

# 共享超分的双分支遥感图像时空融合网络

方 帅<sup>1,2</sup>, 张小溪<sup>1</sup>, 张 晶<sup>1</sup>

(1. 合肥工业大学计算机与信息学院, 安徽合肥 230000; 2. 工业安全与应急技术安徽省重点实验室, 安徽合肥 230000)

**摘 要:** 本文从空间维度和时间维度分析了场景弱变化区域和类型变化区域的融合规律、物理模型的差异性和效果上的互补性, 提出了共享超分辨率的双分支(Shared Super-Resolution Dual-Branch, SSRDB)遥感图像时空融合算法. 该算法具有如下3个特点: (1) 构建了互补性的网络框架, 虽然该框架是端到端的深度学习模型, 但每个模块有各自的物理意义和任务, 通过增加中间监督, 分别实现空间维的超分建模, 时间维的变化预测建模, 以及两者优势互补的融合建模; (2) 对变化预测的数学表示进行推演, 利用一个非线性补偿模块, 使得两分支共享超分模块, 在共享超分模块和递归复用超分单元的双重策略下, 显著降低了网络参数; (3) 递归超分模块使用固定的2倍率超分单元, 在有效监督和有效参考下, 渐进式进行特征增强与图像重建, 这可以有效提高超分精度, 且通过调整超分单元个数, 灵活适应不同倍率差异的时空融合任务. SSRDB算法在空间和光谱特性上以及变化区域上都展现了优秀的融合效果, RMSE (Root Mean Squared Error)、SAM (Spectral Angle Mapper) 和 SSIM (Structural Similarity) 3个定量评价指标显示, 在 CIA (Coleambally Irrigation Area) 数据集上分别优于次优方法 7.067%、2.065%、0.563%; 在 LGC (Lower Gwydir Catchment) 数据集上分别优于次优方法 5.319%、5.490%、1.455%; 在 Nanjing 数据集上分别优于次优方法 6.486%、16.290%、0.481%.

**关键词:** 遥感图像; 时空融合; 双分支; 图像超分; 卷积神经网络

**基金项目:** 国家自然科学基金(No.61872327, No.61175033)

**中图分类号:** TP751.1

**文献标识码:** A

**文章编号:** 0372-2112(2025)02-0581-14

**电子学报 URL:** <http://www.ejournal.org.cn>

**DOI:** 10.12263/DZXB.20240324

## Shared Super-Resolution Dual-Branch Network for Spatiotemporal Fusion of Remote Sensing Images

FANG Shuai<sup>1,2</sup>, ZHANG Xiao-xi<sup>1</sup>, ZHANG Jing<sup>1</sup>

(1. School of Computer and Information, Hefei University of Technology, Hefei, Anhui 230000, China;

2. Anhui Province Key Laboratory of Industry Safety and Emergency Technology, Hefei, Anhui 230000, China)

**Abstract:** In this paper, we analyze the fusion law of scene weak change region and type change region, the difference of physical model and the complementarity of effect from spatial and temporal dimensions, and propose a shared super-resolution dual-branch (Shared Super-Resolution Dual-Branch, SSRDB) remote sensing image spatio-temporal fusion algorithm. The algorithm has the following three characteristics: (1) A complementary network framework is constructed. Although the framework is an end-to-end deep learning model, each module has its own physical meaning and task. By adding intermediate supervision, the super-resolution modeling of spatial dimension, the change prediction modeling of time dimension and the fusion modeling of the two advantages are realized respectively. (2) The mathematical representation of the change prediction is deduced, and a nonlinear compensation module is used to make the two branches share the super-resolution module. Under the dual strategy of sharing super-resolution module and recursive multiplexing super-resolution unit, the network parameters are significantly reduced. (3) The recursive super-resolution module uses fixed 2-magnification super-resolution units to gradually enhance features and reconstruct images under effective supervision and reference, which can effectively improve the precision of super-resolution, and flexibly adapt to spatio-temporal fusion tasks with different magnification differences by adjusting the number of super-resolution units. The SSRDB algorithm shows excellent fusion effect in spatial and spectral characteristics and change regions. The three quantitative evaluation indexes of RMSE (Root Mean Squared Error)、SAM (Spectral Angle Mapper) and SSIM (Structural Similarity) show that it is superior to the sub-

optimal method on the CIA (Coleambally Irrigation Area) dataset by 7.067%, 2.065% and 0.563%, respectively. On the LGC (Lower Gwydir Catchment) dataset, it is superior to the sub-optimal method by 5.319%, 5.490% and 1.455%, respectively. On the Nanjing dataset, it is superior to the suboptimal method by 6.486%, 16.290% and 0.481%, respectively.

**Key words:** remote sensing image; spatiotemporal fusion; dual-branch; image super-resolution; convolutional neural network

**Foundation Item(s):** National Natural Science Foundation of China (No.61872327, No.61175033)

## 1 引言

卫星时序图像在城市资源检测<sup>[1]</sup>、土地覆盖变化检测<sup>[2]</sup>、生态系统动态监测<sup>[3]</sup>等对地观测任务中发挥着重要作用。然而卫星传感器存在光谱分辨率、空间分辨率和回访周期相互制约的情况,再加上云层阻碍和设备故障等原因,单个卫星无法提供高质量密集序列图像<sup>[4]</sup>。为此出现了卫星星座,利用多星之间的互补性,满足对地观察的各种应用需求,这使得融合技术受到广泛关注。本文主要关注遥感图像时空特性的融合,将低空间分辨率高回访周期的卫星图像(简称为粗图像)和高空间分辨率低回访周期卫星图像(简称为细图像)进行融合,生成具有高空间分辨率的密集图像序列。

如图1所示,从空间维度,时空融合建模从粗图像向细图像的映射关系,属于图像超分辨率问题;从时间维度,建模场景随时间改变的关系,属于变化预测问题。近年来提出了大量的时空融合方法,可分为基于加权的、基于解混的、基于学习的(基于稀疏表示和基于深度学习)的方法。从各类方法的底层描述看,加权方法将时空融合表示为变化预测;解混和稀疏表示方法将其表示为超分问题;深度学习方法是数据驱动的端到端模型,但观察其输入、网络结构和特征,可以发现前期侧重于图像超分,后来逐渐关注了变化,兼顾2个方面。下面将依次分析上述4类时空融合方法。

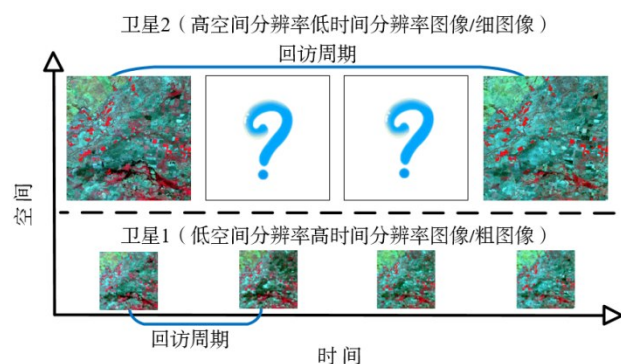


图1 时空融合过程的维度分析示意图

基于权重的方法在邻域内根据像素的相关性确定其对目标像素的权重。典型的算法有STARFM (Spatial and Temporal Adaptive Reflectance Fusion Model)<sup>[5]</sup>及其改

进算法STAARCH (Spatial Temporal Adaptive Algorithm for Mapping Reflectance Change)<sup>[6]</sup>、ESTARFM (Enhanced Spatial and Temporal Adaptive Reflectance Fusion Model)<sup>[7]</sup>和mESTARFM (Modified version of ESTARFM)<sup>[8]</sup>,这类算法稳定性好,但精度不高,加权方式会导致图像有些模糊,尤其在细节丰富的异质区域体现得更为明显。基于解混的方法采用线性光谱混合理论分解粗像素以预测细像素。最早被提出的是MMT (Multisensor Multiresolution Technique)<sup>[9]</sup>算法,利用参考细-粗图像确定端元种类和空间丰度,在“同类同谱”的前提下,估计预测粗图像的端元谱。其改进算法STDDFA (Spatial Temporal Data Fusion Approach)<sup>[10]</sup>引入滑动窗口实现“局部同类同谱”,ESTDFM (Enhanced Spatial and Temporal Data Fusion Model)<sup>[11]</sup>以单步长滑动,进一步提高了同类光谱异质性。这类算法要求在预测时刻不能出现新的端元;且假设“空间域上丰度不随时间变化”,导致在类型变化区域会产生错误的融合结果。此外,还有两者混合的方法,FSDAF (Flexible Spatiotemporal Data Fusion)<sup>[12]</sup>基于谱分解分析和薄板样条插值,通过插值算法预测空间变化,提升异质区域和类型变化的融合效果。Fit-FC (Regression Model Fitting, Spatial Filtering and Residual Compensation)<sup>[13]</sup>采用回归模型进行拟合,通过空间滤波消除块状伪影,并采用残差补偿保持光谱信息。

基于学习的方法可以进一步细分为基于稀疏表示和基于深度学习的方法。基于稀疏表示的时空反射率融合模型SPSTFM (SParse-representation-based Spatio-temporal Reflectance Fusion Model)<sup>[14]</sup>提出2个强假设:参考和预测日期字典表示不变,粗细图像稀疏系数相同。误差有界的半耦合字典学习EBSCDL (Error-bound-regularized Semi-coupled Dictionary Learning)<sup>[15]</sup>算法允许字典微调 and 稀疏系数线性变化。基于结构稀疏和半耦合字典的时空融合模型<sup>[16]</sup>利用系数包含的结构信息,提出单通道聚类结构一致性和多通道联合一致性约束,提高了融合结果的鲁棒性。粗细图像之间不仅存在分辨率差异还包含传感器差异和成像环境差异,因此粗细稀疏系数一致在实际中很难满足。尤其当粗细图像尺度差异大时,即便字典是理想的情况下,融合结果也表现不佳。另外,稀疏表示的方法采用稀疏域非线性优化算法,优化过程具有较高的计算成本。

深度学习方法凭借着优越的特征表征能力和复杂非线性关系映射能力,在遥感时空融合领域展现出了优异的性能. 基于深度卷积神经网络的卫星图像时空融合 STFDCNN (Spatiotemporal Satellite Image Fusion Using Deep Convolutional Neural Network)<sup>[17]</sup>算法,是第一种将卷积神经网络 CNN (Convolutional Neural Network) 技术用于时空融合的算法,由 2 个主要模块构成,其中非线性映射模块解决传感器之间的差异,超分模块提高图像空间分辨率. 深度卷积时空融合网络 DCSTFN (Deep Convolutional Spatiotemporal Fusion Network)<sup>[18]</sup>认为粗细图像特征差异保持不变,该网络包含粗图像特征提取分支和细图像特征提取分支,并将参考粗细图像对特征差与预测粗图像特征合并来预测融合结果. 其增强版 EDCSTFN (Enhanced DCSTFN)<sup>[19]</sup>增加了输入,引入残差特征提取器,改进了特征合并方式以及目标函数. STFGAN (Spatiotemporal Fusion Method Using a GAN)<sup>[20]</sup>是第一种基于生成对抗网络 GAN (Generative Adversarial Network) 的时空融合方法,通过 2 次超分辨率 GAN (Super-resolution Using a GAN, SRGAN)<sup>[21]</sup>缓解粗细图像之间过大的分辨率差异. GAN-STFM (GAN-based Spatio-temporal Fusion Model)<sup>[22]</sup>仅使用预测时刻的粗图像和一张任意时刻的精细图像进行预测,打破了先验图像的时间限制,且引入了条件生成对抗网络 CGAN (Conditional Generation Adversarial Network) 和可切换归一化技术,将预测粗图像作为条件引导网络收敛方向,取得了理想的融合结果. 上面提及的算法偏重于利用空间信息,基于双流卷积神经网络的时空融合算法<sup>[23]</sup>针对 3 对序列图提出时间依赖性和时间一致性先验,并由此构建网络结构和目标函数解决时空融合问题. MLFF-GAN (Multilevel Feature Fusion with Generative Adversarial Network)<sup>[24]</sup>采用类似 U-Net 的网络结构提取多级特征,然后通过特征规范化消除传感器差异,并引入空间注意力学习局部变化权重,关注了变化区域融合,最后进行多级特征重建,表现了突出的性能. OPGAN (Generative Adversarial Network with One Prior Image Pair)<sup>[25]</sup>的生成器由变化模块(时间差值图像),传感器变化模块(粗细差值图像)和基本信息模块(参考细图像)分别提取特征,并对变化信息分支进行监督,加强对变化的关注. MSFusion (Multistage for Remote Sensing Image Spatiotemporal Fusion)<sup>[26]</sup>将 CNN 和 Transformer 结合,提取多尺度的细图像空间特征和粗图像时间差异特征,利用 Transformer 学习全局时间相关性,将时间特征与空间特征进行充分的融合.

基于深度学习方法整体是数据驱动的端到端模型,从早期关注局部空间特征,发展到关注时空特征和全局特征. 对于场景中弱变化区域(地物目标没变),由于场

景轮廓、纹理等没发生变化,先验图像的细节特征应该被充分利用,变化预测模型有利于融合结果;对于类型变化区域(地物目标变化),局部先验图像的细节特征可能对融合结果带来误导,超分模型更有利于融合结果. 空间超分模型和变化预测的物理原理不同,对参考图像的依赖不同,如果将 2 个任务统一建模,模型复杂度,这需要充分的样本保证. 然而,卫星遥感时空融合要求来源于不同卫星的粗细图像在时间上对齐,在时序上相邻,再加上天气和设备本身的问题,面临着大数据少样本的局面. 另外,场景中类型变化区域显著少于弱变化区域,有监督学习过程中少样本的变化区域性能明显下降. 因此,在样本不充分且不均衡的情况下,最终表现为模型在类型变化区域性能差,且泛化能力弱,这是目前时空融合面临的共性问题.

针对“类型变化区域性能差”的问题,本文分别设计了空间维的超分和时间维的变化预测双分支网络结构,2 个分支功能各有侧重,相对于通用模型,表达的关系简单;增加了一个融合模块学习两分支的互补性. 相对于目前主流的时空融合网络结构,本文 2 个分支都重建了融合结果,又增加了融合模块,网络显得较为庞大. 针对这一问题,将变化预测转换为差分图像的超分,上下分支共享超分网络,且在超分网络内部引用递归思想,复用超分子模块,实现逐步渐进超分,减少模型复杂度,提高超分精度. 本文的主要贡献在于如下 2 个方面:

(1) 为了利用时空融合问题中时间维变化预测与空间维超分辨率的互补性,以及克服样本不均衡和不充分条件下对模型性能的负面影响,提出了共享超分的双分支时空融合网络,两分支各有侧重,将两分支预测结果通过融合网络各取所长,提升整体预测精度.

(2) 通过转化变化预测模型的数学表示,构建其与超分问题的关系,让双分支实现了共享超分模块;对于超分模块,利用递归思想将超分任务分解成级联的多次超分,多次超分复用相同的超分单元;通过共享超分模块和递归超分结构,减少了模型参数,提高了融合精度.

## 2 问题描述

从空间维度上,时空融合可看作粗图像向细图像映射的超分,融合结果  $\hat{F}_2^{\text{SR}}$  表示为

$$\hat{F}_2^{\text{SR}} = f_{I_{\text{SR}}}(C_2|F_1) \quad (1)$$

其中,  $F_1$  是参考细图像,希望从  $F_1$  学习到目标特征表示,  $C_2$  是预测粗图像,  $f_{I_{\text{SR}}}$  是图像超分模型,  $\hat{F}_2^{\text{SR}}$  是  $C_2$  对应细图像的估计.

从时间维度上,时空融合可看作是从参考时刻到预测时刻的变化预测,融合结果  $\hat{F}_2^{\text{CP}}$  可表示为

$$\hat{F}_2^{\text{CP}} = f_{C_{\text{SR}}}(C_2 - C_1|F_1) + F_1 \quad (2)$$

$C_1$  是参考粗图像,已知粗图像变化  $C_2 - C_1$ ,估计细

图像变化  $F_2 - F_1$ , 可以解读为对差分图像  $C_2 - C_1$  的超分, 变化超分用  $f_{C,SR}$  表示.  $f_{I,SR}$  和  $f_{C,SR}$  本质都是基于参考图像  $F_1$  的超分模型, 分别为图像超分和差分图像超分. 若两者共用相同的超分模型  $f_{SR}$ , 必须满足下式:

$$f_{SR}(C_2 - C_1|F_1) = f_{SR}(C_2|F_1) - f_{SR}(C_1|F_1) \quad (3)$$

上式是线性系统叠加原理的体现, 当且仅当  $f_{SR}$  是线性系统才能成立. 然而, 超分模型  $f_{SR}$  包含高低分辨率映射关系和传感器映射关系, 满足线性系统的可能性为 0.

为了式(1)和式(2)中的  $f_{I,SR}$  和  $f_{C,SR}$  用相同的超分模型  $f_{SR}$  替代, 引入模块  $g(C_1, C_2, F_1)$  弥补线性叠加的误差, 式(1)和式(2)可重新表示为

$$\hat{F}_2^{SR} = f_{SR}(C_2|F_1) \quad (4)$$

$$\hat{F}_2^{CP} = f_{SR}(C_2 - C_1|F_1) + g(C_1, C_2, F_1) + F_1 \quad (5)$$

其中,  $f_{SR}$  仍然是以  $F_1$  为参考的超分; 式(5)相较于于式(2)增加了  $g(C_1, C_2, F_1)$ , 弥补“图像”超分和“差分图像”超分共享模型的系统误差.

这里说明一下, 本文  $F$  表示细图像,  $C$  表示粗图像, 下标 1 表示参考时刻, 下标 2 表示预测时刻, CP 是变化预测的缩写, SR 是超分的缩写,  $\hat{\cdot}$  表示模型预测结果.

### 3 方法

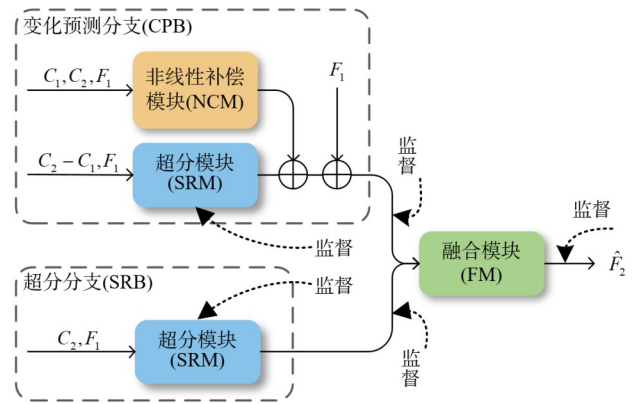
本节详细介绍共享超分的双分支时空融合算法, 包括整体网络架构、3 个主要模块以及损失函数.

#### 3.1 网络整体结构

算法整体结构如图 2 所示, 2 个分支: 上面变化预测分支 CPB (Change Prediction Branch), 下面超分辨率分支 SRB (Super-Resolution Branch); 3 个模块: 两分支共享的超分模块 SRM (Super-Resolution Module), 用于弥补线性叠加误差的非线性补偿模块 NCM (Nonlinear Compensation Module), 学习两分支互补性的融合模块 FM (Fusion Module).

图 2 中虚线是训练阶段施加的监督, 可以看出本算法除了对最终预测结果有监督, 2 个分支还有各自的监督, 让两分支真正做到各司其责, 避免复杂通用模型在样本不充足和不均衡情况下的弊端. 另外, 图 2 的超分模块是由多个超分单元构成的递归超分, 对每个超分单元施加了相应的监督, 保证训练的稳定性.

这里要特别说明一点, 2 个分支共用超分模块的输入是不同的, CPB 中超分模块的输入是粗图像差  $C_2 - C_1$ , SRB 中的输入是粗图像  $C_2$ , 2 个输入的分布差别较大, 为了让两分支更好地共享超分模块, 需要对输入进行规范化处理, 具体做法是让  $C_2 - C_1$  与  $C_2$  的局部均值进行相加.



注:  $C_1$  和  $C_2$  为参考日期和预测日期粗图像,  $F_1$  为参考时刻的细图像,  $\hat{F}_2$  为预测图像.

图 2 整体结构图

#### 3.2 超分模块

超分模块的设计利用了递归思想, 将超分问题进行分解, 在不同尺度上执行相同倍率的带参考的超分, 逐步重建高分辨率图像, 如图 3(a) 所示. 由于每个超分单元的超分倍率相同 (本文是 2 倍率), 因此, 各个超分单元可以复用, 如图 3(b) 的蓝色块, 其网络结构和参数是完全相同的. 下面分别介绍递归的超分模块和被复用的超分单元.

##### 3.2.1 超分模块

超分模块上半部分的输入是参考细图像  $F_1$ , 从右向左以 2 倍倍率对其进行下采样, 为各个超分单元提供相应尺度 (超分单元目标尺度) 的参考细图像  $F_1^{d,k}$  ( $k = 1, 2, \dots, n-1, n$  为超分单元个数), 这有利于高质量特征的提取和学习. 超分模块下半部分的输入是粗图像  $C$ , 从左到右每经过一个超分单元 SRU (Super-Resolution Unit) 实现对图像超分 2 倍的任务, 逐步超分直至获得目标图像分辨率的预测图像  $\hat{F}_C^{SR}$ .

##### 3.2.2 超分单元

超分单元负责在高分辨参考图像信息的帮助下, 完成对图像进行超分 2 倍的任务. 其输入包括低分辨率图像, 特征和参考细图像; 输出包括提高了 2 倍分辨率的特征和超分图像. 除了第一个 SRU 的低分辨率输入是初始粗图像  $C$  和其经简单卷积得到的初步特征, 其余 SRU 的低分辨率图像和特征输入均由其前一个 SRU 提供, 即在 SRU 之间不断进行特征和图像的传递, 逐步完成对特征和图像的超分, 下面分别描述特征超分和图像超分的过程.

(1) 特征超分: 将前一个 SRU 的重建特征  $X_{k-1}$  以及相应的高分辨率参考图像  $F_1^{d,k}$  进行合并,  $X_{k-1}$  含有场景空间信息,  $F_1^{d,k}$  含有高质量的空间信息, 采用如图 3(c) 所示 ResNet 残差网络实现了特征重构. ResNet 使用了局部残差和全局残差, 通过卷积来调整通道数. 相比于

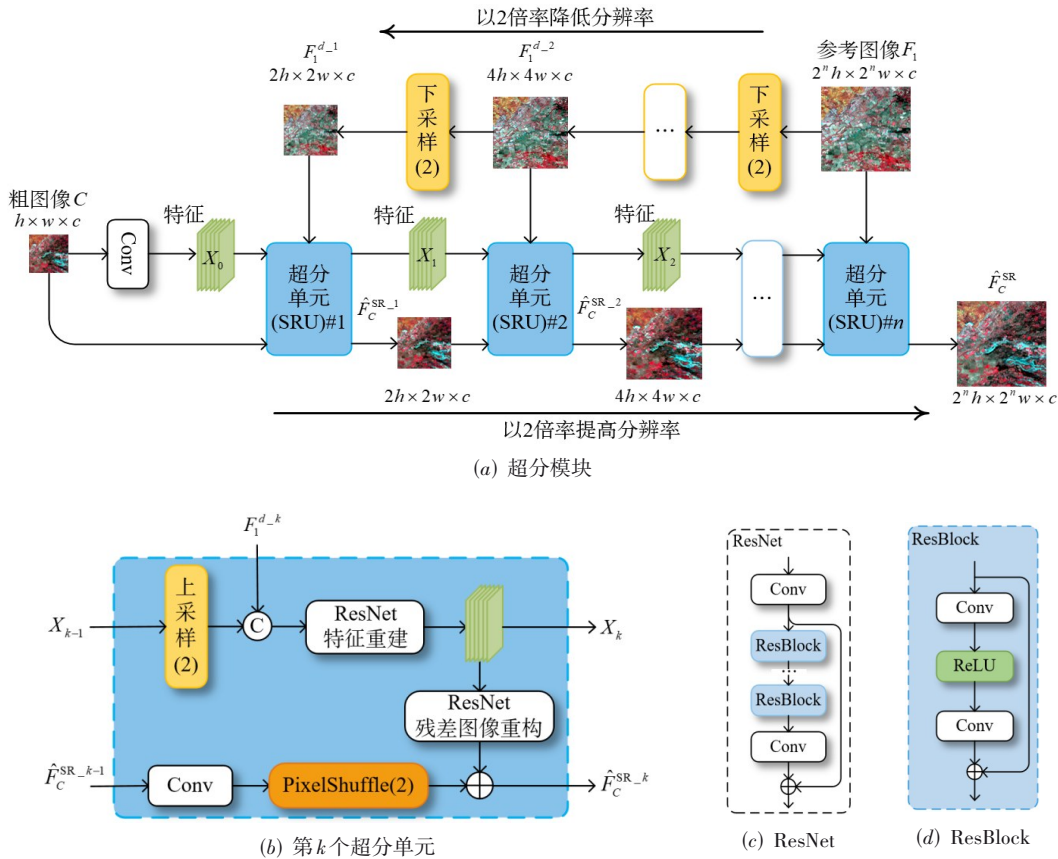


图3 超分模块结构图

串联所有残差块的形式,建立了从残差层之间的跳层连接,加强了特征传播,能有效地提取特征,防止梯度爆炸.此外,残差块 ResBlock 采用由 Conv-ReLU-Conv 组成的简单残差,如图 3(d)所示,相比于原始残差块,去除了批量归一化层,避免破坏图像原本分布,更有利于超分问题.

(2)图像超分:在特征超分辨的基础上,通过图 3(c)所示的 ResNet 实现残差图像重建,与 ResNet 特征重建相比,其残差结构包含较少的残差块.之后将残差图像与前一个 SRU 上采样图像相加,完成图像超分.  $\hat{F}_c^{SR,k-1}$  是低分辨率图像输入,是前一个 SRU 的图像重建结果.先使用卷积操作将通道数变为原本的 4 倍,然后使用亚像素卷积 (Pixel Shuffle) 重新排列通道,由此完成对图像的 2 倍上采样.相比于传统的插值算法,PixelShuffle 参数更少,且它通过像素的重排来进行上采样,有助于保留图像的细节纹理信息.

总的来说,超分模块的特点包括如下 3 点:(a)渐进式的特征注入与融合重建,每次超分 2 倍,这种方式能更好适应大比例因子的超分;(b)超分单元复用,大大减少网络参数;(c)每个超分单元在特征域融合参考细图像信息,进行特征增强,在图像域加入增强后特征的残差

重建图像,进行图像超分.同时对输出图像进行监督,在增强细节的同时保持光谱的有效性,模型更加稳定.

### 3.3 非线性补偿模块

非线性补偿模块旨在让变化预测模型和超分模型能够统一超分模块,弥补线性叠加的误差.其模块结构如图 4 所示,ResNet 结构如图 3(c)所示.将  $C_1$  和  $C_2$  上采样到  $F_1$  的尺寸后,与  $F_1$  在通道维进行拼接,经 2 个 ResNet 后得到非线性补偿值.采用残差网络描述非线性映射关系,一方面是残差网络在特征提取与重建的过程中具有良好的性能,另一方面是残差网络可以有效缓解梯度消失的问题,适合进行非线性补偿.

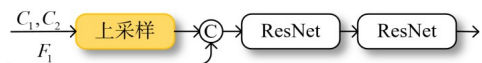


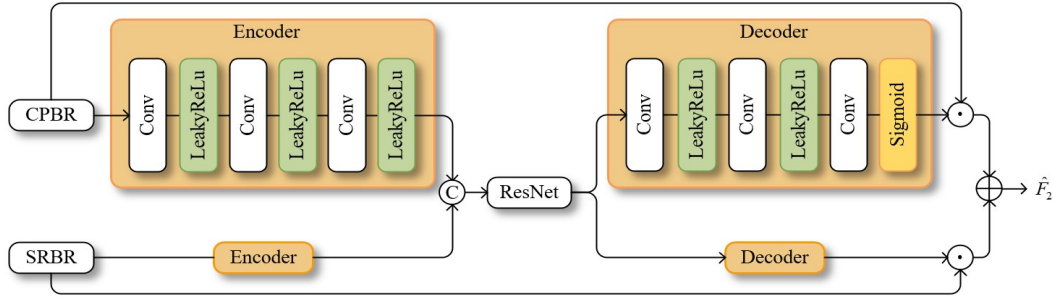
图4 非线性补偿模块结构图

### 3.4 融合模块

融合模块的目的是发挥 2 个分支的优势,该模块可以根据不同的场景来调整两分支融合结果.模块结构如图 5 所示.首先,对两分支结果以相同的编码器 Encoder 进行特征提取;然后,将拼接后的特征通过 ResNet 网络进行特征融合重建;最后,融合特征经过

相同的解码器 Decoder, 得到两分支结果对总结果的贡献权重, 计算两分支加权和即为最终预测结果. 融合模

块充分利用两分支的互补性, 对两分支各取所长, 来提升整体效果.



注: 两输入 CPBR 和 SRBR 分别是变化预测分支和图像超分分支结果. ResNet 结构如图 3(c) 所示.

图 5 融合模块结构图

### 3.5 损失函数

本算法采用联合损失函数优化整体性能, 包括变化预测分支损失, 超分分支损失, 超分模块损失和最终结果损失.

网络最终结果  $\hat{F}_2$  损失  $L_{\text{prediction}}$  如式(6)所示:

$$L_{\text{prediction}} = \text{Loss}(\text{GT}, \hat{F}_2) \quad (6)$$

其中, 本节的 GT (Ground Truth) 均表示预测结果对应的真实值.

变化预测分支 (CPB) 预测结果为 CPBR, 超分分支 (SRB) 预测结果为 SRBR, 两分支损失  $L_{\text{CPB}}$  和  $L_{\text{SRB}}$  如式(7)和式(8)所示:

$$L_{\text{CPB}} = \text{Loss}(\text{GT} \times (1 - \text{mask}), \text{CPBR} \times (1 - \text{mask})) \quad (7)$$

$$L_{\text{SRB}} = \text{Loss}(\text{GT} \times \text{mask}, \text{SRBR} \times \text{mask}) \quad (8)$$

其中, **mask** 是 0-1 矩阵, 分别表示弱变化区域和强变化区域, 依据粗图像的差值  $C_2 - C_1$  进行二值化得到, 让 2 个分支关注不同区域.

双分支超分模块损失  $L_{\text{sr\_CPB}}, L_{\text{sr\_SRB}}$  可表示为  $n$  个超分单元损失的平均, 如式(9)所示:

$$L_{\text{sr\_CPB}}, L_{\text{sr\_SRB}} = \sum_{k=1}^n \frac{1}{n} L_k \quad (9)$$

其中,  $n$  是超分单元个数;  $L_k$  是第  $k$  个超分单元损失, 可表示为  $L_k = \text{Loss}(\text{GT}_k, \hat{F}_C^{\text{SR}_k})$ ,  $\hat{F}_C^{\text{SR}_k}$  和  $\text{GT}_k$  分别为第  $k$  个超分单元的生成图像和真实图像, 真实图像由预测时刻真实值 GT 下采样得到.

本文“融合模块”由  $L_{\text{prediction}}$  反向传播调整其参数; “上分支”由  $L_{\text{prediction}}, L_{\text{CPB}}$  以及  $L_{\text{sr\_CPB}}$  共同作用调整其参数; “下分支”由  $L_{\text{prediction}}, L_{\text{SRB}}$  以及  $L_{\text{sr\_SRB}}$  共同作用调整其参数; 其中上下分支的超分模块同步调整. 上下分支在最终结果损失  $L_{\text{prediction}}$  的基础上, 增加各自关注区域的分支损失来平衡两分支, 让两分支各有所长.

上面公式中的损失函数 Loss 均由内容损失、光谱损失和结构损失 3 个部分组成, 如式(10)~(13)所示.

其中, 内容损失是预测值与真实值的均方误差, 保证基本质量; 光谱损失用余弦相似度表示; 结构损失利用多尺度结构相似性表示.

$$\text{Loss}(\text{GT}, P) = L_{\text{content}}(\text{GT}, P) + L_{\text{spectrum}}(\text{GT}, P) + L_{\text{structure}}(\text{GT}, P) \quad (10)$$

$$L_{\text{content}}(\text{GT}, P) = \frac{1}{N} \sum_{i=1}^N [\text{GT} - P]^2 \quad (11)$$

$$L_{\text{spectrum}}(\text{GT}, P) = I - \frac{\text{GT} \cdot P}{\|\text{GT}\|_2 \|P\|_2} \quad (12)$$

$$L_{\text{structure}}(\text{GT}, P) = I - [l_m(\text{GT}, P)]^{\alpha_M} \cdot \prod_{i=1}^M [c_i(\text{GT}, P)]^{\beta_i} [s_i(\text{GT}, P)]^{\gamma_i} \quad (13)$$

其中, GT 为真实值,  $P$  表示预测图像,  $N$  为训练样本个数,  $I$  表示值为 1 的张量,  $\|\cdot\|_2$  表示 2 范式,  $l_i, c_i, s_i$  分别是第  $i$  个尺度下图像对的亮度相似度、对比度和结构相似度,  $\alpha_M, \beta_i, \gamma_i$  为对应的权重系数.

## 4 实验

本节对提出的 SSRDB 方法在 3 个数据集上进行定性和定量的对比实验分析, 然后进行消融实验验证递归超分结构、非线性补偿模块以及残差结构的有效性, 最后对本文算法的计算负载进行对比评估.

### 4.1 数据集和训练细节

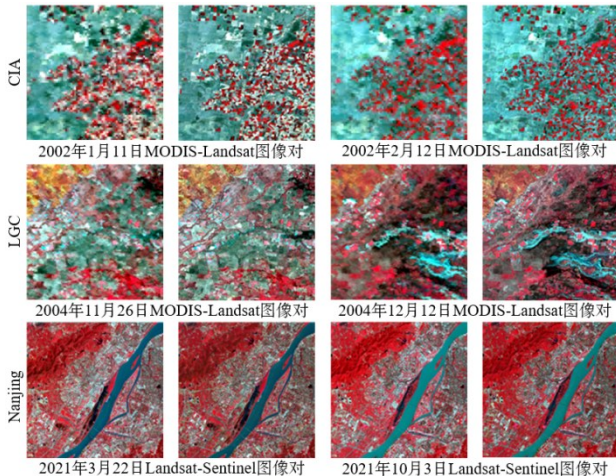
本文选择了 3 个不同场景和不同分辨率差异的数据集, 验证本文 SSRDB 算法对各种场景的适应性和尺度差异上的灵活性.

(1) CIA 数据集. 位于澳大利亚新南威尔士州南部 (34.003 4°E, 145.067 5°S), 使用 2001—2002 年 17 对 MODIS-Landsat 图像, 每张图像 6 个波段, 图像大小为 1 720 × 2 040, 粗细图像之间尺度差异 16 倍. 该场景有一些小面积种植区, 异质区域和物候变化较为显著. 本文实验展示 2002 年 2 月 12 日的 Landsat 图像预测结果,

该组数据如图 6 第 1 行所示。

(2)LGC 数据集<sup>[27]</sup>. 位于澳大利亚新南威尔士州北部(149.281 5°E, 29.085 5°S), 使用 2004—2005 年的 14 对 Landsat-MODIS 图像, 每张图像 6 个波段, 图像大小为 3 200 × 2 720, 粗细图像之间尺度差异 16 倍. 该场景为流域区域, 2004 年 12 月中旬发生了洪水事件, 土地覆盖类型变化显著. 本文实验展示 2004 年 12 月 12 日的 Landsat 图像预测结果, 该组数据如图 6 第 2 行所示。

(3)Nanjing 数据集<sup>[28]</sup>. 位于中国南京(118.803 611°E, 32.075 833°N), 采用 2017—2021 年的 14 对 Landsat8-Sentinel2 图像, 图像取卫星的前 4 个波段, 图像大小为 10 800 × 10 800, 粗细图像之间尺度差异 4 倍. 该场景包含各种人造建筑和自然场景, 有大量的细节和纹理. 本文实验展示 2021 年 10 月 3 日的 Sentinel2 图像预测结果, 该组数据如图 6 第 3 行所示。



注: 图中展示的为波段 4、波段 3 和波段 2 作为 RGB 的假彩色合成图像。  
图 6 数据集展示

CIA 和 LGC 数据集粗细图像空间分辨率差异约 16 倍, 本文算法在 CIA 和 LGC 数据集上采用 4 个超分单元. CIA 数据集 Landsat 图像裁剪为 1 280 × 1 280, MODIS 图像裁剪为 80 × 80; LGC 数据集 Landsat 图像裁剪为 2 400 × 2 400, MODIS 图像裁剪为 150 × 150. 本文算法的超分单元是 2 倍率超分, 而 Nanjing 数据集 Sentinel2 卫星和 Landsat8 卫星尺度倍率约 3 倍, 通过对 Nanjing 数据集中的 Landsat 进一步采样, 将倍率规范为 4 倍, 采用 2 个超分单元. Nanjing 数据集中 Sentinel 图像裁剪为 2 400 × 2 400, Landsat 图像裁剪为 600 × 600. 本文递归超分的特征通道数为 32, 超分单元中“特征重构”的 ResNet 级联了 8 个残差块, “残差图像重构”的 ResNet 级联了 2 个残差块. “非线性补偿模块”和“融合模块”中的 ResNet 级联了 4 个残差块。

模型训练时, 学习率设置为  $2 \times 10^{-4}$ , 批量大小为 8,

模型迭代次数为 200. 本文实验的对比算法中, CIA 数据集中 Landsat 与 MODIS 图像尺寸均为 1 280 × 1 280, 块大小为 128 × 128; LGC 和 Nanjing 数据集中 MODIS、Landsat、Sentinel 图像尺寸均为 2 400 × 2 400, 块大小为 256 × 256, 对比算法中的其他参数使用作者提供的原始代码参数。

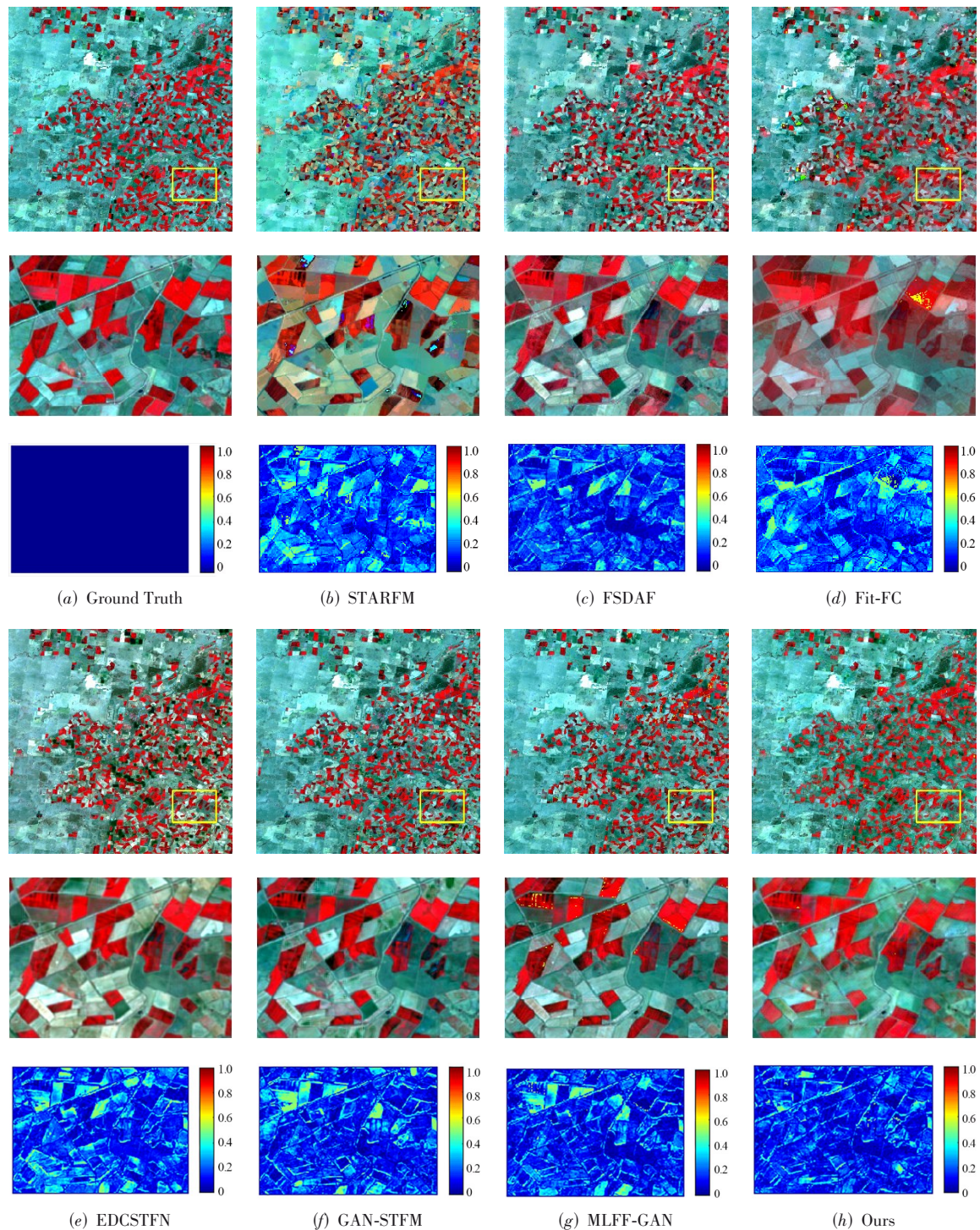
## 4.2 对比方法和评价指标

为了验证本文算法的优越性, 采用 6 种优秀且经典的时空融合算法进行比较, 包括 3 种传统方法 (STARFM<sup>[5]</sup>、FSDAF<sup>[12]</sup>、Fit-FC<sup>[13]</sup>) 和 3 种基于深度学习的融合算法 (EDCSTFN<sup>[19]</sup>、GAN-STFM<sup>[22]</sup>、MLFF-GAN<sup>[24]</sup>). 为了对实验结果进行客观全面的评价, 采用均方根误差 RMSE (Root Mean Square Error)、光谱角映射器 SAM (Spectral Angel Mapping)<sup>[29]</sup>、结构相似度 SSIM (Structural Similarity)<sup>[30]</sup> 和相关系数 CC (Correlation Coefficients) 4 个常用指标定量评估融合性能. 其中, RMSE 和 CC 衡量预测值与真实值间的总体差异, RMSE 理想值为 0, CC 理想值为 1. SAM 计算两光谱夹角, 值范围在 0~1, 理想值为 0. SSIM 从亮度、对比度和结构 3 个方面衡量轮廓纹理相似度, 理想值为 1.

## 4.3 实验结果与分析

图 7 展示了本文算法和对比算法在 CIA 数据集上的融合结果. 从各类算法差值图呈现蓝色主色调可以看出, 各个算法总体上效果都不错; 从其平坦区域的青黄色, 可以看出各类算法在光谱信息上的性能差异, 从细节区域呈现的青黄色可以看出各类算法在空间信息上的性能差异, 总的来说前 3 种传统方法在光谱和空间表征性能上要弱于深度学习的方法, 这也是复杂的模型、大量的样本和高的计算成本带来的收益, 下面 2 个数据集也有相同的体现. 3 个传统算法中, FSDAF 相对效果较好, 从放大图和差值图可以看出 STARFM 和 Fit-FC 在光谱和空间都有较大误差. 3 个深度学习方法, EDCSTFN 模型相对简单一些, 其效果略次于 GAN-STFM 和 MLFF-GAN, 本文算法表现最好, 定量指标如表 1 所示, 在 RMSE、SAM、SSIM、CC 指标上相对于次优方法分别提高了 7.067%、2.065%、0.563%、0.581%, 表明本文算法较好适用于异质区域和物候变化场景。

图 8 展示各类算法在 LGC 数据集上的融合结果. 放大区域以左下右上做对角线, 左上方主要体现为物候变化, 右下方体现为洪水带来的地物变化. 由于物候变化是大面积形式存在, 各类算法重构效果都很好. 而对于右下变化区域, MLFF-GAN 和本文算法在洪水变化区域处理明显优于其他算法. 这是因为 MLFF-GAN 算法设计了注意力模块 AM 专门服务于场景变化, 本文算法针对变化区域和弱变化区域做了相应的策略. 表 1 定量指标表明, 本



注: 每组图的第1行为整体预测结果, 展示的是波段4、3、2作为R、G、B的伪彩色图像; 第2行为第1行中黄色矩形区域的放大, 第3行为第2行预测图像与真实值之间的差值图. 图8和图9与之相同.

图7 2002年2月12日CIA数据集各方法融合结果

文算法略优于MLFF-GAN, 在RMSE、SAM、SSIM、CC上分别优于次优算法5.319%、5.490%、1.455%、1.086%, 表明本文方法对于类型变化场景适应性好.

图9展示各类算法在Nanjing数据集上的融合结果.

该场景细节特别丰富, 异质性非常强, 观察放大图和差值图可以看出, 深度学习类方法体现出了绝对优势, 这可能是由于异质性强, 不管是线性加权还是线性解混, 都容易受周边像素影响. 本文算法采用了带参考的递归

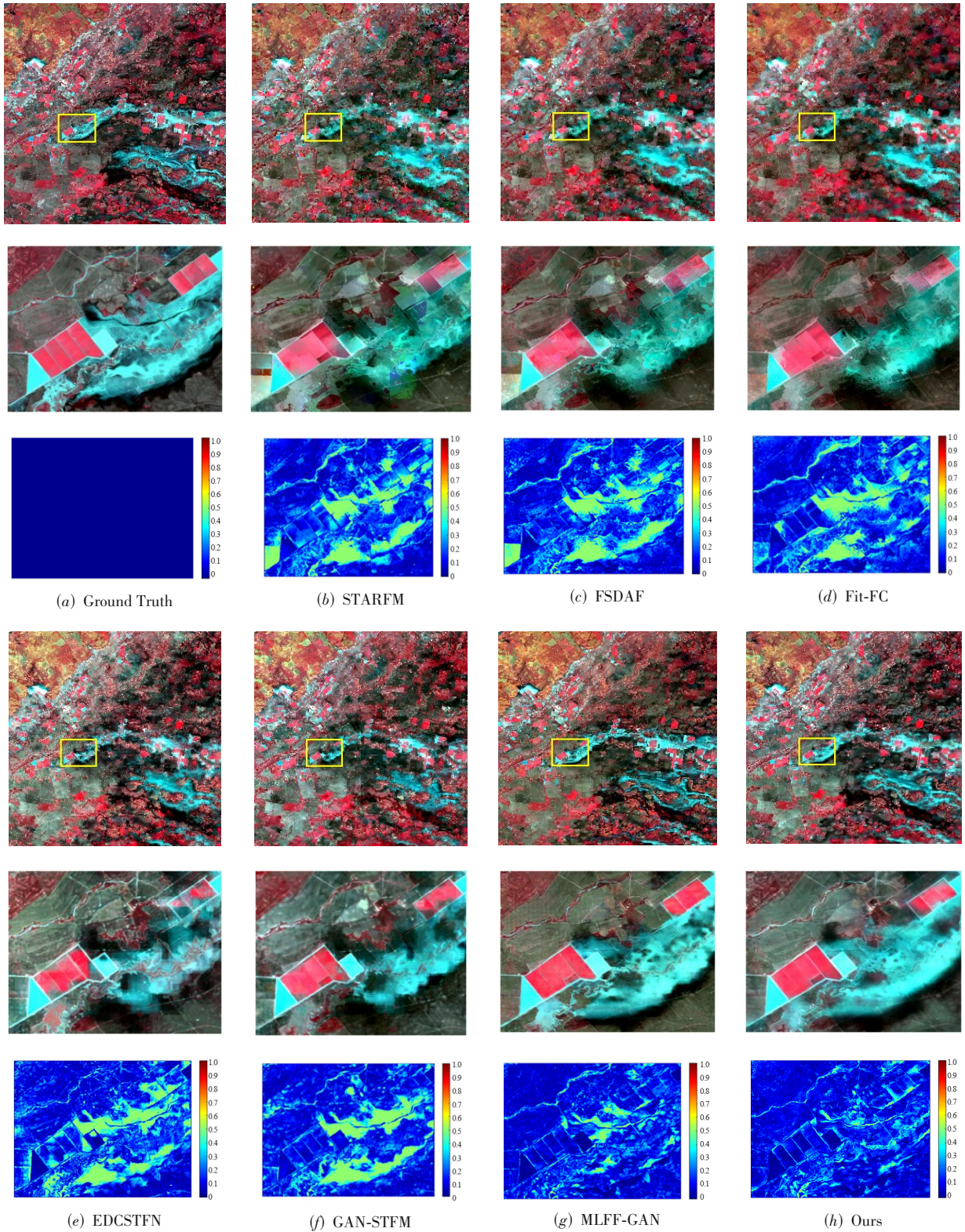


图8 2004年12月12日LGC数据集各方法融合结果

超分,细节重建能力强,该场景内容没有发生大的变化,所以,参考图以及变化预测分支对于细节重建效果都很好.如表1所示,本文算法在RMSE、SAM、SSIM、CC上分

别优于次优算法6.486%、16.290%、0.481%、0.257%.实验结果表明本文算法可以很好适应对细节要求较高的场景,在细节预测和光谱保持上具有优势.

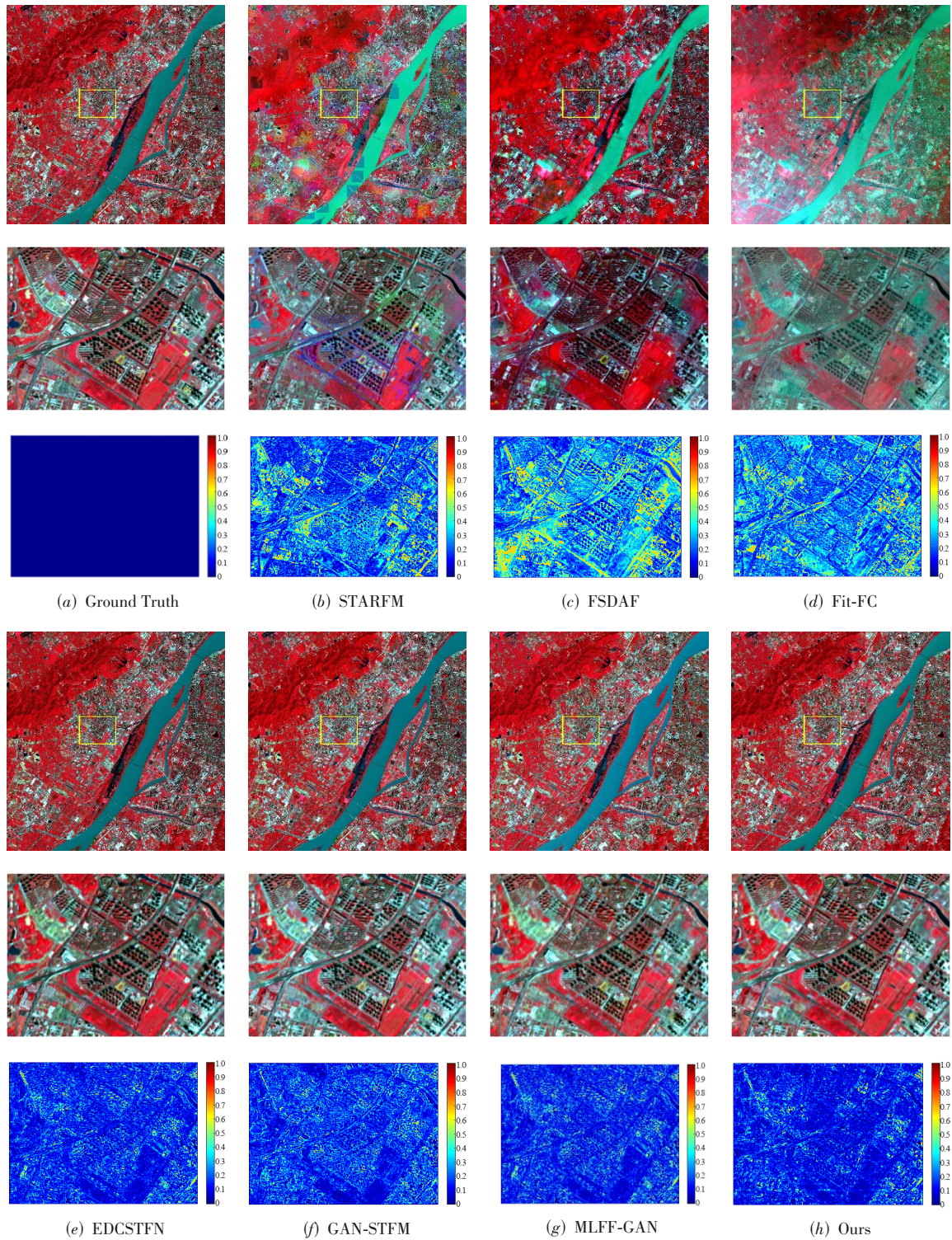


图9 2021年10月3日Nanjing数据集各方法融合结果

#### 4.4 消融实验

在本节进行消融实验,验证模型中递归超分结构、非线性补偿模块以及残差结构的有效性.

为了验证递归超分结构的有效性,修改网络结构,

去除多次的2倍倍率超分单元,仅通过一次超分单元得到预测结果,同时增加此超分单元的复杂度,保证模型正常运行,避免性能显著下降.图10以LGC数据集消融实验结果为例,可直观看出,具有递归超分的网络在

表 1 各方法在 3 个数据集上的定量评估结果(最好的用粗体标注)

数据集	指标	STARFM	FSDAF	Fit-FC	EDCSTFN	GAN-STFM	MLFF-GAN	Ours
CIA	RMSE ↓	0.034 8	0.032 2	0.033 0	0.031 9	0.030 8	0.028 3	<b>0.026 3</b>
	SAM ↓	0.095 7	0.088 2	0.098 0	0.087 2	0.078 5	0.077 5	<b>0.075 9</b>
	SSIM ↑	0.818 2	0.840 2	0.839 5	0.858 5	0.862 8	0.871 0	<b>0.875 9</b>
	CC ↑	0.946 9	0.954 0	0.951 0	0.958 0	0.958 0	0.964 1	<b>0.969 7</b>
LGC	RMSE ↓	0.040 9	0.040 2	0.038 8	0.033 2	0.033 1	0.028 2	<b>0.026 7</b>
	SAM ↓	0.209 7	0.211 6	0.187 9	0.161 2	0.156 0	0.134 8	<b>0.127 4</b>
	SSIM ↑	0.757 0	0.759 5	0.781 1	0.809 2	0.813 3	0.831 9	<b>0.844 0</b>
	CC ↑	0.820 7	0.829 4	0.835 1	0.882 0	0.880 1	0.911 8	<b>0.921 7</b>
Nanjing	RMSE ↓	0.064 2	0.069 5	0.129 4	0.020 0	0.018 5	0.022 7	<b>0.017 3</b>
	SAM ↓	0.321 8	0.343 9	0.222 8	0.070 8	0.066 3	0.082 1	<b>0.055 5</b>
	SSIM ↑	0.432 3	0.548 2	0.602 8	0.898 3	0.914 5	0.874 1	<b>0.918 9</b>
	CC ↑	0.895 2	0.894 8	0.785 6	0.966 7	0.972 2	0.955 8	<b>0.974 7</b>

光谱和水体结构预测上都具有更好的表现. 消融实验定量指标如表 2 所示,从表中可看出,递归超分网络预测结果指标显著优于非递归超分网络预测结果指标. 网络通过多次的 2 倍递归超分,逐步地进行特征与图像的超分重建,加之对中间超分结果图像进行监督,可以更好地恢复细节,保持图像光谱特性,有利于模型预测的精度.

表 2 不同数据集递归超分消融实验的评估结果(最好的用粗体标注)

模块消融	数据集	RMSE ↓	SAM ↓	SSIM ↑	CC ↑
递归超分	CIA	<b>0.026 3</b>	<b>0.075 9</b>	<b>0.875 9</b>	<b>0.969 7</b>
	LGC	<b>0.026 7</b>	<b>0.127 4</b>	<b>0.844 0</b>	<b>0.921 7</b>
	Nanjing	<b>0.017 3</b>	<b>0.055 5</b>	<b>0.918 9</b>	<b>0.974 7</b>
非递归超分	CIA	0.028 7	0.088 7	0.872 1	0.963 3
	LGC	0.030 7	0.150 5	0.821 4	0.900 5
	Nanjing	0.020 6	0.072 4	0.890 5	0.964 6

为了验证非线性补偿模块的有效性,图 11 给出了 LGC 数据集“有/无”非线性补偿模块的融合结果,可以看出图 11(b)对红色块的边缘预测更清晰;表 3 的量化指标也可以看出,具有非线性补偿模块的预测结果指标更好. 该模块可以使得非线性情况下,“差分图像超分”和“图像超分”共享超分模型成立,从而简化模型,保证精度.

为了验证本文残差结构的有效性,将“有/无全局”

表 3 不同数据集非线性补偿消融实验的评估结果(最好的用粗体标注)

模块消融	数据集	RMSE ↓	SAM ↓	SSIM ↑	CC ↑
有非线性补偿	CIA	<b>0.026 3</b>	<b>0.075 9</b>	<b>0.875 9</b>	<b>0.969 7</b>
	LGC	<b>0.026 7</b>	<b>0.127 4</b>	<b>0.844 0</b>	<b>0.921 7</b>
	Nanjing	<b>0.017 3</b>	<b>0.055 5</b>	<b>0.918 9</b>	<b>0.974 7</b>
无非线性补偿	CIA	0.029 7	0.091 4	0.866 0	0.961 0
	LGC	0.028 0	0.139 5	0.832 0	0.915 2
	Nanjing	0.021 0	0.073 5	0.886 9	0.962 8

残差连接的残差结构进行对比,即图 3(c)“有/无”跳跃连接,图 12 给出了 LGC 数据集的融合结果,图 12(b)在变化大的水体部分的光谱和结构上表现更好;表 4 的量化指标也可以看出,全局残差预测结果在各数据集上的表现都更好. 全局残差连接使细节直接传播到后面层,减少了信息丢失,有利于预测结果.

表 4 不同数据集残差结构消融实验的评估结果(最好的用粗体标注)

模块消融	数据集	RMSE ↓	SAM ↓	SSIM ↑	CC ↑
有全局残差连接	CIA	<b>0.026 3</b>	<b>0.075 9</b>	<b>0.875 9</b>	<b>0.969 7</b>
	LGC	<b>0.026 7</b>	<b>0.127 4</b>	<b>0.844 0</b>	<b>0.921 7</b>
	Nanjing	<b>0.017 3</b>	<b>0.055 5</b>	<b>0.918 9</b>	<b>0.974 7</b>
无全局残差连接	CIA	0.027 0	0.076 6	0.872 3	0.963 6
	LGC	0.027 8	0.133 4	0.838 4	0.914 9
	Nanjing	0.018 4	0.063 0	0.910 1	0.971 5

#### 4.5 计算负载

表 5 展示了本文算法和对比算法的模型参数量(百万)和乘法累加操作次数 MACs(十亿),其中参数量反映了模型空间复杂度,MACs 反映了模型计算复杂度. EDCSTFN 采用简单的编-解码器结构,参数量和 MACs 都较低,但从表 1 可以看出其预测精度较其他深度学习融合算法低;GAN-STFM 的生成器同样采用编-解码结构,由多个残差模块级联而成,参数量和 MACs 都均是 EDCSTFN 的 2 倍. MLFF-GAN 生成器类似 U-Net 结构,在编-解码器之间增加了特征融合模块,以解决传感器之间的信息转换,并关注变化区域,参数量大,但是其多级特征融合操作中包含了许多小尺寸的操作,显著减少了 MACs. 本文算法 SSRDB 虽然是复杂的双分支网络结构,通过共享参数和小尺寸操作策略,其参数量和 MACs 综合表现最好. 首先,上下分支的超分模块和超分模块内部超分单元皆共享参数,其模型参数比 GAN-STFM 略多,远少于 MLFF-GAN. 其

次,模型从小尺寸开始进行逐步预测恢复,即包含了很多小尺寸上的计算,减少计算量,其MACs表现最佳.

表5 不同深度学习时空融合算法的参数数量和MACs

算法	参数量/M	MACs/G
EDCSTFN	0.284	18.585
GAN-STFM	0.583	37.770
MLFF-GAN	8.701	17.369
SSRDB	0.856	16.051

## 5 结论

遥感图像时空融合问题在类型变化区域融合性能较差,本文提出了共享超分的双分支时空融合网络

SSRDB,将时空融合问题解耦成时间维变化预测和空间维超分辨率,简化表达关系,利用两者互补性融合得到最终结果,缓解了类型变化区域性能差的问题.对本文算法有如下总结:首先,分析时空融合问题在时间维和空间维上的互补性,依此构建了双分支网络框架,解耦复杂任务,增加中间监督让每个分支关注自身任务,并通过融合模块使两分支取长补短.其次,两分支共享超分模块且超分模块复用超分单元,大大减少模型参数;渐进式的超分方式,保证了超分精度;复用超分单元可灵活适应不同倍率差异的超分.将本文算法与STARFM、FSDAF、Fit-FC、EDCSTFN、GAN-STFM、MLFF-GAN等方法在3个数据集上进行比较,实验结果展现了SSRDB在光谱保持、细节增强和变化区

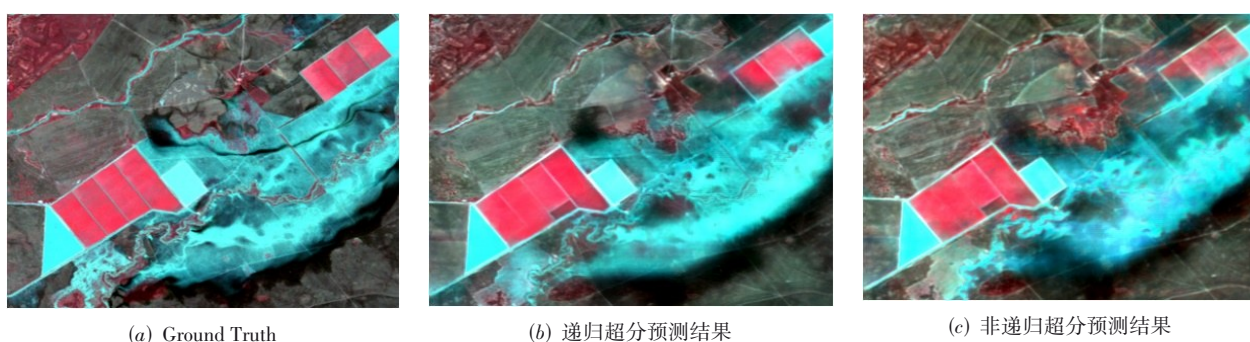


图10 LGC数据集2004年12月12日部分区域真实值和递归超分消融实验预测结果

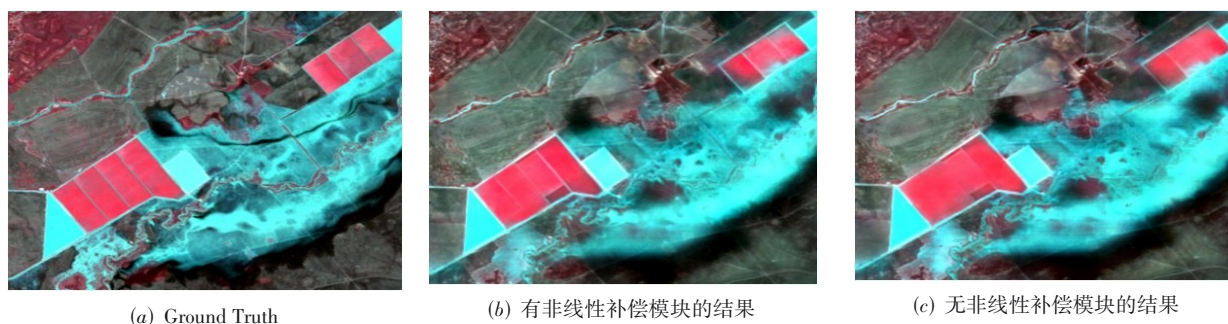


图11 LGC数据集2004年12月12日部分区域真实值和非线性补偿消融实验预测结果

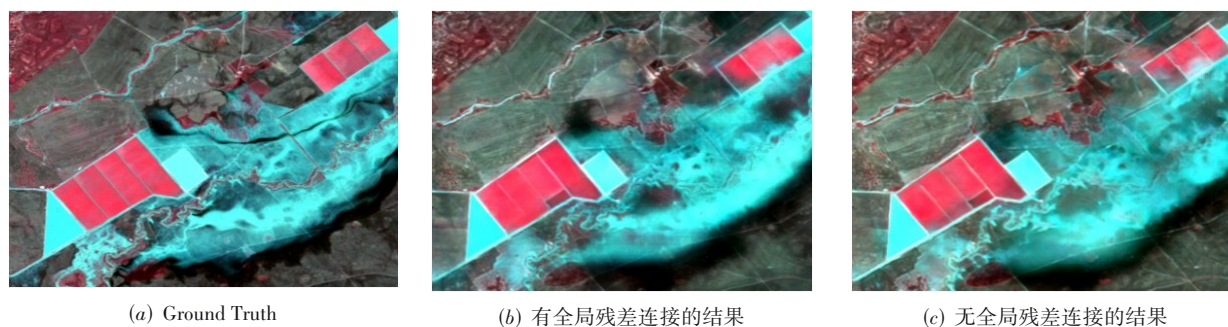


图12 LGC数据集2004年12月12日部分区域真实值和残差结构消融实验预测结果

域优秀的性能,同时在消融实验中证明了 SSRDB 中递归超分结构、非线性补偿模块以及残差结构的有效性。

由于时空融合模型暗含了传感器之间的映射关系,这种关系具有设备依赖性,从而导致面对不同应用场景,模型需要重新训练,未来工作将关注时空融合模型的泛化能力。目前想法是将本文的超分模块独立出来,利用不同卫星的大量数据进行预训练,作为重要的预训练好的组件,提供给不同时空融合策略使用。面对不同应用场景时,超分部分固定,利用少量的样本,学习设备依赖的关系,快速建立时空融合模型。

#### 参考文献

- [1] ZHANG Y. A time-series approach to detect urbanized areas using biophysical indicators and landsat satellite imagery[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2021, 14: 9210-9222.
- [2] ARÉVALO P, OLOFSSON P, WOODCOCK C E. Continuous monitoring of land change activities and post-disturbance dynamics from landsat time series: A test methodology for REDD+ reporting[J]. *Remote Sensing of Environment*, 2020, 238: 111051.
- [3] YANG Y, ANDERSON M C, GAO F, et al. Studying drought-induced forest mortality using high spatiotemporal resolution evapotranspiration data from thermal satellite imaging[J]. *Remote Sensing of Environment*, 2021, 265: 112640.
- [4] LUO X, WANG M C, DAI G M, et al. A novel technique to compute the revisit time of satellites and its application in remote sensing satellite optimization design[J]. *International Journal of Aerospace Engineering*, 2017, 2017: 6469439.
- [5] GAO F, MASEK J, SCHWALLER M, et al. On the blending of the landsat and MODIS surface reflectance: Predicting daily landsat surface reflectance[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2006, 44(8): 2207-2218.
- [6] HILKER T, WULDER M A, COOPS N C, et al. A new data fusion model for high spatial- and temporal-resolution mapping of forest disturbance based on Landsat and MODIS[J]. *Remote Sensing of Environment*, 2009, 113(8): 1613-1627.
- [7] ZHU X L, CHEN J, GAO F, et al. An enhanced spatial and temporal adaptive reflectance fusion model for complex heterogeneous regions[J]. *Remote Sensing of Environment*, 2010, 114(11): 2610-2623.
- [8] FU D J, CHEN B Z, WANG J, et al. An improved image fusion approach based on enhanced spatial and temporal the adaptive reflectance fusion model[J]. *Remote Sensing*, 2013, 5(12): 6346-6360.
- [9] ZHUKOV B, OERTEL D, LANZL F, et al. Unmixing-based multi-sensor multiresolution image fusion[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 1999, 37(3): 1212-1226.
- [10] WU M Q, NIU Z, WANG C Y, et al. Use of MODIS and Landsat time series data to generate high-resolution temporal synthetic Landsat data using a spatial and temporal reflectance fusion model[J]. *Journal of Applied Remote Sensing*, 2012, 6(1): 063507.
- [11] ZHANG W, LI A N, JIN H A, et al. An enhanced spatial and temporal data fusion model for fusing landsat and MODIS surface reflectance to generate high temporal landsat-like data[J]. *Remote Sensing*, 2013, 5(10): 5346-5368.
- [12] ZHU X L, HELMER E H, GAO F, et al. A flexible spatiotemporal method for fusing satellite images with different resolutions[J]. *Remote Sensing of Environment*, 2016, 172: 165-177.
- [13] WANG Q M, ATKINSON P M. Spatio-temporal fusion for daily Sentinel-2 images[J]. *Remote Sensing of Environment*, 2018, 204: 31-42.
- [14] HUANG B, SONG H H. Spatiotemporal reflectance fusion via sparse representation[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2012, 50(10): 3707-3716.
- [15] WU B, HUANG B, ZHANG L P. An error-bound-regularized sparse coding for spatiotemporal reflectance fusion[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2015, 53(12): 6791-6803.
- [16] WEI J B, WANG L Z, LIU P, et al. Spatiotemporal fusion of remote sensing images with structural sparsity and semi-coupled dictionary learning[J]. *Remote Sensing*, 2017, 9(1): 21.
- [17] SONG H H, LIU Q S, WANG G J, et al. Spatiotemporal satellite image fusion using deep convolutional neural networks[J]. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 2018, 11(3): 821-829.
- [18] TAN Z Y, YUE P, DI L P, et al. Deriving high spatiotemporal remote sensing images using deep convolutional network[J]. *Remote Sensing*, 2018, 10(7): 1066.
- [19] TAN Z Y, DI L P, ZHANG M D, et al. An enhanced deep convolutional model for spatiotemporal image fusion[J]. *Remote Sensing*, 2019, 11(24): 2898.
- [20] ZHANG H Y, SONG Y Y, HAN C, et al. Remote sensing image spatiotemporal fusion using a generative adversarial network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 59(5): 4273-4286.
- [21] LEDIG C, THEIS L, HUSZÁR F, et al. Photo-realistic single image super-resolution using a generative adversarial network[C]// 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). Piscataway: IEEE, 2017: 105-114.
- [22] TAN Z Y, GAO M L, LI X H, et al. A flexible reference-insensitive spatiotemporal fusion model for remote sensing images using conditional generative adversarial network[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2021, 60: 5601413.
- [23] LIU X, DENG C W, CHANUSSOT J, et al. StfNet: A two-stream convolutional neural network for spatiotemporal image fusion[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2019, 57(9): 6552-6564.
- [24] SONG B Z, LIU P, LI J, et al. MLFF-GAN: A multilevel feature fusion with GAN for spatiotemporal remote sensing images[J]. *IEEE Transactions on Geoscience and Remote Sensing*, 2022, 60: 4410816.

- [25] SONG Y Y, ZHANG H Y, HUANG H, et al. Remote sensing image spatiotemporal fusion via a generative adversarial network with one prior image pair[J]. IEEE Transactions on Geoscience and Remote Sensing, 2022, 60: 5528117.
- [26] YANG G Q, QIAN Y R, LIU H, et al. MSFusion: Multistage for remote sensing image spatiotemporal fusion based on texture transformer and convolutional neural network[J]. IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing, 2022, 15: 4653-4666.
- [27] EMELYANOVA I V, MCVICAR T R, VAN NIEL T G, et al. Assessing the accuracy of blending landsat-MODIS surface reflectances in two landscapes with contrasting spatial and temporal dynamics: A framework for algorithm selection[J]. Remote Sensing of Environment, 2013, 133: 193-209.
- [28] WANG Z Y, FANG S, ZHANG J. Spatiotemporal fusion model of remote sensing images combining single-band and multi-band prediction[J]. Remote Sensing, 2023, 15(20): 4936.
- [29] YUHAS R H, GOETZ A F H, BOARDMAN J W. Discrimination among semi-arid landscape endmembers using the spectral angle mapper (SAM) algorithm[C]//Summaries of the 3rd Annual JPL Airborne Geoscience Workshop. Pasadena: JPL, 1992: 147-149.
- [30] 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, 13(4): 600-612.

### 作者简介



方 帅 女,1978年1月出生于安徽省淮南市.现为合肥工业大学计算机与信息学院教授、博士生导师.主要研究方向为图像复原和视觉检测.

E-mail: fangshuai@hfut.edu.cn



张 晶 女,1976年10月出生于安徽省合肥市.现为合肥工业大学计算机与信息学院副教授.主要研究方向为人工智能、数据挖掘.

E-mail: zhangjing@hfut.edu.cn



张小溪 女,2001年12月出生于河南省周口市.现为合肥工业大学计算机与信息学院硕士研究生.主要研究方向为图像复原和视觉检测.

E-mail: 2022171263@mail.hfut.edu.cn