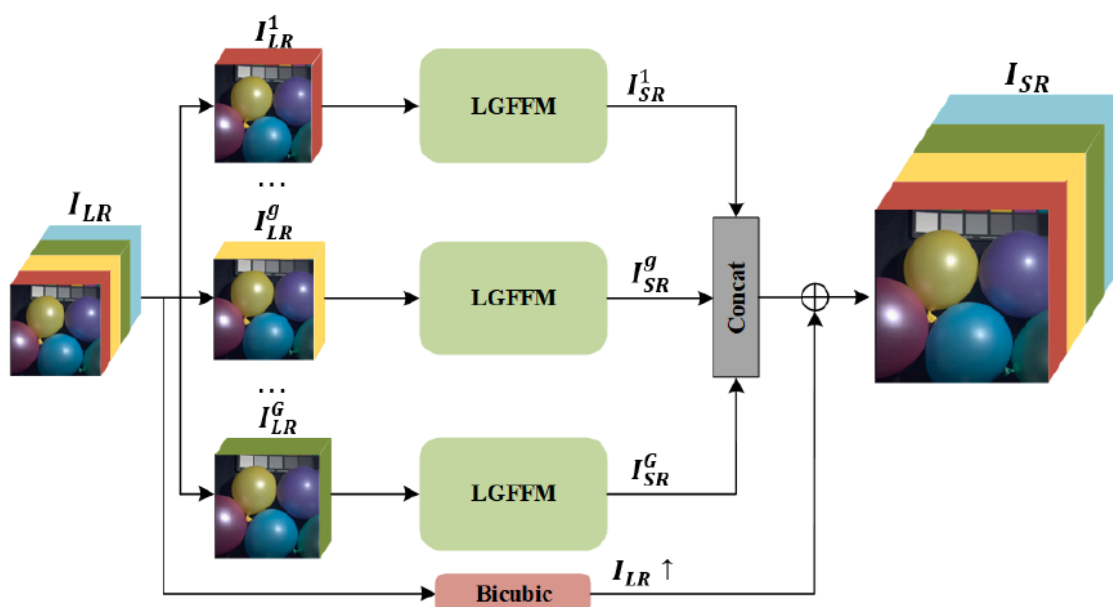


用于高光谱图像超分辨率的局部-全局特征融合网络

摘要

近年来，许多高光谱图像超分辨率方法已经取得了非凡的成就。然而，在深层网络巨大的计算和资源的消耗在实际应用中并不适用。因此，一个轻量化且高效的局部-全局特征融合网络

(LGFFMN) 被提出用于减少计算负担同时实现超分辨率重构性能。特别的，提出了一个局部特征融合模块 (LFFM)，利用卷积神经网络 (CNN) 进行有效的局部特征提取。与此同时，为了探索全局信息和避免因传统的Transformer带来的计算负担，采用基于多尺度表示的特征调制机制，构建了一种视觉变换器的全局特征融合模块 (GFFM)。为了在利用相邻波段之间具有较高的相似性的同时降低HSI的高维性质，该网络采用分组策略。



👉 图1 所提出的LGFFMN模型，在模型前期讲HSI的整体数据划分为 G 个组。

局部特征融合模块

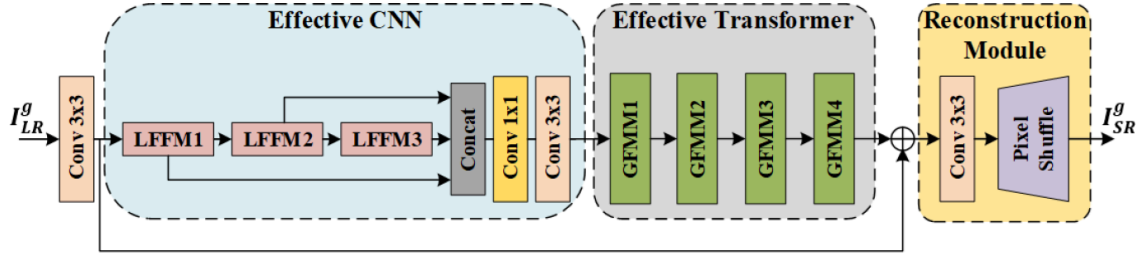


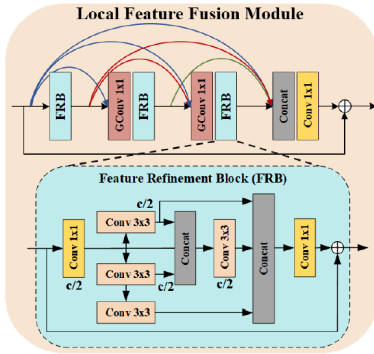
图2 LGFFM网络结构图，主要包含有效的CNN网络，有效的Transformer模块以及重建模块

图2展示了所提出的LFFM模块。首先，浅层特征提取 F_{sf}^g 被送入LFFMs中，具体表示为：

$$F_{LFFM}^1 = f_{LFFM}^1(F_{sf}^g), i = 1$$

$$F_{LFFM}^i = f_{LFFM}^i(F_{LFFM}^{i-1}), i = 2, \dots, n$$

其中 $f_{LFFM}^i(\cdot)$ 表示第 i 个LFFM模块。之后输出 F_{LFFM}^i 通过拼接操作将所有的LFFMs模块进行拼接。最终，采用两个 1×1 和 3×3 的卷积层实现特征融合和压缩得到局部特征 F_{CNN}^g 。



LFFM的网络结构如图3所示。该模块在DenseBlock的基础上做了几点改进：

- (1) 使用FRB模块代替卷积层来增强模型的特征提取能力。
- (2) 在每个FRB之前使用 1×1 组卷积来降低特征维度。
- (3) 在FRB内部采用局部残差学习来进一步进行信息的传递。

以上三点可公式化为：

$$\begin{aligned} F_{FRB}^1 &= f_{FRB}^1(F_{LFFM}^1) \\ F_{FRB}^2 &= f_{FRB}^2(f_{gc}^1[F_{LFFM}^{m-1}, F_{FRB}^1]) \\ F_{FRB}^3 &= f_{FRB}^3(f_{gc}^2[F_{LFFM}^{m-1}, F_{FRB}^1, F_{FRB}^2]) \\ F_{LFFM}^m &= F_{LFFM}^{m-1} + f_{1 \times 1}([F_{LFFM}^{m-1}, F_{FRB}^1, F_{FRB}^2, F_{FRB}^3]) \end{aligned}$$

其中 F_{FRB}^i 表示输出的第 i ($i = 1, 2, 3$)个FRB模块。第 j ($j = 1, 2$)个分组卷积表示为 $f_{gc}^j(\cdot)$ 。输入和输出第 m 个LFFM模块表示为 F_{LFFM}^{m-1} 和 F_{LFFM}^m 。

特征细化模块（FRB）：

首先，通过将特征送入多分支网络中，在不同分支上使用不同的卷积层来改变感受野的大小来过的不同尺度的特征。 $c/2$ 表示通过卷积运算将输入特征减半，以压缩特征维度并减少网络参数。最后，将不同感受野的特征进行拼接，重构细节特征。



图3 LFFM和FRB的网络结构

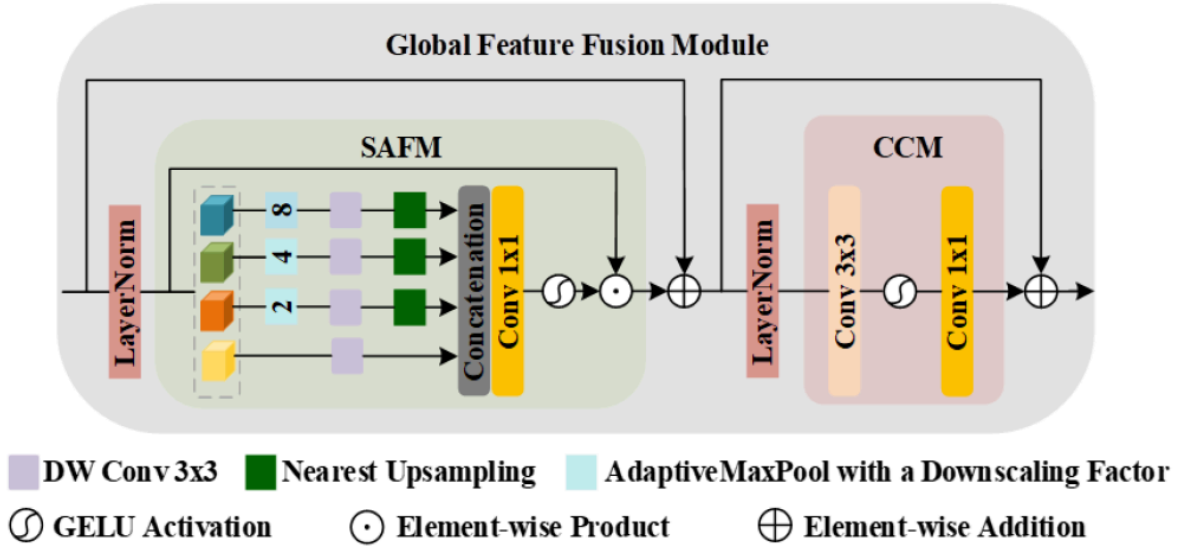


图4 GFFM网络架构，主要包含SAFM和CCM模块

全局特征融合模块

全局特征融合模块如图4所示。主要包含一个空间自适应特征调制（SAFM）和一个通道混合卷积（CCM）模块。SAFM模块引入一种多尺度表示的特征调制来动态进行特征提取，获得具有长程依赖关系的全局特征。CCM是一种用于特征细化的前馈网络，通过编码特征和混合通道来增强特征的表达能力。GFFM网络公式化为：

$$Y = SAFM(LN(X)) + X$$

$$Z = CCM(LN(Y)) + Y$$

其中 $LN(\cdot)$ 表示归一化层，且 X, Y, Z 表示网络中的中间特征。

空间自适应特征调制（SAFM）

SAFM采用空间自适应调制机制来获得多尺度特征表达依赖长程关系，这样能够解决在Transformer中自注意力机制的计算负担。在图4中，SAFM使用一个特征金字塔模块来获得注意力映射图。首先，归一化后的输入特征被通道分离来获得4个分量，之后第一个分量通过 3×3 的深度卷积进行处理，且剩余部分被送入多尺度特征单元中来获得金字塔特征。此过程可表示为：

$$[X_0, X_1, X_2, X_3] = Split(X)$$

$$\hat{X}_0 = DWConv_{3 \times 3}(X_0)$$

$$\hat{X}_i = \uparrow_p (DWConv_{3 \times 3}(\downarrow_{\frac{p}{2^i}})(X_i))$$

其中 $Split(\cdot)$ 表示通道分离操作，且 $DWConv_{3 \times 3}(\cdot)$ 表示深度卷积操作。通过最近邻插值上采样得到原始分辨率 p 表示为 $\uparrow_p(\cdot)$ 。 $\downarrow(\frac{p}{2^i})$ 表示将输入下采样到 $(\frac{p}{2^i})$ 。有区别的特征通过学习非局部的依赖关系，并对输入应用自适应最大值合并用于生成多尺度特征。在拼接这些多尺度特征后，通过 1×1 卷积来聚合局部和全局关系。此过程表示为：

$$\hat{X} = Conv_{1 \times 1}(Concat([\hat{X}_0, \hat{X}_1, \hat{X}_2, \hat{X}_3]))$$

其中 $Concat(\cdot)$ 表示在通道维度上的拼接操作。之后使用GELU非线性激活函数来估计获得的精确的特征 \hat{X} 。最终，估计的注意力映射图通过与输入 X 进行元素乘积。

$$\bar{X} = \Phi(\hat{X}) \cdot X$$

损失函数

L_1 损失可以惩罚较小的误差并且在保存空间信息方面表现良好。为了保证光谱的保真度，在本文中引入SAM损失。因此，采用 L_1 损失和SAM损失相结合能够同时减小数值误差和光谱失真，具体表示为：

$$L_{joint}(\Theta) = L_1 + \lambda_s L_s$$

其中 λ_s 表示平衡因子，在本文中我们设置为 $\lambda_s=0.5$ 。