评价是今后需要进一步研究的方向, 如某些花草类植物的识别更注重MS图像部分光谱波段的信息, 所以常用的定量评价指标就需要根据具体的应用进行调整以满足对光谱信息的有效评价(罗代清等, 2016)。

(5) 轻量级全色图像锐化深度模型。现有的基于深度学习的全色图像锐化模型, 在主观效果和客观指标上一般优于传统的融合方法。然而, 基于深度学习的全色图像锐化模型需要对大量的网络参数进行训练, 另外, 为追求性能的进一步提升, 研究者通常构建结构更为复杂的模型, 而造成全色图像锐化模型容量更大、训练时间更久, 这限制了全色图像锐化深度模型在实际中的应用。轻量级深度模型具有较小的网络容量和较低的时间复杂度, 同时对硬件的要求也较低, 为使基于深度学习的全色图像锐化模型更为实用, 利用诸如通道权重参数共享、快速引导滤波器等方法构建轻量级的全色图像锐化模型也是今后一个值得探索的方向。

## 7 结语

深度学习技术对非线性、结构复杂的数据具

有强大的学习表征能力, 在计算机视觉等研究领域取得了较大进展。受此启发, 许多研究者将该技术运用于像素级全色图像锐化的研究中, 并取得了一系列的研究成果。然而, 专门的、系统全面的基于深度学习的像素级全色图像锐化研究综述尚未见于文献。为此, 本文提出从经典方式和协同方式两个方面对现有基于深度学习的像素级全色图像锐化模型研究成果进行详尽的综述、分析挖掘其演进规律路径, 并结合深度学习最新的发展趋势, 凝练了基于深度学习的全色图像锐化研究领域的5个未来研究发展方向。

当前, 基于深度学习的像素级全色图像锐化研究尚处于初步阶段, 其研究成果日新月异。可以预见, 在未来的研究中, 面向标准数据集的且高效的基于深度学习的全色图像锐化模型将大量涌现。本文针对基于深度学习的全色图像锐化方法进行的综述研究希望能够为相关研究人员提供一些参考和启发。

## 参考文献(References)

- Aguilar M A, Saldaña M M and Aguilar F. 2013. GeoEye-1 and world-view-2 pan-sharpened imagery for object-based classification in urban environments. *International Journal of Remote Sensing*, 34(7):2583-2606 [DOI: 10.1080/01431161.2012.747018]
- Aiazzi B, Baronti S and Selva M. 2007. Improving component substitution pansharpening through multivariate regression of ms+pan data. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10):3230-3239 [DOI: 10.1109/tgrs.2007.901007]
- Alparone L, Aiazzi B, Baronti S and Garzelli A. 2015. *Remote sensing image fusion*. London, UK: CRC Press
- Alparone L, Wald L, Chanussot J, Thomas C, Gamba P and Bruce L M. 2007. Comparison of pansharpening algorithms: outcome of the 2006 grs-s data-fusion contest. *IEEE Transactions on Geoscience and Remote Sensing*, 45(10): 3012-3021 [DOI: 10.1109/TGRS.2007.904923]
- Azarang A and Ghassemian H. 2017. A new pansharpening method using multi resolution analysis framework and deep neural networks. // *Proceedings of the International Conference on Pattern Recognition and Image Analysis*. Shahrekord, Iran, 1-6 [DOI: 10.1109/pria.2017.7983017]
- Azarang A, Manoochehri H E and Kehtarnavaz N. 2019. Convolutional autoencoder-based multispectral image fusion. *IEEE Access*, 7: 35673-35683 [DOI: 10.1109/ACCESS.2019.2905511]
- Benzenati T, Kessentini Y, Kallel A and Hallabia H. 2020. Generalized laplacian pyramid pan-sharpening gain injection prediction based on cnn. *IEEE Geoscience and Remote Sensing Letters*, 17(4):651-655 [DOI: 10.1109/LGRS.2019.2928181]
- Bovolo F, Bruzzone L, Capobianco L, Garzelli A, Marchesi S and Nencini F. 2010. Analysis of the effects of pansharpening in change detection on VHR images. *IEEE Geoscience and Remote Sensing Letters*, 7(1): 53-57 [DOI: 10.1109/LGRS.2009.2029248]
- Bowman S R, Angeli G, Potts C and Manning C D. 2015. A large annotated corpus for learning natural language inference // *Proceedings of the 2015 Conference on Empirical Methods in Natural Language Processing (EMNLP)*. Lisbon, Portugal: 632-642 [DOI: 10.18653/v1/D15-1075]
- Cai W T, Xu Y, Wu Z B, Liu H Y, Qian L and Wei Z H. 2018. Pansharpening based on multilevel coupled deep network. // *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*. Valencia, Spain: 7046-7049 [DOI: 10.1109/IGARSS.2018.8518121]
- Carper W J, Lillesand T M and Kiefer P W. 1990. The use of intensity-hue-saturation transformations for merging spot panchromatic and multispectral image data. *Photogrammetric Engineering and Remote Sensing*, 56(4):459 - 467
- Chen C, Fu J Q, Gai Y Y, Li J, Chen L, Mantravadi V S and Tan A H. 2018. Damaged bridges over water: using high-spatial-resolution remote-sensing images for recognition, detection, and assessment. *IEEE Geoscience and Remote Sensing Magazine*, 6(3): 69-85 [DOI: 10.1109/MGRS.2018.2852804]
- Chen C, Fu J Q, Lu N, Chu Y L, Hu J C, Guo B Y and Zhao X. 2019. Knowledge-based identification and damage detection of bridges spanning water via high-spatial-resolution optical remotely sensed imagery. *Journal of the Indian Society of Remote Sensing*, 47(12): 1999-2008 [DOI: 10.1007/s12524-019-01036-z]
- Chen F J, Zhu F, Wu Q X, Hao Y M, Wang E D and Cui Y G. 2021. A survey about image generation with generative adversarial nets. *Chinese Journal of Computers*, 44(2): 347-369 (陈佛计, 朱枫, 吴清潇, 郝颖明, 王恩德, 崔芸阁. 2021. 生成对抗网络及其在图像生成中的应用研究综述. *计算机学报*, 44(2):347-369) [DOI: 10.11897/SP.J.1016.2021.00347]
- Chen J F, Pan Y and Chen Y. 2020. Remote sensing image fusion based on Bayesian GAN. *ArXiv Preprint ArXiv: 2009.09465*
- Chen M M, Guo Q, Liu M L and Li A. 2021. Pan-sharpening by residual network with dense convolution for remote sensing images. *Journal of Remote Sensing*, 25(6): 1270-1283 (陈毛毛, 郭擎, 刘明亮, 李安. 2021. 密集卷积残差网络的遥感图像融合. *遥感学报*, 25(6): 1270-1283) [DOI: 10.11834/jrs.20219411]
- Da Cunha A L, Zhou J P and Do M N. 2006. The nonsubsampling contourlet transform: theory, design, and applications. *IEEE Transactions on Image Processing*, 15(10):3089-3101 [DOI: 10.1109/TIP.2006.877507]
- Deng L J, Vivone G, Jin C and Chanussot J. 2020. Detail injection-based deep convolutional neural networks for pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 59(8): 6995-7010 [DOI: 10.1109/TGRS.2020.3031366]
- Dong C, Loy C C, He K M and Tang X O. 2016. Image super-resolution using deep convolutional networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 38(2):295-307 [DOI:

- 10.1109/TPAMI.2015.2439281]
- Eghbalian S and Ghassemian H. 2018. Multi spectral image fusion by deep convolutional neural network and new spectral loss function. *International Journal of Remote Sensing*, 39(12): 3983-4002 [DOI: 10.1080/01431161.2018.1452074]
- Garzelli A, Capobianco L and Nencini F. 2009. On the effects of pan-sharpening to target detection//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Cape Town, South Africa: II-136 – II-139 [DOI: 10.1109/IGARSS. 2009. 5418022]
- Garzelli A, Nencini F and Capobianco L. 2008. Optimal mmse pan-sharpening of very high resolution multispectral images. *IEEE Transactions on Geoscience and Remote Sensing*, 46(1):228-236 [DOI: 10.1109/TGRS.2007.907604]
- Gianinetto M, Rusmini M, Candiani G, Via G D, Frassy F, Maianti P, Marchesi A, Nodari FR and Dini L. 2017. Hierarchical classification of complex landscape with VHR pan-sharpened satellite data and obia techniques. *European Journal of Remote Sensing*, 47(1): 229-250 [DOI: 10.5721/EuJRS20144715]
- Goodfellow I J, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A C and Bengio Y. 2014. Generative adversarial nets//Proceedings of the Advances in Neural Information Processing Systems (NIPS). Montreal, Quebec, Canada: 2672-2680
- Guo Y C, Ye F and Gong H. 2019. Learning an efficient convolution neural network for pansharpening. *Algorithms*, 12(1):16:1-16:14 [DOI: 10.3390/a12010016]
- He G Q, Xing S Y, Xia Z Q, Huang Q Q and Fan J P. 2018. Panchromatic and multi-spectral image fusion for new satellites based on multi-channel deep model. *Machine Vision and Applications*, 29(6):933-946 [DOI: 10.1007/s00138-018-0964-5]
- He K M and Sun J. 2015. Convolutional neural networks at constrained time cost//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Boston, MA, USA: 5353-5360 [DOI: 10.1109/CVPR.2015.7299173]
- He K M, Zhang X Y, Ren S Q and Sun J. 2016. Deep residual learning for image recognition.//Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: 770-778 [DOI: 10.1109/CVPR.2016.90]
- He L, Rao Y Z, Li J, Chanussot J, Plaza A, Zhu J W and Li B. 2019. Pansharpening via detail injection based convolutional neural networks. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 12(4): 1188-1204 [DOI: 10.1109/JSTARS.2019.2898574]
- He L, Xi D H, Li J and Zhu J W. 2020. A spectral-aware convolutional neural network for pansharpening. *Applied Sciences*, 10(17):5809: 1-5809:18 [DOI: 10.3390/app10175809]
- Hinton G E and Salakhutdinov R R. 2006. Reducing the dimensionality of data with neural networks. *Science*, 313: 504-507 [DOI: 10.1126/science.1127647]
- Hinton G E, Osindero S and Teh Y W. 2006. A fast learning algorithm for deep belief nets. *Neural Computation*, 18(7):1527-1554 [DOI: 10.1162/neco.2006.18.7.1527]
- Hochreiter S and Schmidhuber J. 1997. Long short-term memory. *Neural Computation*, 9(8): 1735-1780 [DOI: 10.1162/neco.1997.9.8.1735]
- Hu J, He Z and Wu J M. 2019. Deep self-learning network for adaptive pansharpening. *Remote Sensing*, 11(20): 2395: 1-2395: 22 [DOI: 10.3390/rs11202395]
- Huang G, Liu Z, Maaten L V D, and Weinberger K Q. 2017. Densely connected convolutional networks//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Honolulu, HI, USA: 2261-2269 [DOI: 10.1109/CVPR.2017.243]
- Huang W, Fei X, Yin J R and Liu Y. 2018. A multi-direction subbands and deep neural networks based pan-sharpening method//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: 5139-5142 [DOI: 10.1109/IGARSS. 2018.8517817]
- Huang W, Xiao L, Wei Z H, Liu H Y and Tang S Z. 2015. A new pansharpening method with deep neural networks. *IEEE Geoscience and Remote Sensing Letters*, 12(5): 1037-1041 [DOI: 10.1109/LGRS.2014.2376034]
- Javan F D, Samadzadegan F, Mehravar S, Toosi A, Khatami R and Stein A. 2021. A review of image fusion techniques for pansharpening of high-resolution satellite imagery. *ISPRS Journal of Photogrammetry and Remote Sensing*, 171:101-117 [DOI: 10.1016/j.isprsjprs.2020.11.001]
- Jiang M H, Li J, Yuan Q Q, Shen H F, Liu X X and Xu M M. 2019. Differential information residual convolutional neural network for pansharpening. // Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Yokohama, Japan: 4865-4868 [DOI: 10.1109/IGARSS.2019.8900270]
- Kang X D, Li S T and Benediktsson J A. 2014. Pansharpening with matting model. *IEEE Transactions on Geoscience and Remote Sensing*, 52(8):5088-5099 [DOI: 10.1109/TGRS.2013.2286827]
- Khan M M, Chanussot J, Condat L and Montanvert A. 2008. Indusion: fusion of multispectral and panchromatic images using the induction scaling technique. *IEEE Geoscience and Remote Sensing Letters*, 5(1):98-102 [DOI: 10.1109/LGRS.2007.909934]
- Kim J, Lee J K and Lee K M. 2016. Accurate image super-resolution using very deep convolutional networks. // Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Las Vegas, NV, USA: 1646-1654 [DOI: 10.1109/CVPR. 2016.182]
- Laben C A and Brower B V. 2000. Process for enhancing the spatial resolution of multispectral imagery using pan-sharpening. *U. S.*, No.6011875
- Lee J, Seo S and Kim M. 2021. Sipsa-net: shift-invariant pan sharpening with moving object alignment for satellite imagery//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Online: 10161-10169 [DOI: 10.1109/CVPR46437. 2021.01003]
- Li H, Liu F, Yang S Y and Zhang K. 2016. Remote sensing image fusion based on deep support value learning networks. *Chinese Journal of Computers*, 39(8): 1583-1596 (李红, 刘芳, 杨淑媛,

- 张凯. 2016. 基于深度支撑值学习网络的遥感图像融合. 计算机学报, 39(8): 1583-1596 [DOI: 10.11897/SP.J.1016.2016.01583]
- Li J, Li Y F, He L, Chen J and Plaza A. 2020a. Spatio-temporal fusion for remote sensing data: an overview and new benchmark. *Science China Information Sciences*, 63(4) : 7-23 [DOI: 10.1007/s11432-019-2785-y]
- Li N, Huang N and Xiao L. 2017. Pan-sharpening via residual deep learning//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Fort Worth, TX, USA: 5133-5136 [DOI: 10.1109/IGARSS.2017.8128158]
- Li S T, Li C Y and Kang X D. 2021. Development status and future prospects of multi-source remote sensing image fusion. *Journal of Remote Sensing*, 25(1): 148-166 (李树涛, 李聪好, 康旭东. 2021. 多源遥感图像融合发展现状与未来展望. 遥感学报, 25(1):148-166) [DOI: 10.11834/jrs.20210259]
- Li S. 2018. Detection Method of Urban Land Cover Change Based on Combining Domain Knowledge and Deep Learning. Wuhan: Wuhan University (李胜. 2018. 联合领域知识和深度学习的城市地表覆盖变化检测方法. 武汉: 武汉大学)
- Li X, Pan Y, Gao A, Li L X, Mei S H and Yue S G. 2018. Pansharpening based on joint gaussian guided upsampling //Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: 7184-7187 [DOI: 10.1109/IGARSS.2018.8518150]
- Li X, Xu F, Lyu X, Tong Y, Chen Z Q, Li S Y and Liu D F. 2020b. A remote-sensing image pan-sharpening method based on multi-scale channel attention residual network. *IEEE Access*, 8: 27163-27177 [DOI: 10.1109/ACCESS.2020.2971502]
- Li Z Q and Cheng C Q. 2019. A cnn-based pan-sharpening method for integrating panchromatic and multispectral images using landsat 8. *Remote Sensing*, 11(22): 2606: 1-2606: 17 [DOI: 10.3390/rs11222606]
- Liu J M and Liang S L. 2016. Pan-sharpening using a guided filter. *International Journal of Remote Sensing*, 37(8):1777-1800 [DOI: 10.1080/01431161.2016.1163749]
- Liu L, Wang J, Zhang E L, Li B, Zhu X, Zhang Y Q and Peng J Y. 2020. Shallow - deep convolutional network and spectral-discrimination-based detail injection for multispectral imagery pan-sharpening. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 13:1772-1783 [DOI: 10.1109/JSTARS.2020.2981695]
- Liu P F, Xiao L and Li T. 2018a. A variational pan-sharpening method based on spatial fractional-order geometry and spectral-spatial low-rank priors. *IEEE Transactions on Geoscience and Remote Sensing*, 56(3):1788-1802 [DOI: 10.1109/TGRS.2017.2768386]
- Liu X Y, Wang Y H and Liu Q J. 2018b. Psgan: a generative adversarial network for remote sensing image pan-sharpening//Proceedings of the IEEE International Conference on Image Processing. Athens, Greece: 873-877 [DOI: 10.1109/ICIP.2018.8451049]
- Liu Y, Chen X, Wang Z F, Wang Z J, Ward R K and Wang X S. 2018c. Deep learning for pixel level image fusion: recent advances and future prospects. *Information Fusion*, 42:158-173 [DOI: 10.1016/j.inffus.2017.10.007]
- Lu X C, Zhang J P and Zhang Y. 2017. An improved non-subsampled contourlet transform-based hybrid pan-sharpening algorithm//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Fort Worth, TX, USA: 3393-3396 [DOI: 10.1109/IGARSS.2017.8127726]
- Luo D Q, Huang Y Q, Wang S Z, Fang Y P and Jin W B. 2016. Spectral signature analysis and band selection of Macheng rhododendron based on Landsat5 TM. *Hubei Agricultural Sciences*, 55(19): 4991-4994 (罗代清, 黄勇奇, 王书珍, 方元平, 金卫斌. 2016. 基于 Landsat5 TM 的麻城杜鹃花光谱分析与波段选择. 湖北农业科学, 55(19): 4991-4994) [DOI: 10.14088/j.cnki.issn0439-8114.2016.19.023]
- Ma J Y, Yu W, Chen C, Liang P W, Guo X J and Jiang J J. 2020. Pagan: an unsupervised pan-sharpening method for remote sensing image fusion. *Information Fusion*, 62: 110-120 [DOI: 10.1016/j.inffus.2020.04.006]
- Ma L, Liu Y, Zhang X L, Ye Y X, Yin G F and Johnson B A. 2019. Deep learning in remote sensing applications: a meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152:166-177 [DOI: 10.1016/j.isprsjprs.2019.04.015]
- Masi G, Cozzolino D, Verdoliva L and Scarpa G. 2016. Pansharpening by convolutional neural networks. *Remote Sensing*, 8(7): 594: 1-594:22 [DOI: 10.3390/rs8070594]
- Masi G, Cozzolino D, Verdoliva L and Scarpa G. 2017. CNN-based pansharpening of multi-resolution remote-sensing images// Proceedings of the Joint Urban Remote Sensing Event. Dubai, United Arab Emirates: 1-4 [DOI: 10.1109/JURSE.2017.7924534]
- Mohammadzadeh A, Tavakoli A and Valadan Z M. 2006. Road extraction based on fuzzy logic and mathematical morphology from pan-sharpened ikonos images. *The photogrammetric record*, 21(113): 44-60 [DOI: 10.1111/j.1477-9730.2006.00353.x]
- Otazu X, González-Audicana M, Fors O and Nunez J. 2005. Introduction of sensor spectral response into image fusion methods: application to wavelet-based methods. *IEEE Transactions on Geoscience and Remote Sensing*, 43(10): 2376-2385 [DOI: 10.1109/TGRS.2005.856106]
- Ozcelik F, Alganci U, Sertel E and Unal G. 2020. Rethinking cnn-based pansharpening: guided colorization of panchromatic images via gans. *IEEE Transactions on Geoscience and Remote Sensing*, 59(4):3486-3501 [DOI: 10.1109/TGRS.2020.3010441]
- Pålsson F, Sveinsson J R, Benediktsson J A and Aanaes H. 2012. Classification of pansharpened Urban satellite images. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 5(1):281-297 [DOI: 10.1109/JSTARS.2011.2176467]
- Pålsson F, Sveinsson J R, Ulfarsson M O and Benediktsson J A. 2015. Model based pansharpening method based on tv and mtf deblurring//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Milan, Italy: 33-36 [DOI: 10.1109/IGARSS.2015.7325690]
- Rao Y Z, He L and Zhu J W. 2017. A residual convolutional neural network for pan-sharpening//2017 International Workshop on Re-

- ote Sensing with Intelligent Processing. 1-4 [DOI: 10.1109/RSIP.2017.7958807]
- Ripley B D. 1996. Pattern Recognition and Neural Networks. Cambridge: Cambridge University Press
- Scarpa G, Vitale S and Cozzolino D. 2018. Target-adaptive cnn-based pansharpening. *IEEE Transactions on Geoscience and Remote Sensing*, 56(9):5443-5457 [DOI: 10.1109/TGRS.2018.2817393]
- Shah V P, Younan N H and King R L. 2008. An efficient pan-sharpening method via a combined adaptive pca approach and contourlets. *IEEE Transactions on Geoscience and Remote Sensing*, 46(5):1323-1335 [DOI: 10.1109/TGRS.2008.916211]
- Shahdoosti H R and Javaheri N. 2017. Pansharpening of clustered ms and pan images considering mixed pixels. *IEEE Geoscience and Remote Sensing Letters*, 14(6): 826-830 [DOI: 10.1109/LGRS.2017.2682122]
- Shao Z F and Cai J J. 2018. Remote sensing image fusion with deep convolutional neural network. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 11(5): 1656-1669 [DOI: 10.1109/JSTARS.2018.2805923]
- Shao Z M, Lu Z X, Ran M S, Fang L Y, Zhou J L and Zhang Y. 2020. Residual encoder-decoder conditional generative adversarial network for pansharpening. *IEEE Geoscience and Remote Sensing Letters*, 17(9):1573-1577 [DOI: 10.1109/LGRS.2019.2949745]
- Shi X J, Chen Z R, Wang H, Yeung D Y, Wong W K and Woo W C. 2015. Convolutional lstm network: a machine learning approach for precipitation nowcasting//Proceedings of the International Conference on Neural Information Processing Systems (NIPS). Montreal, Quebec, Canada: 802-810
- Souza C, Firestone L, Silva L M and Roberts D. 2003. Mapping forest degradation in the eastern amazon from spot 4 through spectral mixture models. *Remote Sensing of Environment*, 87(4):494-506 [DOI: 10.1016/j.rse.2002.08.002]
- Tsagkatakis G, Aidini A, Fotiadou K, Giannopoulos M, Pentari A and Tsakalides P. 2019. Survey of deep-learning approaches for remote sensing observation enhancement. *Sensors*, 19(18):3929:1-3929:39 [DOI: 10.3390/s19183929]
- Tu T M, Su S C, Shyu H C and Huang P S. 2001. A new look at IHS-like image fusion methods. *Information Fusion*, 2(3): 177-186 [DOI: 10.1016/S1566-2535(01)00036-7]
- Vinothini D S and Bama B S. 2019. Residual dense network for pansharpening satellite data. *IEEE Sensors Journal*, 19(24): 12279-12285 [DOI: 10.1109/JSEN.2019.2939844]
- Vitale S, Ferraioli G and Scarpa G. 2018. A CNN-based model for pansharpening of worldview-3 images//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Valencia, Spain: 5108-5111 [DOI: 10.1109/IGARSS.2018.8519202]
- Vitale S. 2019. A cnn-based pansharpening method with perceptual loss//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Yokohama, Japan: 3105-3108 [DOI: 10.1109/IGARSS.2019.8900390]
- Vivone G, Alparone L, Chanussot J, Mura M D, Garzelli A, Licciardi G A, Restaino R and Wald L. 2015. A critical comparison among pansharpening algorithms. *IEEE Transactions on Geoscience and Remote Sensing*, 53(5): 2565-2586 [DOI: 10.1109/TGRS.2014.2361734]
- Wald L, Ranchin T and Mangolini M. 1997. Fusion of satellite images of different spatial resolutions: assessing the quality of resulting images. *Photogrammetric Engineering & Remote Sensing*, 63(6): 691-699
- Wald L. 2000. Quality of high resolution synthesised images: Is there a simple criterion? // Proceedings of the Fusion of Earth data: merging point measurements, raster maps and remotely sensed images. Sophia Antipolis, France: 99-103
- Wang D, Bai Y P and Li Y. 2020. Multispectral pan-sharpening via dual-channel convolutional network with convolutional lstm based hierarchical spatial-spectral feature fusion. *ArXiv Preprint ArXiv: 2007.10060*
- Wang F, Guo Q and Ge X Q. 2021. Pan-sharpening by deep recursive residual network. *Journal of Remote Sensing*, 25(6): 1244-1256 (王芬, 郭擎, 葛小青. 深度递归残差网络的遥感图像空谱融合. 2021. 遥感学报, 25(6): 1244-1256) [DOI: 10.11834/jrs.20219250]
- Wei Y C and Yuan Q Q. 2017. Deep residual learning for remote sensed imagery pansharpening//Proceedings of the International Workshop on Remote Sensing with Intelligent Processing. Shanghai, China: 1-4 [DOI: 10.1109/RSIP.2017.7958794]
- Wei Y C, Yuan Q Q, Shen H F and Zhang L P. 2017. Boosting the accuracy of multispectral image pansharpening by learning a deep residual network. *IEEE Geoscience and Remote Sensing Letters*, 14(10):1795-1799 [DOI: 10.1109/LGRS.2017.2736020]
- Xiang Z K, Xiao L, Liu P F and Zhang Y F. 2019. A multi-scale densely deep learning method for pansharpening//Proceedings of the IEEE International Geoscience and Remote Sensing Symposium. Yokohama, Japan: 2786-2789 [DOI: 10.1109/IGARSS.2019.8898095]
- Xing Y H, Wang M, Yang S Y and Jiao L C. 2018a. Pan-sharpening via deep metric learning. *ISPRS Journal of Photogrammetry and Remote Sensing*, 145: 165-183 [DOI: 10.1016/j.isprsjprs.2018.01.016]
- Xing Y H, Wang M, Yang S Y and Zhang K. 2018b. Pansharpening with multiscale geometric support tensor machine. *IEEE Transactions on Geoscience and Remote Sensing*, 56(5):2503-2517 [DOI: 10.1109/TGRS.2017.2742002]
- Yang J F, Fu X Y, Hu Y W, Huang Y, Ding X H and Paisley J. 2017. Pannet: a deep network architecture for pan-sharpening// Proceedings of the IEEE International Conference on Computer Vision (ICCV). Venice, Italy: 1753-1761 [DOI: 10.1109/ICCV.2017.193]
- Yang S Y, Zhang K and Wang M. 2018a. Learning low-rank decomposition for pansharpening with spatial-spectral offsets. *IEEE Transactions on Neural Networks & Learning Systems*, 29(8): 3647-3657 [DOI: 10.1109/TNNLS.2017.2736011]
- Yang Y, Li L Y, Huang S Y, Zhang Y M and Lu H Y. 2020. Remote sensing image fusion with convolutional sparse representation based on adaptive dictionary learning. *Journal of Signal Processing*, 36(1): 125-138 (杨勇, 李露奕, 黄淑英, 张迎梅, 卢航远. 2020. 自适应

- 字典学习的卷积稀疏表示遥感图像融合. 信号处理, 36(1):125-138 [DOI: 10.16798/j.issn.1003-0530.2020.01.016]
- Yang Y, Tu W, Huang S Y and Lu H Y. 2020a. Pcdm: progressive cascade deep residual network for pansharpening. Remote Sensing, 12(4):676:1-676:20 [DOI:10.3390/rs12040676]
- Yang Y, Wu L, Huang S Y, Sun J C, Wan W G and Wu J H. 2018b. Compensation details-based injection model for remote sensing image fusion. IEEE Geoscience and Remote Sensing Letters, 15(5):734-738 [DOI: 10.1109/LGRS.2018.2810219]
- Yang Y, Wu L, Huang S Y, Tang Y J and Wan W G. 2018c. Pansharpening for multiband images with adaptive spectral - intensity modulation. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(9): 3196-3208 [DOI: 10.1109/JSTARS.2018.2849011]
- Yang Y, Wu L, Huang S Y, Wan W G, Tu W and Lu H Y. 2020b. Multi-band remote sensing image pansharpening based on dual-injection model. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 13: 1888-1904 [DOI: 10.1109/JSTARS.2020.2981975]
- Yao W, Zeng Z G, Lian C and Tang H M. 2018. Pixel-wise regression using u-net and its application on pansharpening. Neurocomputing, 312:364-371 [DOI: 10.1016/j.neucom.2018.05.103]
- Ye F J, Li X F and Zhang X L. 2018. Fusioncnn: a remote sensing image fusion algorithm based on deep convolutional neural networks. Multimedia Tools and Applications, 78(11): 14683-14703 [DOI: 10.1007/s11042-018-6850-3]
- Ye F, Guo Y C and Zhuang P X. 2019. Pan-sharpening via a gradient-based deep network prior. Signal Processing: Image Communication, 74:322-331 [DOI: 10.1016/j.image.2019.03.004]
- Yuan Q Q, Wei Y C, Meng X C, Shen H F and Zhang L P. 2018. A multiscale and multidepth convolutional neural network for remote sensing imagery pan-sharpening. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 11(3):978-989 [DOI: 10.1109/JSTARS.2018.2794888]
- Zeiler M D and Fergus R. 2014. Visualizing and understanding convolutional networks//Proceedings of the Conference on European Conference on Computer Vision (ECCV). Zurich, Switzerland: 818-833 [DOI:10.1007/978-3-319-10590-1\_53]
- Zhang H Q, Liu X Y, Yang S and Li Y. 2017. Retrieval of remote sensing images based on semisupervised deep learning. Journal of Remote Sensing, 21(3): 406-414 (张洪群,刘雪莹,杨森,李宇. 2017. 深度学习的半监督遥感图像检索. 遥感学报, 21(3):406-414) [DOI: 10.11834/jrs.20176105]
- Zhang L B, Zhang J, Lyu X R and Ma J. 2019a. A new pansharpening method using objectness based saliency analysis and saliency guided deep residual network//Proceedings of the IEEE International Conference on Image Processing. Taipei, China: 4529-4533 [DOI: 10.1109/ICIP.2019.8803477]
- Zhang L P, Li W S, Zhang C and Lei D J. 2020. A generative adversarial network with structural enhancement and spectral supplement for pan-sharpening. Neural Computing and Applications, 32(24): 18347-18359 [DOI: 10.1007/s00521-020-04973-w]
- Zhang Y J, Liu C, Sun M W and Ou Y J. 2019b. Pan-sharpening using an efficient bidirectional pyramid network. IEEE Transactions on Geoscience and Remote Sensing, 57(8):5549-5563 [DOI: 10.1109/TGRS.2019.2900419]
- Zhang Y L, Tian Y P, Kong Y, Zhong B E and Fu Y. 2018. Residual dense network for image super-resolution//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Salt Lake City, UT, USA: 2472-2481 [DOI: 10.1109/CVPR.2018.00262]
- Zhang Y T, Li X H, Zhou J L. 2019c. Sftgan: a generative adversarial network for pan-sharpening equipped with spatial feature transform layers. Journal of Applied Remote Sensing, 13(2): 026507: 1-026507:16 [DOI: 10.1117/1.JRS.13.026507]
- Zhong J Y, Yang B, Huang G Y, Zhong Fei and Chen Z Z. 2016. Remote sensing image fusion with convolutional neural network. Sensing and Imaging, 17(1):6:1-6:16 [DOI: 10.1007/s11220-016-0135-6]
- Zhou M, Yan K Y, Huang J, Yang Z H, Fu X Y and Zhao F. 2022. Mutual information-driven pan-sharpening//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR). New Orleans, Louisiana, USA: 1788-1798 [DOI: 10.1109/CVPR52688.2022.00184]

## Deep-learning approaches for pixel-level pansharpening

**YANG Yong<sup>1</sup>, SU Zhao<sup>1</sup>, HUANG Shuying<sup>2</sup>, WAN Weiguo<sup>3</sup>, TU Wei<sup>1</sup>, LU Hangyuan<sup>1</sup>**

*1.School of Information Technology, Jiangxi University of Finance and Economics, Nanchang 330032, China;*

*2.School of Software, Tiangong University, Tianjin 300387, China;*

*3.School of Software and Internet of Things Engineering, Jiangxi University of Finance and Economics, Nanchang 330032, China*

**Abstract:** Pansharpening is a fundamental problem in the field of remote sensing data processing. It has important research significance and application value in ground object classification and ground surface change detection. In recent years, Deep Learning (DL) has made great progress in natural language processing, computer vision, etc. and has promoted the development of pixel-level pansharpening technology. This work presents a systematic review of the research of DL in pansharpening from two aspects (classical and collaborative

approaches) and makes a prospect on this basis. First, the common datasets of pansharpening and the objective evaluation indexes of pansharpening, including reference and non-reference quality evaluation indexes, are provided. Second, the latest research results of DL-based pansharpening are introduced in two different categories from the classical and collaborative methods, and the performance of their algorithms is compared, analyzed, and summarized. The classical methods mainly include AE-based pansharpening, CNN-based pansharpening, DRN-based pansharpening, and GAN-based pansharpening methods. Meanwhile, the collaborative methods mainly include DL+CS-based pansharpening, DL+MRA-based pansharpening, DL+MB-based pansharpening, DL+injection model-based pansharpening, CNN+DRN-based pansharpening, and RNN+CNN-based pansharpening methods. In the comparative analysis of the classical and collaborative methods, the common point is that the DL technology can automatically learn the advantages of complex data features and extract the feature information of the MS or PAN image (i.e., the information that needs to be retained in the HRMS fusion image). The difference is that the structure of the classical mode is more concise, while that of the collaborative mode is more complex because it is the combination of multiple methods or frameworks. In addition, most early DL-based pansharpening methods utilized the powerful data fitting ability of the DL model and seldom paid attention to the field of pansharpening problems. With the gradual deepening of research, such as using DL methods combined with traditional pansharpening methods, this designed fusion model considers spectral and spatial distortions. Accordingly, the DL methods can further enhance the pansharpening effect. Thirdly, the three main application fields of pansharpening are analyzed, such as object classification, target recognition, and surface change detection. Finally, this work discusses the future research direction of DL-based pansharpening in combination with remote sensing knowledge to fully tap the potential of DL to obtain fused images with richer details and more natural spectra. For example, for the evaluation of pansharpening application, the performance of pansharpening in a certain application is related not only to the high quality of fusion image but also to the knowledge of a specific application field. Accordingly, the application-oriented pansharpening evaluation algorithms will be the focus of future study. Furthermore, DL-based pansharpening needs to train a large number of network parameters, resulting in a longer training time for the pansharpening model. The lightweight depth model has a smaller network capacity, lower time complexity, and lower hardware requirements. Therefore, constructing a lightweight pansharpening model is a promising future direction.

**Key words:** pansharpening, deep learning, classical mode, collaborative mode, quality evaluation, remote sensing image fusion

**Supported by** National Natural Science Foundation of China (No.62072218, 61862030); Natural Science Foundation of Jiangxi Province (No.20192ACB20002, 20192ACBL21008)