基于 CNN-Transformer 混合构架的轻量图像超分辨率方法*

林承浩，吴丽君
（福州大学　物理与信息工程学院，福建　福州　350108）

摘　要：针对基于混合构架的图像超分模型通常需要较高计算成本的问题，提出了一种基于 CNN-Transformer 混合构架的轻量图像超分网络 STSR（Swin-Transformer-based Single Image Super-Resolution）。首先，提出了一种并行特征提取的特征增强模块（Feature Enhancement Block，FEB），由卷积神经网络（Convolutional Neural Network，CNN）和轻量型 Transformer 网络并行地对输入图像进行特征提取，再将提取到的特征进行特征融合。其次，设计了一种动态调整模块（Dynamic Adjustment，DA），使得网络能根据输入图像来动态调整网络的输出，减少网络对无关信息的依赖。最后，采用基准数据集来测试网络的性能，实验结果表明 STSR 在降低模型参数量的前提下仍然保持较好的重建效果。

关键词：图像超分辨率；轻量化；卷积神经网络；Transformer
中图分类号：TP391　　**文献标识码：**A　　**DOI：**10.19358/j.issn.2097-1788.2024.03.005
引用格式：林承浩，吴丽君. 基于 CNN-Transformer 混合构架的轻量图像超分辨率方法［J］. 网络安全与数据治理，2024，43(3)：27-33.

A lightweight image super-resolution method based on a hybrid CNN-Transformer architecture

Lin Chenghao，Wu Lijun
（School of Physics and Information Engineering，Fuzhou University，Fuzhou 350108，China）

Abstract：In order to address the problem that image super-segmentation models based on hybrid architectures usually require high computational cost，this study proposes a lightweight image super-segmentation network STSR（Swin-Transformer-based Single Image Super-Resolution）based on a hybrid CNN-Transformer architecture. Firstly，this paper proposes a Feature Enhancement Block（FEB）for parallel feature extraction，which consists of a Convolutional Neural Network（CNN）and a lightweight Transformer Network to extract features from the input image in parallel，and then the extracted features are fused to the features. Secondly，this paper designs a Dynamic Adjustment（DA）module，which enables the network to dynamically adjust the output of the network according to the input image，reducing the network's dependence on irrelevant information. Finally，some benchmark datasets are used to test the performance of the network，and the experimental results show that STSR still maintains a better reconstruction effect under the premise of reducing the number of model parameters.

Key words：image super-resolution；lightweighting；Convolutional Neural Network；Transformer

0　引言

图像超分辨率（Super Resolution，SR）是一项被广泛关注的计算机视觉任务，其目的是从低分辨率（Low Resolution，LR）图像中重建出高质量的高分辨率（High Resolution，HR）图像[1]。由于建出高质量的高分辨率图像具有不适定的性质，因此极具挑战性[2]。随着深度学习等新兴技术的崛起，许多基于卷积神经网络（CNN）的方法被引入到图像超分任务中[3-6]。SRCNN[3]首次将卷积神经网络引入到图像超分任务中，用卷积神经网络来学习图像的特征表示，并通过卷积层的堆叠来逐步提取更高级别的特征，使得重建出的图像具有较高的质量。在后续研究中，Kaiming He 等人提出了残差结构 ResNet[5]，通过引入跳跃连接，允许梯度能够跨越层进行传播，有助于减轻梯度消失的问题，使得模型在较深的网络情况下仍然能保持较好的性能。Bee Lim 等人在 ED-

*基金项目：国家自然科学基金项目（62271151）；福建省自然科学基金项目（2021J01580）

SR[6]中也引入了残差结构，EDSR 实际上是 SRResnet[7]的改进版，去除了传统残差网络中的 BN 层，在节省下来的空间中扩展模型尺寸来增强表现力。RCAN[8]中提出了一种基于 Residual in Residual 结构（RIR）和通道注意力机制（CA）的深度残差网络。虽然这些模型在当时取得了较好的效果，但本质上都是基于 CNN 网络的模型，网络中卷积核的大小会限制可以检测的空间范围，导致无法捕捉到长距离的依赖关系，意味着它们只能提取到局部特征，无法获取全局的信息，不利于纹理细节的恢复，使得图像重建的效果不佳[5]。

由于 Transformer 在自然语言处理（Natural Language Processing，NLP）领域中取得了较好的成效[9]，Alexey Dosovitskiy 等人将 Transformer 引入到计算机视觉（CV）领域中，即 ViT（Vision Transformer），并且在多个视觉任务中取得了成功[10]。ViT 的优势在于 Transformer 对于全局的信息更加敏感，模型中的注意力模块能够在输入序列的所有位置上进行全局交互，从而捕捉到长距离的依赖关系。但是，用于图像超分辨率的 ViT，需要将输入图像分割成固定大小的块，并对每个小块进行独立的处理[11]，这种处理策略就会产生两个弊端：一是修复后的图像可能会在每个小块周围引入边界伪影；二是每个补丁的边界像素可能会丢失信息，影响到重建图像的质量。Ze Liu 等人提出了 Swin-Transformer[11]，将滑动窗口机制引入到 Transformer 中，其中滑窗操作包括了不重叠的 Local-window 和重叠的 Cross-window，将注意力计算限制在一个窗口内可以大幅度节省计算量，并且通过滑窗操作也能使得注意力机制能够注意到全局的特征。

Jingyun Liang 等人结合卷积神经网络和 Swin-Transformer 两者的优点提出了 SwinIR[12]，将两个构架以串行的方式应用在超分领域，展示出了混合构架在超分任务中的巨大应用前景。Peng 等人提出了一种以 CNN-Transformer 并行的方法 Conformer[13]，它由一个 CNN 分支和一个 Transformer 分支组成，依靠特征耦合单元（Feature Coupling Unit，FCU）以交互的方式在不同分辨率下融合局部特征和全局特征，其结果表明了并行结构能以最大限度地保留局部特征和全局特征。虽然 Conformer 结合了两个网络的优势，但训练的参数量和训练时长也相应地增加了。并且，图像超分任务通常需要输入较高分辨率的图像，占用大量 GPU 内存，限制了模型的灵活性。要想取得更加高清的图像势必会增加网络模型中的参数量[14]。因此，在结合两种构架优势的同时，降低训练成本，使得图像超分辨率模型轻量化成为了本文需要解决的重大问题。

针对上述问题，本文提出了基于 Swin-Transformer 的单图像超分网络 STSR（Swin-Transformer-based Single Image Super-Resolution）。具体贡献如下：结合 CNN 和 Transformer 的优势，设计了并行特征提取的特征增强模块（Feature Enhancement Block，FEB），能够有效地捕捉图像的局部细节特征，同时也能够捕捉长距离的依赖关系，使模型具有全局上下文建模的能力。本文采用了轻量化的 Transformer 模块，在达到较好重建效果的同时还能保持较低的计算成本。此外，通过设计动态调整模块（Dynamic Adjustment，DA），可根据输入图像的特征对输出进行动态的调整，从而增强网络的拟合能力，使得重建图像的纹理细节更加贴近于真实图像。

1　本文算法

本文设计的基于 Swin-Transformer 的单图像超分网络 STSR 的整体网络结构如图 1 所示。结合 CNN 和 Transformer 两者的优势，本文设计了一个并行特征提取的特征增强模块 FEB，采用卷积神经网络和 Transformer 网络并行提取图像特征，为了减少训练中的参数量，FEB 中的轻量化 Transformer 模块只采用 Swin-Transformer 中的 STL（Swin Transformer Layer）来提取特征，并且在网络中堆叠 3 层 FEB 以获取更深层次的特征，有助于提升重建图像的质量。为了进一步的提升重建图像的质量，本文设计了一个动态调整模块 DA，动态调整模块主要采用的是通道注意力机制，通过这一机制，网络可以根据输入图像的内容和结构，自适应地增强或者减弱不同通道的影响，也可以减少网络对无关信息的依赖，使得在训练过程中网络能够快速收敛，减少训练时间，同时能够使得重建图像更加贴近于真实图像。STSR 主要包括了四个模块：浅层特征提取、深层特征提取，动态调整模块和高质量图像重建模块。

（1）浅层特征提取

首先，使用一个 3×3 的卷积层来提取图像的浅层特征 F_0，如式（1）所示。

$$F_0 = H_{3\times1}(I_{LR}) \tag{1}$$

（2）深层特征提取

因为浅层特征提取模块只经过一次简单的卷积操作，提取到较为初级的特征。将提取到的浅层特征 F_0，使用深层特征提取模块进一步地提取图像特征。深层特征提取模块由 K 个 FEB 模块和一个 3×3 的卷积层构成，用来提取深层特征 F_{DF}，如式（2）所示。

$$F_{DF} = H_{DF}(F_0) \tag{2}$$

其中每个 FEB 的输出 F_1、F_2、F_K 以及输出的深层特征 F_{DF}，如式（3）(4）所示。式中 H_{KF_i} 表示第 i 个 FEB 模块，H_{conv} 表示最终的卷积层。

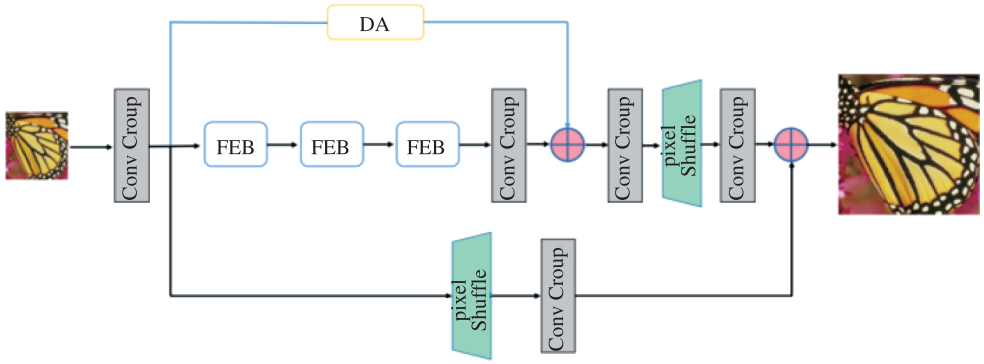


图 1　STSR 的整体网络框架

$$F_i = H_{KF_i}(F_{i-1}),\ i = 1, 2, \cdots, K \quad (3)$$

$$F_{DF} = H_{conv}(F_K) \quad (4)$$

（3）动态调整模块

动态调整模块主要采用通道注意力机制，通过计算通道间的相关性和重要性，网络可以自适应地增强或减弱不同通道的影响。这有助于提取输入图像中最相关和有用的特征，使得重建图像与输入图像具有更高的相似度，从而改善超分辨率重建的效果。

动态调整模块接收来自浅层特征提取模块输出的浅层特征 F_0，经过处理后得到动态调整特征 F_{DA}，如式（5）所示。

$$F_{DA} = H_{DA}(F_0) \quad (5)$$

（4）高质量图像重建模块

图像重建模块其实就是卷积与上采样的组合，本文中采用的是“卷积 + pixel shuffle + 卷积”的方式来进行图像重建。作为网络的最后部分，将接收浅层特征 F_0，深层特征 F_{DF} 和动态调整特征 F_{DA}，以获得重建图像 I_{SR}。f 和 f_p 分别代表的是卷积层和亚像素卷积层，计算 I_{SR} 的公式如下：

$$I_{SR} = f(f_p(f(F_{DF}))) + f(f_p(f(F_{DA}))) + f(f_p(F_0)) \quad (6)$$

1.1　特征增强模块（FEB）

本文方法中提出的 FEB 模块主要是由卷积神经网络 CNN、轻量型 Transformer 和特征融合模块三个部分组成。特征增强模块 FEB 的大致结构如图 2（a）所示。

（1）卷积神经网络

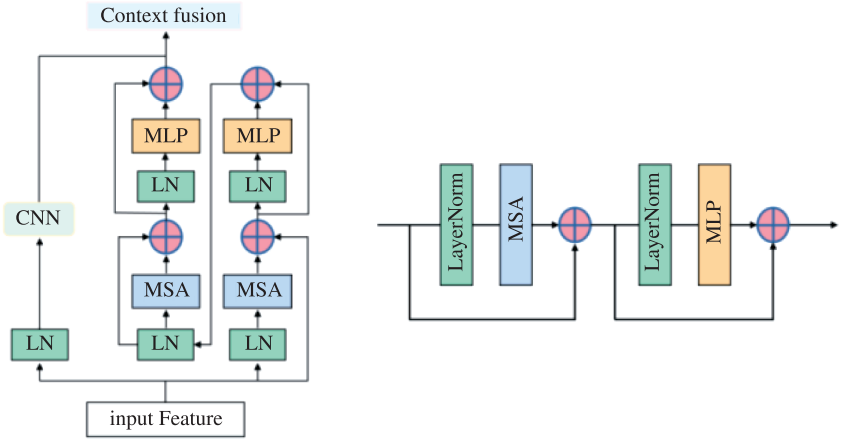
首先，利用卷积层来对浅层提取的特征 F_0 进行进一步的特征提取，得到卷积层提取的特征 F_C，如式（7）所示。

$$F_C = f_c(F_0) \quad (7)$$

其中 f_c 为 FEB 模块中卷积层中特征提取的映射关系。

（2）轻量型 Transformer 网络

轻量型 Transformer 是由 2 个 STL（Swin Transformer Layer）模块（如图 2（b）所示）串联组成的，第一个 STL 模块中 MSA（Multi-headed Self-Attention）采用的是 W-MSA（Window based Multi-headed Self-Attention），W-MSA 是窗口化的多头自注意力机制，相较于传统全局注意力机制减少了计算量。第二个 STL 模块采用的则是 SW-MSA（Shifted Window based Multi-headed Self-Attention），由于 W-MSA 只能关注窗口本身的内容，无法跨窗口连接，这就导致了窗口之间的特征信息无法传递，而



（a）FEB模块　　（b）STL模块

图 2　FEB 模块和 STL 模块结构

SW-MSA 可以通过引入移位窗口的方式，在保持窗口化的情况下兼顾了全局特征，并提高了计算效率，使得模型达到轻量化的效果。使用多层感知机（Multi-Layer Perception，MLP），其中包括了两个全连接层和 GELU 激活函数，来进行进一步的特征变化。在 MSA 和 MLP 之前都添加了 LN（LayerNorm）层，并且这两个模块都引入了残差连接。最后得到轻量化 Transformer 层提取的特征 F_T，如公式（8）所示。

$$F_T = H_{Swin}(F_0) \tag{8}$$

其中H_{Swin}为 FEB 模块中的轻量型 Transformer 层。

（3）特征融合模块

特征融合模块含有多尺度卷积块和特征重建部分，多尺度卷积块利用不同尺寸的卷积核来进行特征的多尺度提取，获得不同感受野的特征。特征融合模块采用残差连接的方式，将多个维度的特征相加，从而使网络更容易学习到低频和高频细节之间的映射关系，有助于加速训练过程。

在这个模块中，将卷积层提取的特征 F_C 和轻量型 Transformer 层提取的特征 F_T 来进行特征融合，得到 FEB 模块的输出特征 F_{Cf}，如式（9）所示。

$$F_{Cf} = H_{Cf}(F_C + F_T) \tag{9}$$

其中H_{Cf}为特征融合模块的函数。

1.2 动态调节模块（DA）

图 3 所展示的是动态调整模块，主要运用到的是全局平均池化（Global Average Pooling，GAP）操作。用于将卷积神经网络的特征图转换为固定长度的向量表示。GAP 通过将每个通道的特征图转换为一个标量值。具体来说，假设输入图像经过卷积层后得到的特征图为 F，其尺寸为 $H \times W \times C$，其中 H 和 W 是特征图的高度和宽度，C 是通道数，GAP 会对每个通道的特征图计算平均值，得到一个 C 维的向量表示，记为向量 $\boldsymbol{A}$，即每个通道的平均值。总而言之，通过 GAP 操作，特征图的空间信息被压缩成一个固定长度的向量。对向量 $\boldsymbol{A}$ 进行一系列的线性变换和非线性激活操作，得到一个长度为 C 的通道权重向量 $\boldsymbol{S}$。将特征图 F 与通道权重向量 $\boldsymbol{S}$ 进行逐通道乘法操作，得到调整后的特征图 F'。最后再将调整后的特征图 F'输入到后续的网络层中，继续进行超分辨率重建的处理。

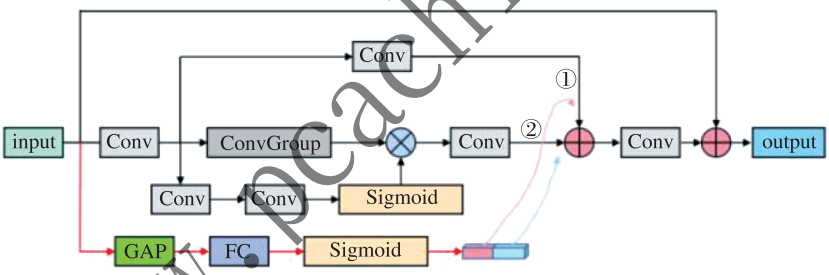


图 3　动态调整模块作用流程图

通过动态调整模块，网络可以根据输入图像的内容和结构，自适应地增强或者减弱不同通道的影响，以提取到输入图像中最为主要的特征，同时减少网络对无关或者冗余信息的依赖，从而改善超分辨率重建的质量和细节保留的能力。

1.3 损失函数

网络训练采用的损失函数是平均绝对误差（Mean Absolute Error，MAE），也称为 L_1 损失函数，如式（10）所示。

$$L_1 = \frac{1}{HWC}\sum_{i,j,k}|\hat{I}_{i,j,k} - I_{i,j,k}| \tag{10}$$

其中，H、W、C 分别是图像的高度、宽度、通道数，$I_{i,j,k}$和$\hat{I}_{i,j,k}$分别是 HR 图像和 SR 图像中第 i 行、第 j 列、第 k 条通道处的像素值。采用 L_1 损失函数训练模型会有较好的收敛能力。

2 实验结果及分析

2.1 数据集和评价标准

在训练阶段，采用 DIV2K 数据集来训练模型。该数据集包含有 1 000 张高清图像（2K 分辨率），其中包括 800 张图像作为训练数据，100 张图像作为验证数据，100 张图像作为测试数据。在评估阶段，采用的基准数据集是：Set5、Set14、BSD100、Urban100 和 Manga109。

在评价标准方面，采用峰值信噪比（PSNR）和结构相似度（SSIM）来作为衡量模型效果的评价指标。PSNR 可以评价两幅图像之间的相似程度。PSNR 值越高，说明重建出来的图像中的失真或者误差越小。较高的 PSNR 值表示重建图像质量较好，而数值较低意味着重建图像存在着更明显的伪影。SSIM 与专注于对比像素差异的 PSNR 不同，SSIM 主要比较图像的亮度、对比度和结构。通过计算三项的平均值来衡量相似性：亮度相似性、对比度

相似性和结构相似性。所以 SSIM 更加侧重于图像的结构信息和感知质量。

2.2 训练细节

网络训练所用平台为 Ubuntu20.04，所有的实验均在单张 NVIDIA GeForce GTX 3060Ti 显卡上完成训练。在训练之前，对 HR 图像进行不同缩放因子的双三次下采样，生成对应的 LR 图像。本文利用 L1 损失和 Adam 算法对模型进行优化。初始学习率定义为 1×10^{-4}，算法中参数设置为 $\beta_1=0.9$、$\beta_2=0.999$，训练周期定义为 300K iterations，并且采用余弦退火衰减的学习方案来加速收敛，动量 momentum 为 0.9。

2.3 模型比较

本文将 STSR 网络与一些经典的图像超分网络和近年来的轻量型图像超分网络进行对比，其中包括 FSRCNN[4]、VDSR[15]、MemNet[16]、EDSR[6]、CARN[8]、IMDN[17]、LatticeNet[18]、ESRT[19]。表 1 为在同一实验环境下所获得测试结果。在每一行中，最好的结果用加粗的方式突出显示。从测试结果可以看出在模型较小的情况下，量化指标也取得了较好的结果，在模型性能和计算成本之间取得了较好的平衡。

表 1 图像超分辨率重建效果量化比较结果

方法名	采样倍率	参数量	Set5	Set14	BSD100	Urban100	Manga109
			PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
Bicubic	×3	—	30.39/0.868 2	27.55/0.774 2	27.21/0.738 5	24.46/0.734 6	26.95/0.855 6
FSRCNN		13K	33.18/0.914 0	29.37/0.824 0	28.53/0.791 0	26.43/0.808 0	31.10/0.921 0
VDSR		666K	33.66/0.921 3	29.77/0.831 4	28.82/0.797 6	27.14/0.827 9	32.01/0.934 0
MemNet		678K	34.09/0.924 8	30.00/0.835 0	28.96/0.800 1	27.56/0.837 6	32.51/0.936 9
EDSR		1，555K	34.37/0.927 0	30.28/0.841 7	29.09/0.805 2	28.15/0.852 7	33.45/0.943 9
CARN		1，592K	34.29/0.925 5	30.29/0.840 7	29.06/0.803 4	28.06/0.849 3	33.50/0.944 0
IMDN		703K	34.36/0.927 0	30.32/0.841 7	29.09/0.804 6	28.17/0.851 9	33.61/0.944 5
LatticeNet		765K	**34.53**/0.928 1	30.39/0.842 4	29.15/0.805 9	28.33/0.858 3	-/-
ESRT		770K	34.42/0.926 8	30.43/0.843 3	29.15/0.806 3	28.46/0.857 4	33.95/0.945 5
STSR（ours）		**561K**	34.52/**0.928 6**	**30.53/0.846 0**	**29.18/0.808 8**	**28.48/0.859 2**	**33.96/0.946 9**
Bicubic	×4	-	28.42/0.810 4	26.00/0.702 7	25.96/0.667 5	23.14/0.657 7	24.89/0.786 6
FSRCNN		13K	30.72/0.866 0	27.61/0.755 0	26.98/0.715 0	24.62/0.728 0	27.90/0.861 0
VDSR		666K	31.35/0.883 8	28.01/0.767 4	27.29/0.725 1	25.18/0.752 4	28.83/0.887 0
MemNet		678K	31.74/0.889 3	28.26/0.772 3	27.40/0.728 1	25.50/0.763 0	29.42/0.894 2
EDSR		1，518K	32.09/0.893 8	28.58/0.781 3	27.57/0.735 7	26.04/0.784 9	30.35/0.906 7
CARN		1，592K	32.13/0.893 7	28.60/0.780 6	27.58/0.734 9	26.07/0.783 7	30.47/0.908 4
IMDN		715K	32.21/0.894 8	28.58/0.781 1	27.56/0.735 3	26.04/0.783 8	30.45/0.907 5
LatticeNet		777K	32.30/0.896 2	28.68/0.783 0	27.62/0.736 7	26.25/0.787 3	-/-
ESRT		751K	32.19/0.894 7	28.69/0.783 3	27.69/0.737 9	26.39/0.796 2	30.75/0.910 0
STSR（ours）		**570K**	**32.32/0.896 8**	**28.69/0.784 2**	**27.66/0.740 6**	**26.43/0.796 1**	**30.90/0.913 2**

2.4 消融实验

为了验证特征增强模块 FEB 以及动态调整模块 DA 的实际作用，本文将 FEB 模块和 DA 模块分别单独引入到 STSR 模型中，通过实验来验证这两个模块在量化指标 PSNR 和 SSIM 上的效果，可以验证出不同模块对网络性能提升的有效性。方法 STSR-F 是将 STSR 模型中的特征增强模块 FEB 替换为单一的卷积神经网络结构，由此来验证 FEB 模块的作用。方法 STSR-D 是去掉 STSR 模型中的动态调整模块 DA，由此来验证 DA 模块在网络中的作用。

可以从表 2 中的量化指标得出特征增强模块 FEB 以及动态调整模块 DA 对于本文网络性能提升的有效性。

2.5 视觉效果评价

图 4 中给出了 STSR 与其他图像超分模型的视觉效果对比。本文设计模型重建出的高分辨率图像包含了更多纹理细节，对于边缘和线条等细节方面的重建，也展现出了更高质量的重建效果。

表 2　消融实验

方法名	采样倍率	Set5	Set14	BSD100	Urban100
		PSNR/SSIM	PSNR/SSIM	PSNR/SSIM	PSNR/SSIM
STSR-F	×3	34. 26/0. 925 2	30. 24/0. 839 2	29. 01/0. 802 9	28. 02/0. 849 0
STSR-D		34. 45/0. 927 8	30. 42/0. 843 6	29. 12/0. 806 0	28. 43/0. 856 9
STSR		**34. 52/0. 928 6**	**30. 53/0. 846 0**	**29. 18/0. 808 8**	**28. 48/0. 859 2**

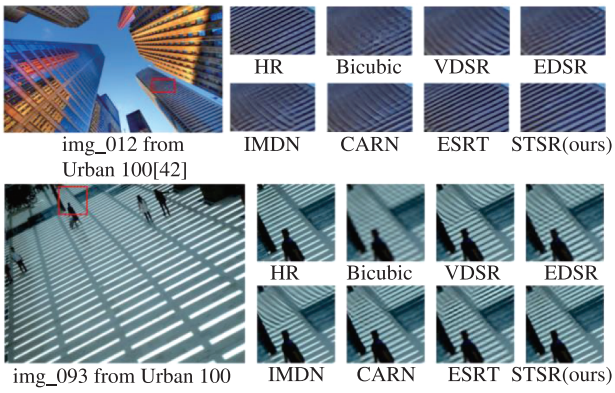


图 4　STSR 与其他模型的视觉对比

3　结论

针对基于混合构架的图像超分模型较高计算成本的问题，本文提出了一种基于卷积神经网络 CNN 和 Transformer 混合模型，结合了两种构架的优势，提高模型对局部细节特征和全局信息的建模能力，增强了上下文信息的利用效率，并且通过轻量化 Transformer 网络，在达到较好重建效果的同时保持较低的计算成本。此外，本文还设计了 DA 模块，根据输入图像来动态调整网络的输出，使重建的图像更加贴近于真实图像。量化指标表明，本文方法在减少模型参数量的情况下，依然取得了较好的图像重建效果。

参考文献

[1] WANG Z, CHEN J, HOI S C H. Deep learning for image super-resolution: a survey [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2020, 43 (10): 3365 – 3387.

[2] GUO Y, CHEN J, WANG J, et al. Closed-loop matters: dual regression networks for single image super-resolution [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2020: 5407 – 5416.

[3] DONG C, LOY C C, HE K, et al. Image super-resolution using deep convolutional networks [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015, 38 (2): 295 – 307.

[4] DONG C, LOY C C, TANG X. Accelerating the super-resolution convolutional neural network [J]. Springer International Publishing, 2016: 391 – 407.

[5] HE K, ZHANG X, REN S, et al. Deep residual learning for image recognition [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 770 – 778.

[6] LIM B, SON S, KIM H, et al. Enhanced deep residual networks for single image super-resolution [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 136 – 144.

[7] LEDIG C, THEIS L, HUSZAR F, et al. Photo-realistic single image super-resolution using a generative adversarial network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 4681 – 4690.

[8] ZHANG Y, LI K, LI K, et al. Image super-resolution using very deep residual channel attention networks [C]//Proceedings of the European Conference on Computer Vision, 2018: 286 – 301.

[9] VASWANI A, SHAZEER N, PARMAR N, et al. Attention is all you need [J]. arXiv preprint arXiv: 1706. 03762, 2017.

[10] DOSOVITSKIY A, BEYER L, KOLESNIKOV A, et al. An image is worth 16x16 words: transformers for image recognition at scale [J]. arXiv preprint arXiv: 2010. 11929, 2020.

[11] LIU Z, LIN Y, CAO Y, et al. Swin transformer: hierarchical vision transformer using shifted windows [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision. 2021: 10012 – 10022.

[12] LIANG J, CAO J, SUN G, et al. Swinir: image restoration using swin transformer [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 1833 – 1844.

[13] PENG Z, HUANG W, GU S, et al. Conformer: local features coupling global representations for visual recognition [C]//Proceedings of the IEEE/CVF International Conference on Computer Vision, 2021: 367 – 376.

[14] TAI Y, YANG J, LIU X. Image super-resolution via deep recursive residual network [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2017: 3147 – 3155.

[15] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, 2016: 1646 – 1654.

[16] TAI Y, YANG J, LIU X, et al. Memnet: a persistent memory

network for image restoration [C]//Proceedings of the IEEE International Conference on Computer Vision, 2017: 4539 -4547.

[17] HUI Z, GAO X, YANG Y, et al. Lightweight image super-resolution with information multi-distillation network [C]//Proceedings of the 27th ACM International Conference on Multimedia, 2019: 2024 -2032.

[18] LUO X, XIE Y, ZHANG Y, et al. Latticenet: towards lightweight image super-resolution with lattice block [J]. Springer International Publishing, 2020: 272 -289.

[19] LU Z, LI J, LIU H, et al. Transformer for single image super-resolution [C]//Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2022: 457 -466.

(收稿日期: 2024 -02 -21)

作者简介:

林承浩 (1999 -), 男, 硕士研究生, 主要研究方向: 计算机视觉。

吴丽君 (1984 -), 通信作者, 女, 博士, 副教授, 主要研究方向: 计算机视觉。E-mail: lijun. wu@ fzu. edu. cn。

版权声明

凡《网络安全与数据治理》录用的文章，如作者没有关于汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权等版权的特殊声明，即视作该文章署名作者同意将该文章的汇编权、翻译权、印刷权及电子版的复制权、信息网络传播权与发行权授予本刊，本刊有权授权本刊合作数据库、合作媒体等合作伙伴使用。同时，本刊支付的稿酬已包含上述使用的费用，特此声明。

《网络安全与数据治理》编辑部