



计算机科学与探索

Journal of Frontiers of Computer Science and Technology

ISSN 1673-9418, CN 11-5602/TP

## 《计算机科学与探索》网络首发论文

题目: 多尺度自适应上采样的图像超分辨率重建算法  
作者: 吕佳, 许鹏程  
网络首发日期: 2021-10-25  
引用格式: 吕佳, 许鹏程. 多尺度自适应上采样的图像超分辨率重建算法[J/OL]. 计算机科学与探索. <https://kns.cnki.net/kcms/detail/11.5602.TP.20211020.1013.002.html>



**网络首发:** 在编辑部工作流程中, 稿件从录用到出版要经历录用定稿、排版定稿、整期汇编定稿等阶段。录用定稿指内容已经确定, 且通过同行评议、主编终审同意刊用的稿件。排版定稿指录用定稿按照期刊特定版式 (包括网络呈现版式) 排版后的稿件, 可暂不确定出版年、卷、期和页码。整期汇编定稿指出版年、卷、期、页码均已确定的印刷或数字出版的整期汇编稿件。录用定稿网络首发稿件内容必须符合《出版管理条例》和《期刊出版管理规定》的有关规定; 学术研究成果具有创新性、科学性和先进性, 符合编辑部对刊文的录用要求, 不存在学术不端行为及其他侵权行为; 稿件内容应基本符合国家有关书刊编辑、出版的技术标准, 正确使用和统一规范语言文字、符号、数字、外文字母、法定计量单位及地图标注等。为确保录用定稿网络首发的严肃性, 录用定稿一经发布, 不得修改论文题目、作者、机构名称和学术内容, 只可基于编辑规范进行少量文字的修改。

**出版确认:** 纸质期刊编辑部通过与《中国学术期刊 (光盘版)》电子杂志社有限公司签约, 在《中国学术期刊 (网络版)》出版传播平台上创办与纸质期刊内容一致的网络版, 以单篇或整期出版形式, 在印刷出版之前刊发论文的录用定稿、排版定稿、整期汇编定稿。因为《中国学术期刊 (网络版)》是国家新闻出版广电总局批准的网络连续型出版物 (ISSN 2096-4188, CN 11-6037/Z), 所以签约期刊的网络版上网络首发论文视为正式出版。

# 多尺度自适应上采样的图像超分辨率重建算法

吕佳<sup>1,2+</sup>, 许鹏程<sup>1,2</sup>

1. 重庆师范大学 计算机与信息科学学院, 重庆 401331

2. 重庆师范大学 重庆市数字农业服务工程技术研究中心, 重庆 401331

+ 通信作者 E-mail: lvjia@cqnu.edu.cn

**摘要:** 针对现有基于卷积神经网络的图像超分辨率重建网络中存在的训练难以收敛、无法适配多个放大系数和无法进行非整数放大系数上采样的问题, 提出了一种基于多尺度和分而治之思想的自适应上采样图像超分辨率重建算法。该算法通过改进后的多尺度通道注意力特征提取模块对低分辨率图像进行多尺度特征提取以生成不同尺度下的特征图, 再将其输入瓶颈层实现全局特征融合, 使用基于分而治之的自适应上采样模块获得超分辨率图像, 从而解决了不同放大系数的适配问题和非整数放大系数的上采样问题。在对比实验中, 本文算法在不使用任何初始化方法时仍然具有良好的收敛性。在整数放大系数下本文算法的图像重建性能超过当前主流的超分辨率网络, PSNR 和 SSIM 性能相比 MRFN 分别提升了 0.34db 和 0.0391。在非整数放大系数下, 其 PSNR 性能相比双三次插值方法平均提升 1.24db, 且不需要对每一个放大系数都进行训练。

**关键词:** 图像处理; 超分辨率重建; 卷积神经网络; 多尺度特征提取; 分而治之

**文献标志码:** A **中图分类号:** TP391.41

## Image Super-resolution Reconstruction Algorithm Based on Adaptive Upsampling of Multi-scale

LYU Jia<sup>1,2+</sup>, XU Pengcheng<sup>1,2</sup>

1. College of Computer and Information Sciences, Chongqing Normal University, Chongqing 401331, China

2. Chongqing Center of Engineering Technology Research on Digital Agriculture Service, Chongqing Normal University, Chongqing 401331, China

**Abstract:** In view of the problems existing in the existing convolutional neural network-based super-resolution network, such as difficult convergence of training, unable to adapt to multiple upsampling coefficients and unable to sample non-integer upsampling coefficients, an image super-resolution reconstruction algorithm of adaptive upsampling module based on multi-scale and the idea of divide and conquer is proposed in this paper. In the algorithm, the

**基金项目:** 国家自然科学基金重大项目 (11991024); 重庆市高校创新研究群体资助 (CXQT20015); 重庆市研究生科研创新项目 (CYS20241)。

This work was supported by the Major projects of National Natural Science Foundation of China (11991024), the Chongqing innovative research group Foundation of China (CXQT20015) and the Chongqing postgraduate research and innovation Project of China (CYS20241).

improved multi-scale channel attention feature extraction module is used to extract multi-scale features of low-resolution images to generate feature maps at different scales, and then the global feature fusion is realized by inputting them into the bottleneck layer. The super-resolution images are obtained by using the adaptive upsampling module based on divide-and-conquer to solve the adaptation problem of different upsampling coefficients and the upsampling problem of non-integer upsampling coefficients. In the contrast experiment, the proposed algorithm still has good convergence without any initialization methods. Under the integer magnification factor, the image reconstruction performance of the proposed algorithm exceeds the current mainstream super-resolution network, and the PSNR and SSIM performance are improved by 0.34 db and 0.0391 respectively compared with MRFN. Under the non-integer magnification factor, the average PSNR performance is improved by 1.24 db compared with the bicubic interpolation method, and there is not necessary to train each magnification factor.

**Key words:** image processing; super-resolution reconstruction; convolution neural network; multi-scale feature extraction; divide-and-conquer

图像超分辨率 (Super Resolution, SR) 重建是计算机视觉和图像处理中的一类重要的图像处理方法, 是将低分辨率 (Low-Resolution, LR) 图像恢复成高分辨率 (High-Resolution, HR) 图像的过程<sup>[1]</sup>。SR 重建主要包括基于插值的方法、基于重建的方法和基于学习的方法等<sup>[1]</sup>。基于插值的方法有双三次插值法<sup>[2]</sup>, 最为简单且计算速度最快, 但模糊了边缘; 基于重建的方法有迭代反向投影法<sup>[3]</sup>等, 具有较快的计算速度, 但无法对复杂图像进行有效的重建; 基于学习的方法有稀疏表示法<sup>[4]</sup>等, 因其重建效果最好, 成为当前研究热点并广泛被采用。2014 年, Dong 等<sup>[5]</sup>将卷积神经网络首次应用到了超分辨率重建领域当中, 提出了超分辨率卷积神经网络 (Super-Resolution Convolutional Neural Network, SRCNN), 但是计算速度较慢; 2016 年, Dong 等<sup>[6]</sup>和 Shi 等<sup>[7]</sup>在 SRCNN 的基础上分别提出了基于反卷积的反卷积的快速图像超分辨率卷积神经网络 (Fast Super-Resolution Convolutional Neural Network, FSRCNN) 和基于亚像素卷积的高效亚像素卷积神经网络 (Efficient Sub-Pixel Convolutional Neural Network, ESPCN), 两者均采用后上采样方法, 使得超分辨率网络的计算量大幅度降低, FSRCNN 的运算速度相对 SRCNN 提升了 17.36 倍<sup>[2]</sup>; 同年, Kim 等<sup>[8]</sup>提出了基于残差学习的深层超分辨率 (Very Deep Super-Resolution, VDSR) 网络, 解决了在超分辨率网络中随着深度增加而导致的梯度消失现象; 2018 年, Liu 等<sup>[9]</sup>将注意力机制和

密集连接引入超分辨率网络中, 以 SRResNet<sup>[10]</sup>作为主干网络, 取得了较好的效果。

然而, 在上述方法中采用后上采样的超分辨率网络存在变更放大系数时需要更改网络结构的问题。另外, 大部分网络对结构的变化非常敏感, 直接修改网络结构会导致网络失效<sup>[11]</sup>。因此, 直接修改放大系数以使一个网络适配多个放大系数的方法并不可行。2019 年, Hu 等<sup>[12]</sup>基于元学习提出了使用采样核动态上采样的 Meta-Upscale 模块, 在不更改参数的情况下使一个网络适配任意放大系数成为可能, 并在 Meta-RDN 网络上取得了不错的效果。但是, 随着网络的加深以及元学习的使用, 导致网络的训练变得十分困难, 初始化方法和训练技巧的重要性甚至超过了网络结构<sup>[11]</sup>。因此, 搭建一个简单有效且能够自适应地适配任意放大系数的网络成为了目前 SR 重建领域的研究热点。

为此, 本文提出一种基于多尺度特征提取的超分辨率重建方法。采用多个亚像素层作为上采样模块, 在网络中采用分而治之的思想选择不同的亚像素层适配对应的放大系数。同时, 使用一个下采样模块控制输出图像的大小, 从而实现对一个网络进行任意系数的上采样。实验结果表明, 本文算法在不使用任何权重初始化方法的情况下仍然具有良好的收敛性, 且整数放大系数下的重建性能超过当前的主流 SR 网络。在非整数放大系数下性能超过双三次插值, 且只需训练比当前放大系数大的整数放大系数。

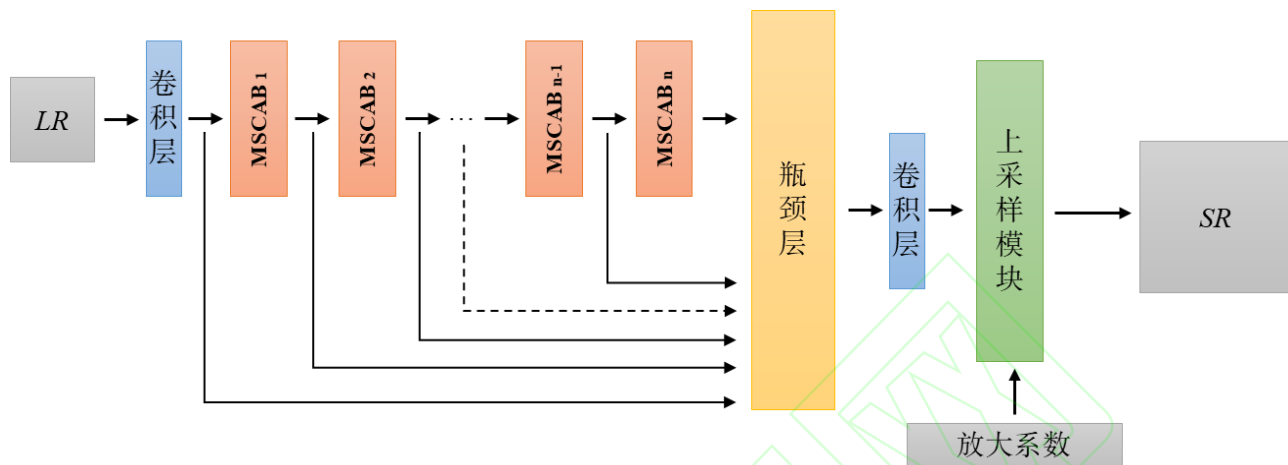


Fig.1 Network structure diagram of our algorithm

图1 本文算法网络结构图

## 1 基本原理

### 1.1 特征提取模块

目前,超分辨率网络通常采用三种特征提取模块,分别是残差块、稠密块和 Inception 块<sup>[13]</sup>。残差块采取了跳跃连接的方法,有效地避免了网络过深导致的特征消失问题。稠密块采用密集连接的方式,其每一层的输入都是前面所有层的输出,增强了特征传播,减轻了梯度消失。然而,随着网络深度的增加,稠密块的计算复杂度会急剧增加,会降低计算速度并使得网络的训练变得更加困难。Inception 块中包含多个分支,每个分支对应不同大小的卷积核以提取不同尺度下的特征。之后通过级联的方式将每个分支生成的特征图进行合并,并通过  $1 \times 1$  卷积进行降维以减少计算量。这是一种有效的特征提取方法,在确保参数量不会大幅增加的同时能够实现多尺度特征提取。

### 1.2 上采样方法

上采样是 SR 重建的核心,目前虽已有多种传统的上采样算法,但利用神经网络学习端到端的上采样过程可降低计算量和空间复杂度,因此成为当前主流的上采样方法<sup>[1]</sup>,如反卷积<sup>[6]</sup>和亚像素卷积<sup>[7]</sup>等。本文采用亚像素卷积作为上采样方法。

亚像素层是一个端到端的可学习上采样层,它通过卷积生成多个通道,然后对其进行排列组合来

执行上采样<sup>[1]</sup>。由于其采用端到端的上采样方式,亚像素层被 SR 模型广泛使用。与反卷积层相比,亚像素层的优点是感受野更大,能提供更多的上下文信息以生成更精确的细节。然而,亚像素层的感受野分布是不均匀的,同一块状区域共享相同的感受野,这可能导致在不同块的边界附近会出现一些噪声,在实际应用中表现为噪点的出现。

目前,基于学习的上采样层在 SR 重建中被广泛应用。尤其是在后上采样模型中,这些层通常在最后的上采样阶段用于从低维空间提取高阶特征来实现 SR 重建,在避免因高维空间大量计算的同时实现端到端的 SR 重建。

### 1.3 注意力机制

注意力机制在一系列计算机视觉任务中表现出良好的性能,其根据特征的重要性来分配权重,使得网络可以更加关注与输出相关的特征,并弱化无关特征<sup>[9]</sup>。注意力机制可分为空间注意力<sup>[9]</sup>和通道注意力<sup>[14]</sup>等。本文针对特征提取模块采用局部通道注意力机制,以提高特征提取模块的性能。

通道注意力机制通过对特征通道之间的相互依赖性建模来自适应地重新缩放每个通道的特征,从而提高网络的特征表达能力<sup>[14]</sup>。在 SR 任务中,引入通道注意力机制来区别对待不同的通道,可以有效提高网络重建性能。



## 2 本文算法

本文在分而治之思想的基础上对现有的多尺度残差模块(Multi-scale Residual Block, MSRB)<sup>[11]</sup>进行改进, 利用改进后的多尺度通道注意力特征提取模块(Multi-scale with Channel Attention block, MSCAB)在不同尺度提取特征。为了最大化利用 LR 图像特征和解决传输过程中的特征消失问题, 将每一个 MSCAB 的特征和全局特征相结合, 并通过使用  $1 \times 1$  卷积的瓶颈层进行全局特征融合。使用自适应上采样模块将特征图放大至相应的放大系数得到 SR。整体网络结构见图 1, 网络输入 LR 和放大系数, LR 经过  $n$  个 MSCAB 进行多尺度特征提取后通过瓶颈层进行全局特征融合, 上采样模块采用分而治之的思想, 根据输入的放大系数自适应地选择相应的亚像素层进行上采样处理, 通过一层卷积层将通道数转换为 RGB 通道后最终输出 SR。

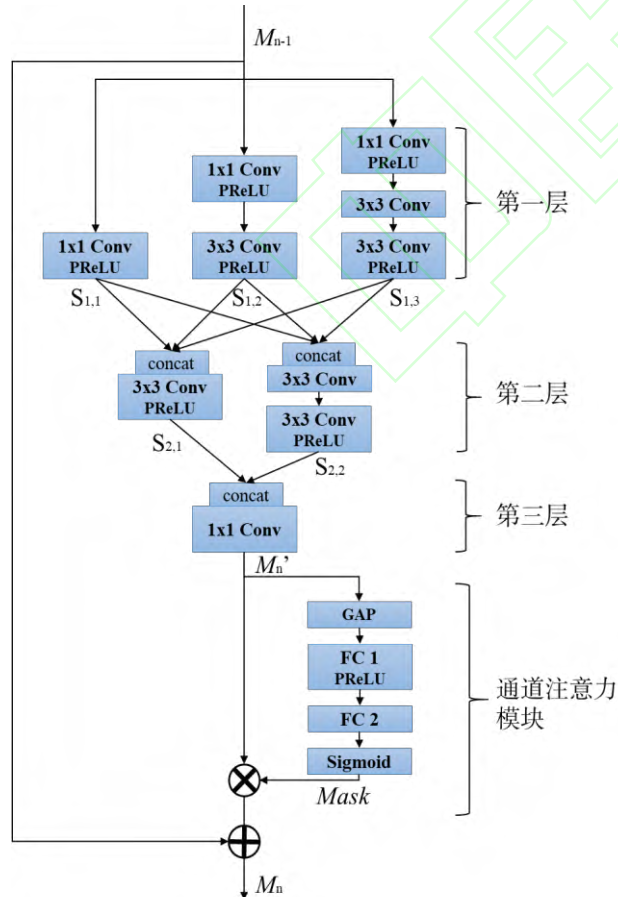


Fig.2 The structure of MSCAB

图 2 多尺度通道注意力模块结构

### 2.1 多尺度局部通道注意力特征提取模块

MSCAB 在 MSRB 基础上改进得到, 采取类似于 Inception 块的方法, 通过引入了不同尺度的卷积核, 用于自适应检测不同尺度下的图像特征。同时, 在不同尺度的特征之间采用跳跃连接, 使得特征信息可以相互共享和重用<sup>[11]</sup>, 这有助于充分利用图像的局部特征。MSCAB 的结构如图 2 所示。

MSCAB 由多尺度特征提取模块和局部通道注意力模块两个部分组成。在特征提取模块中增加卷积层进行升降维, 配合带参数的线性整流函数 (Parametric Rectified Linear Unit, PReLU)<sup>[15]</sup>以提升特征图的非线性特性。局部通道注意力模块对特征提取模块依据通道特性进行自适应校准, 强化与特征相关的通道, 同时抑制无效信息, 从而提升特征提取模块的性能。

#### 2.1.1 多尺度特征提取模块

本文在特征提取模块中构造了一个多分支结构的多层网络。第一层网络中包含了三个分支, 第二层包含两个分支, 不同的分支采用不同尺寸的卷积核, 从而可以提取出图像不同尺度下的特征。最后级联两个分支, 通过  $1 \times 1$  卷积操作使得输出特征图的通道数与输入相等。由于两个串联的  $3 \times 3$  卷积的感受野与一个  $5 \times 5$  卷积相同, 但参数量比后者小很多<sup>[16]</sup>, 约为后者的 72%, 因此本文使用两个  $3 \times 3$  卷积串联来代替  $5 \times 5$  卷积。模块按照如下公式执行:

$$\sigma(x) = \max(0, x) + \alpha \min(0, x) \quad (1)$$

$$S_{1,1} = \sigma(W_{1 \times 1}^{1,1} \otimes M_{n-1} + b^{1,1}) \quad (2)$$

$$S_{1,2} = \sigma(W_{3 \times 3}^{1,3} \otimes \sigma(W_{1 \times 1}^{1,2} \otimes M_{n-1} + b^{1,2}) + b^{1,3}) \quad (3)$$

$$S' = W_{3 \times 3}^{1,5} \otimes \sigma(W_{1 \times 1}^{1,4} \otimes M_{n-1} + b^{1,4}) + b^{1,5} \quad (4)$$

$$S_{1,3} = \sigma(W_{3 \times 3}^{1,6} \otimes S' + b^{1,6}) \quad (5)$$

$$S_{2,1} = \sigma(W_{3 \times 3}^{2,1} \otimes [S_{1,1}, S_{1,2}, S_{1,3}] + b^{2,1}) \quad (6)$$

$$S_{2,2} = \sigma(W_{3 \times 3}^{2,3} \otimes (W_{3 \times 3}^{2,2} \otimes [S_{1,1}, S_{1,2}, S_{1,3}] + b^{2,2}) + b^{2,3}) \quad (7)$$

$$M'_n = W_{1 \times 1}^{3,1} \otimes [S_{2,1}, S_{2,2}] + b^{3,1} \quad (8)$$

$$M_n = M_{n-1} + M'_n \quad (9)$$

其中, 公式(1)表示 PReLU 函数, 函数中的  $\alpha$  由模型学习得来。公式(2)到公式(8)中的  $W$  和  $b$  表示权

重和偏置参数,  $\otimes$  代表卷积操作,  $b$  上标的第一个参数表示其所在的层数, 第二个参数为该层的第几个卷积;  $W$  的下标表示卷积核大小,  $W$  上标的第一个参数为所在层数, 第二个参数为该层的第几个卷积。第一层中包含了 6 个卷积, 第二层包含 3 个卷积, 第三层包含 1 个卷积。本文在对图像进行  $3 \times 3$  卷积时, 对图像边界进行补零操作, 以确保输入和输出特征图的维度相等。公式(6)到公式(8)中的  $[S_{1,1}, S_{1,2}, S_{1,3}]$ ,  $[S_{2,1}, S_{2,2}]$  表示级联操作, 沿通道维度将特征图进行级联。 $M_n$  为经过多尺度特征提取模块处理后的特征图。

在每一个多尺度特征提取模块中, 输入特征图  $M_{n-1}$  分别通过公式(2)、(3)和(4)进行升降维并提取特征获得特征图  $S_{1,1}$ 、 $S_{1,2}$  和  $S'$ ; 公式(5)对  $S'$  进行进一步提取特征以提升感受野, 获得感受野与  $5 \times 5$  卷积相同的特征图  $S_{1,3}$ ; 公式(6)和(7)分别将特征图  $S_{1,1}$ 、 $S_{1,2}$  和  $S_{1,3}$  级联后进一步执行特征提取工作, 级联操作融合了第一层提取的多尺度特征, 确保在第二层的特征提取中得以充分利用, 同时公式(7)包含两个  $3 \times 3$  卷积, 进一步增加了特征图的感受野。公式(8)将公式(6)和(7)的输出级联, 在执行  $1 \times 1$  卷积去除特征图中的冗余信息的同时对残差图  $M_n'$  降维, 使其与输入特征图  $M_{n-1}$  通道一致。最后, 公式(9)将残差图与输入特征图相加, 获得最终的输出  $M_n$ 。3.5.1 的消融实验证明, 本文提出的多尺度特征提取模块是有效的。

用  $M$  表示输入多尺度特征提取模块特征图的通道数。因此第一层的输入通道数为  $M$ , 输出为  $3M$ 。第二层的输入通道数为  $3M$ , 输出为  $2M$ 。第二层输出的所有特征图被级联并输入第三层。第三层将这些特征图的通道数减少到  $M$ 。因此 MSCAB 的输入和输出具有相同通道数量的特征图, 该结构允许多个 MSCAB 一起串联使用。

### 2.1.2 局部通道注意力模块

局部通道注意力模块首先通过全局平均池化(Global Average Pooling, GAP)获取特征提取模块输出特征图中每个通道的全局信息, 然后通过多个全连接

层和激活函数自适应地建立通道间的相互关系, 最后使用 Sigmoid 函数将权重压缩至 0~1 之间并对原特征通道进行加权处理。模块按照如下公式执行:

$$X = \text{vec}(\text{GAP}(M_n')) \quad (10)$$

$$\text{Mask} = W_2 * \sigma(W_1 * X + b_1) + b_2 \quad (11)$$

$$M_n' = M_n * \text{Sigmoid}(\text{Mask}) \quad (12)$$

其中,  $X$  表示经过 GAP 的通道维度向量,  $\text{vec}$  表示向量化操作, 将 GAP 后的特征图转换为向量, 向量中的每一个值表示每个通道的全局信息。 $W$  和  $b$  表示全连接层的权重和偏置参数, 其下标表示其来自第几个全连接层。 $\text{Mask}$  为局部通道注意力模块生成的权重向量。多尺度特征提取模块输出的残差图  $M_n'$  首先经过公式(10)执行 GAP 并被拉直为向量  $X$ , 之后通过公式(11)对特征通道之间的相互依赖性进行建模并获得每个通道的权重向量  $\text{Mask}$ , 最终将权重向量压缩至 0~1 之间并与原残差图点乘以获得校准后的残差图  $M_n'$ 。

局部通道注意力模块的输入和输出通道数与多尺度特征提取模块的输出通道数相同, 因此可以直接添加进多尺度特征提取模块中。

## 2.2 瓶颈层

在单张图像的 SR 重建中, 输入和输出图像高度相关, 因此在重建时充分利用输入图像的特征非常重要。然而, 随着网络的加深, 这些特征在传输过程中逐渐消失<sup>[11]</sup>。跳跃连接和密集连接都是简单有效的解决该问题的方法, 这些方法都通过在不同层之间创建不同的连接, 但并不能充分利用输入图像的特征, 且会产生过多的冗余信息。

随着网络深度的增加, 网络的空间表达能力逐渐下降, 而语义表达能力逐渐增强。因此, 如何充分利用浅层空间特征和深层语义特征, 将直接影响模型重建的质量。本文将每个 MSCAB 的输出传至网络的末端进行重建。但是, 这些特征映射包含大量冗余信息, 直接使用它们进行重建会大幅增加计算复杂度。为了自适应地提取其中的有用信息, 采用了一个瓶颈层, 利用  $1 \times 1$  卷积降低输入特征图的通道数。该部分的输出可以表示为:

$$F_{LR} = W \otimes [M_0, M_1, M_2, \dots, M_{n-1}, M_n] + b \quad (13)$$

其中,  $M_0$  表示第一个卷积层的输出,  $M_i$  ( $i \neq 0$ ) 表示第  $i$  个 MSCAB 的输出,  $[M_0, M_1, M_2, \dots, M_{n-1}, M_n]$  表示沿通道维度的级联操作。

### 2.3 上采样模块

传统基于卷积神经网络的 SR 重建问题通常采用反卷积层和亚像素层进行上采样。但是, 采用反卷积的方法由于其放大系数已经固定到其参数内, 因此在变更放大系数时无法通过简单的设置来使其进行适配。亚像素层虽然在适配任意放大系数时较为简单, 但是在该过程中仍然需要对网络的参数进行调整。因此, 本文提出了一种采用基于分而治之的自适应上采样模块。该模块内使用两个亚像素层来执行上采样操作, 通过输入的放大系数来自动对其排列组合使得输出的图像与之适配。图 3 为该模块中两个亚像素层的结构。

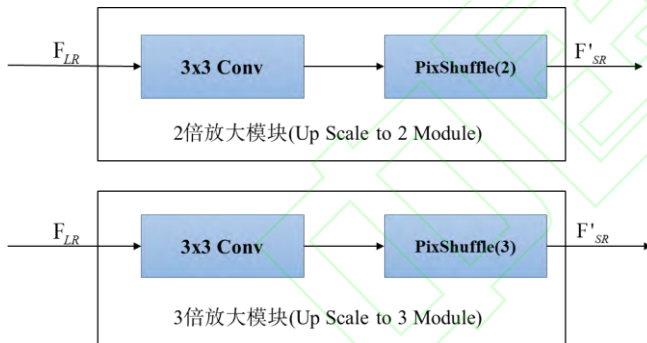


Fig.3 The structure of two sub-pixel layers

图 3 两个亚像素层的结构

如果 LR 的维度为  $W \times H$ , 通道数为  $r$ ; 放大系数为  $s$ , 亚像素层就将  $s^2 r$  个通道的  $W \times H$  特征图转换为  $r$  个通道的  $sW \times sH$  特征图, 从而实现上采样操作。而在每个亚像素层前的卷积层的作用是提升特征图的通道数, 从  $r$  通道提升为  $s^2 r$  通道。在 2 倍放大模块中, 卷积层输入  $r$  通道特征图, 输出  $4r$  通道特征图; 3 倍放大模块中, 卷积层输入  $r$  通道特征图, 输出  $9r$  通道特征图。图 4 为放大系数为 2、3、4 时该模块的结构图。

分支结构在输入放大系数后使用两个放大模块进行匹配, 使用迭代的方法实现更大的放大倍

数, 如 4 倍放大系数可由 2 倍放大模块执行两次得到; 6 倍放大系数可先用 2 倍放大模块放大后, 再用 3 倍放大模块进一步放大得到。但是, 这种方法产生的目标放大系数必须以两个放大模块的放大系数为约数, 无法实现任意放大系数的上采样。因此, 引入了一个下采样模块在上采样完成后对其下采样至目标放大系数。同时, 该方法也可以实现非整数的放大系数。

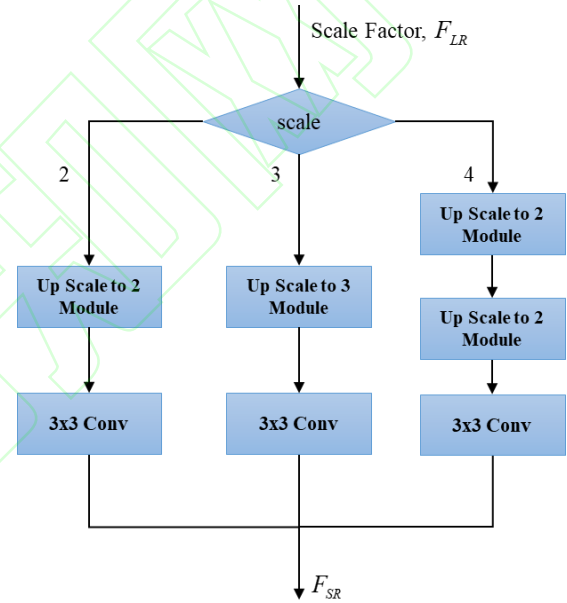


Fig.4 Upsampling module structure under the amplification factor 2, 3 and 4

图 4 放大系数为 2, 3, 4 时上采样模块结构

Table 1 Algorithm flow of upsampling module

表 1 上采样模块算法流程

算法: 上采样模块

输入: scale\_factor、 $F_{LR}$

1. 分别计算 scale\_factor 和 2, 3 的倍数, 生成 scalelist 数组和 target\_scale.
2. 遍历 scalelist 依次获得其中的元素值, 并与 2 和 3 进行比较.
3. 若值为 2 则进行  $F' = \text{UpTo2}(F_{LR})$ ; 为 3 则进行  $F' = \text{UpTo3}(F_{LR})$ .
4. 对  $F'$  进行一次卷积操作以降维到 RGB 色彩空间.
5. 根据 scale\_factor / target\_scale 计算出 ratio, 若 ratio 不为 1 则通过  $\text{Bilinear}(F', \text{ratio})$  计算  $F'$ .
6. 将  $F'$  赋值给  $F_{SR}$  得到最终结果.

输出:  $F_{SR}$



上采样模块包含  $scalelist$  数组,  $target\_scale$ ,  $F'$  和  $ratio$  等参数。其输入为放大系数  $scale\_factor$  和  $F_{LR}$ , 输出为  $F_{SR}$ , 算法流程如表 1 所示。 $scalelist$  为只包含 2 和 3 的上采样数组, 通过计算  $scale\_factor$  和 2, 3 之间的因数得到, 无法匹配时则匹配到与其绝对值最小的公倍数。如放大系数为 5 时,  $scalelist=\{2,3\}$ 。 $target\_scale$  是  $scalelist$  中所有元素相乘的结果, 即该模块实际放大的倍数, 在上述情况下  $target\_scale$  为 6。利用循环匹配  $scalelist$  中的每个元素, 并进行相应系数的上采样。两个放大模块分别是 2 倍放大模块和 3 倍放大模块(Up Scale to 3 Module, UpTo3), 其输入为原特征图  $F_{LR}$ , 输出为经过上采样的中间结果  $F'$ 。在上采样结束后对  $F'$  进行一次  $3 \times 3$  卷积操作使其降维至 RGB 色彩空间。当实际放大系数  $target\_scale$  与  $scale\_factor$  不匹配时, 通过下采样将其降维到目标维数。 $ratio$  表示放大系数和实际放大系数的比值;  $Bilinear(F, scale)$  表示双线性插值层, 该层输入原特征图  $F$  和放大系数  $scale$ , 输出目标维度的特征图, 当  $scale$  大于 1 时为上采样操作,  $scale$  小于 1 时为下采样操作。双线性插值层输入为经过上采样的特征图  $F'$  和比率  $ratio$ , 由于  $ratio$  恒小于 1, 故该层可以将特征图下采样至目标放大系数。

## 2.4 损失函数

在 SR 重建问题中通常采用均方误差 (Mean Square Error, MSE) 作为损失函数, 其对离群点更为敏感, 但会导致整体图像变得较为平滑。因此, 常用平均绝对误差 (Mean Absolute Error, MAE) 作为损失函数。MAE 相较于 MSE 对离群点的敏感度较小, 在对于图像中轮廓边缘等高频特征有着更好的表现<sup>[17]</sup>。然而, MAE 的函数特征会导致其在零点处不可导, 因此在重建时会对模型产生一些不可预估的影响。

为解决上述问题, 并抑制亚像素层感受野分布不均匀产生的噪声, 本文采用了一种新的损失函数 Charbonnier Loss<sup>[18]</sup>, 其通过添加了一个常量  $\varepsilon$ , 使得该损失函数在零点附近可导, 稳定了模型的训练

过程, 同时加快了模型的收敛速度, 计算公式为:

$$L_C(\hat{I}, I) = \frac{1}{hwc} \sum_{i,j,k} \sqrt{(\hat{I}_{i,j,k} - I_{i,j,k})^2 + \varepsilon^2} \quad (14)$$

其中,  $\hat{I}$  为网络生成的图像, 即 SR;  $I$  为真实图像, 即 HR。 $i, j, k$  分别代表图像长、宽、色彩通道中对应的像素。 $h, w, c$  分别代表图像的长、宽、色彩通道。 $\varepsilon$  是一个很小的常量, 作用是使得损失函数的值稳定且在零点附近可导, 在后续的实验中该值被设置为 0.01。

## 3 实验及结果分析

为了验证本文算法的有效性, 采用了搭载 AMD Ryzen 9 3900X 的电脑进行实验, 64GB 内存, Nvidia GeForce RTX2080 和 Nvidia Tesla M40 24GB 显卡; Windows10 21H1 专业版 64 位操作系统, PyTorch 1.5.1 框架, 并配置了 CUDA 10.2 和 CUDNN 8.0.4 加速, 采用 Matlab 2019b 对算法进行测试。

### 3.1 数据集

本文算法的训练部分采用 SR 重建的公开数据集 DIV2K<sup>[19]</sup>。对于测试部分, 采用了 SR 重建问题中广泛采用的基准数据集: Set5<sup>[20]</sup>, Set14<sup>[21]</sup>, BSD100<sup>[22]</sup> 和 Urban100<sup>[23]</sup>。本文所有的训练基于 RGB 通道进行。

### 3.2 训练参数

本文实验在放大系数为 2、3、4 时对 DIV2K 数据集下每个放大系数进行 2000 轮遍历, 每轮包含 113 次反向传播迭代。在每次迭代中对每张图像随机采用大小为  $60 \times 60$  的 LR 切片作为输入, 对应的 HR 大小为  $60s \times 60s$ ,  $s$  为放大系数, 每次迭代包含 8 张图像。设置初始学习率为 0.0001, 并设置模型训练轮次每 1000 轮后学习率衰减为原来的 0.5 倍, 采用 Charbonnier Loss 损失函数和 Adam 优化器<sup>[24]</sup>来训练模型。在模型中, 采用了 8 个 MSCAB, 其输出均为 64 通道; 瓶颈层的输出为 64 通道。在每个放大系数训练完成后将当前放大系数的模型作为下一个放大系数的预训练模型继续训练。参照文献[11]的策略, 本文实验也不采用任何权重初始化方法。



Table 2 PSNR|SSIM result of ablation experiment of our algorithm under 3x amplification factor

表 2 3 倍放大系数下本文算法消融实验 PSNR|SSIM 结果

MSRB	5x5 多尺度特征提取模块	多尺度特征提取模块	局部空间注意力模块	局部通道注意力模块	自适应上采样模块	模型参数量	Set5	BSD100
√						6.08M	34.13 0.9245	28.97 0.8013
√					√	6.26M	34.14 0.9250	28.99 0.8018
	√					12.40M	34.28 0.9263	29.04 0.8043
		√				9.98M	34.22 0.9256	29.02 0.8036
		√	√			10.58M	34.28 0.9263	29.03 0.8039
		√		√		10.05M	34.31 0.9263	29.04 0.8039
		√		√	√	10.23M	<b>34.40 0.9273</b>	<b>29.09 0.8051</b>

Table 3 PSNR|SSIM result of ablation experiment of our algorithm under 4x amplification factor

表 3 4 倍放大系数下本文算法消融实验 PSNR|SSIM 结果

MSRB	5x5 多尺度特征提取模块	多尺度特征提取模块	局部空间注意力模块	局部通道注意力模块	自适应上采样模块	模型参数量	Set5	BSD100
√						6.33M	31.76 0.8889	27.42 0.7298
√					√	6.26M	31.82 0.8903	27.44 0.7311
	√					12.66M	31.97 0.8923	27.50 0.7340
		√				10.24M	31.93 0.8920	27.49 0.7340
		√	√			10.84M	31.97 0.8923	27.50 0.7341
		√		√		10.31M	32.07 0.8931	27.50 0.7343
		√		√	√	10.23M	<b>32.18 0.8947</b>	<b>27.57 0.7363</b>

### 3.3 评价指标

本文采用峰值信噪比 (Peak Signal to Noise Ratio, PSNR) 和结构相似性指标 (Structural Similarity, SSIM) 作为重建结果的客观评价标准。考虑到比较的相等性, 测试时将图像色彩空间从 RGB 转换为 YCbCr, 并在其 Y 通道进行, 再从图像的每个边缘裁剪  $s$  个像素<sup>[11]</sup>。

### 3.4 实验结果分析

本文将从所提出的模块的有效性、模型收敛性、整数放大系数下的重建性能和非整数放大系数下的重建性能四个方面进行仿真实验。其中, 实验一的对比算法有 MSRN<sup>[11]</sup>; 实验二的对比算法有 FSRCNN 和 MSRN; 实验三的对比如算法有双三次插值、FSRCNN、VDSR、MSRN、SRFBN-S<sup>[21]</sup>和 MRFN<sup>[25]</sup>; 实验四的对比算法为双三次插值和 EDSR<sup>[12]</sup>。

#### 3.4.1 实验一

为了验证本文算法中每个模块的有效性, 同时验证不同注意力模块对算法重建性能的影响, 本文在 3 和 4 倍放大系数下使用本文的训练参数和 MAE 损失函数对每个模块分别进行了消融实验, 并使用

Set5 和 BSD100 对模型性能进行测试。3 倍放大系数下的消融实验结果如表 2 所示, 4 倍放大系数下的消融实验结果如表 3 所示。其中, 多尺度特征提取模块表示采用两个  $3 \times 3$  卷积代替  $5 \times 5$  卷积的特征提取模块;  $5 \times 5$  多尺度特征提取模块表示分别采用  $3 \times 3$  卷积和  $5 \times 5$  卷积的特征提取模块; 局部空间注意力模块表示文献[9]中的空间注意力模块。打勾表示在实验中使用了该模块。

由表 2 和表 3 的实验结果可以看出, 多尺度特征提取模块相对于 MSRB 添加了  $1 \times 1$  卷积, 可以对不同尺度的信息进行了升降维操作。同时, 第一层中的第一个  $1 \times 1$  卷积将输入特征图降维后直接传入第二层中, 可以更加充分地利用输入特征, 因此其性能优于 MSRB。同时, 将  $5 \times 5$  卷积替换为两个  $3 \times 3$  卷积的操作可以在对模块性能有轻微影响的前提下有效减少参数量。在使用  $5 \times 5$  卷积的情况下每个特征提取模块的参数量为 1.50M, 总参数量为 12.40M 和 12.66M。而替换为  $3 \times 3$  卷积后虽然 PSNR 和 SSIM 性能平均下降了 0.03db 和 0.0004, 但每个特征提取模块的参数量为 1.19M, 是原先的 79%;

总参数量下降至 9.98M 和 10.24M，相比平均减少了 2.42M 参数量，是原先的 81%。

注意力模块可以针对通道维度或空间维度进行自适应调整，因此使用注意力模块后算法重建性能好于不使用注意力模块，相较于不使用注意力模块，使用空间注意力模块的算法性能在 PSNR 指标平均提升 0.03db，在 SSIM 指标平均提升 0.0004；使用通道注意力模块的算法性能在 PSNR 指标平均提升 0.07db，在 SSIM 指标平均提升 0.0006。相对于局部空间注意力模块，局部通道注意力模块对每

个特征提取模块的通道进行选择，在强化与特征相关的通道的同时抑制与特征无关的通道，而特征提取模块输出特征图的不同通道对应着不同的特征，因此通道注意力模块使得特征提取模块能够更关注于与特征相关的通道，从而在筛选有用特征的同时屏蔽无用特征。因此相较于局部空间注意力模块，算法性能在 PSNR 指标平均提升了 0.04db，在 SSIM 指标平均提升了 0.0002。图 5 所示为 3 倍放大系数下第 8 个 MSCAB 输出特征图。

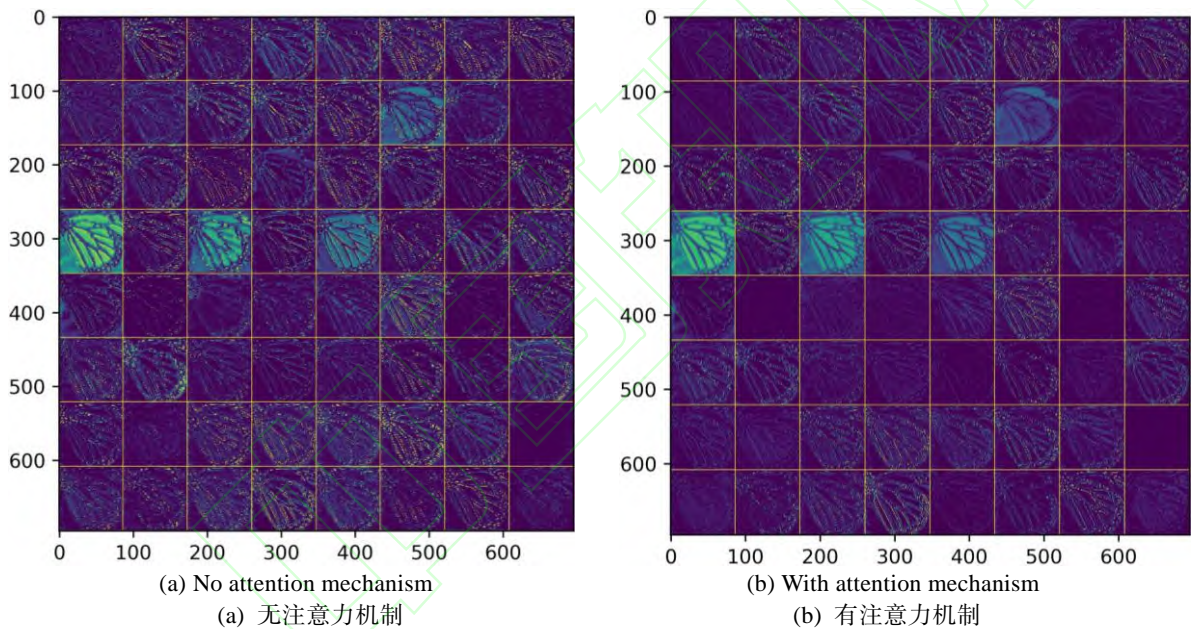


Fig.5 Comparison of channel attention module's effect on output feature map

图 5 通道注意力模块对输出特征图影响对比

从图 5 可以看出，在使用局部通道注意力模块时，每个通道输出的特征图的特征清晰，且对重建没有帮助的通道被屏蔽；而不使用局部通道注意力模块时每个通道的特征图中包含了一些杂乱的特征，对重建会造成一定的影响。因此，局部通道注意力模块对模型的性能有一定的提升作用。

由于自适应上采样模块采用多个亚像素层对图像执行上采样操作，因此在放大系数为 2 或 3 时，其结果与亚像素层相近。但在 4 倍放大系数下自适应上采样模块采用递增式上采样的方法对图像执行两次放大系数为 2 的上采样，避免了高放大系数下上采样

层学习困难的情况发生<sup>[1]</sup>，因此结果好于亚像素层，其 PSNR 和 SSIM 指标平均提升了 0.4db 和 0.0018。

Table 4 PSNR performance comparison of our algorithms under different loss functions

表 4 本文算法在不同损失函数下的 PSNR 性能比较

损失函数	放大系数	Set5	Set14	BSD100	Urban100
MAE	2	38.10	33.64	32.20	32.20
L1_C	2	<b>38.12</b>	<b>33.76</b>	<b>32.21</b>	<b>32.25</b>
MAE	3	34.40	<b>30.32</b>	<b>29.09</b>	28.10
L1_C	3	<b>34.46</b>	<b>30.32</b>	<b>29.09</b>	<b>28.12</b>
MAE	4	32.18	28.55	<b>27.57</b>	26.10
L1_C	4	<b>32.22</b>	<b>28.62</b>	<b>27.57</b>	<b>26.16</b>

为了验证不同损失函数对实验结果的影响,对模型分别采用 Charbonnier Loss 和 MAE 作为损失函数进行训练,并分别进行了测试。表 4 所示为本文算法在不同损失函数下的 PSNR 性能对比。其中, L1\_C 表示 Charbonnier Loss。

从表 4 的结果可以得出采用 Charbonnier Loss 作为损失函数,由于常量的引入,使得损失函数在在零点处可导,因此其在零点附近的稳定性大大增加,所以效果优于使用 MAE 作为损失函数。

为了验证不同数量 MSCAB 对算法性能的影响,分别使用 1、2、4、8、12、16 个 MSCAB 在 3 倍放大系数下使用本文的训练参数和 Charbonnier Loss 进行训练,并使用 Set14 和 BSD100 作为测试集对模型性能进行测试,其结果如表 5 所示。

Table 5 The influence of the number of MSCABs on the performance of the PSNR|SSIM in our algorithm

表 5 MSCAB 数量对本文算法 PSNR|SSIM 性能影响

MSCAB 数量	参数量	Set14	BSD100
1	1.74M	30.00 0.8352	28.86 0.7993
2	2.95M	30.12 0.8373	28.98 0.8022
4	5.38M	30.26 0.8400	29.05 0.8041
8	10.23M	30.32  <b>0.8412</b>	<b>29.09 0.8054</b>
12	15.08M	<b>30.33 0.8409</b>	<b>29.09 0.8051</b>
16	19.94M	30.32 0.8406	29.08 0.8052

由表 5 的结果可知,随着 MSCAB 数量的增加,模型性能逐步提升,但是参数量也随之增长。当 MSCAB 数量为 8 后模型性能到达最大值,进一步增加 MSCAB 数量后其性能几乎不变。因此,本文算法采用 8 个 MSCAB。

### 3.4.2 实验二

为了验证本文算法的收敛性,使用本文的训练集在 2 倍放大系数下对对比算法进行重新训练,两个对比算法都采用了后上采样的方式进行超分辨率重建,与本文算法一致,可以有效地表示出后上采样 SR 重建方法的收敛特性。使用 Set5 作为测试集;PSNR 作为评价指标每 20 轮对算法进行测试。实验结果如图 6 所示。

由图 6 的实验结果可以看出,本文算法的 PSNR

指标在 200 轮之前快速上升,约 400 轮即可收敛,快于 FSRCNN;由于都是采用了多尺度特征提取方法,本文算法的收敛曲线与 MSRN 相似,但是本文算法第一层的第一个  $1 \times 1$  卷积将上一模块的特征直接传播到本模块的第二层中,减轻了模块内特征传播困难的发生,因此网络在 200 轮后 PSNR 性能超过了 MSRN。自适应上采样模块通过亚像素层进行上采样,与 MSRN 一致,对算法收敛性没有影响。综上,本文算法具有良好的收敛性,且不需要任何权重初始化方法,易于训练。

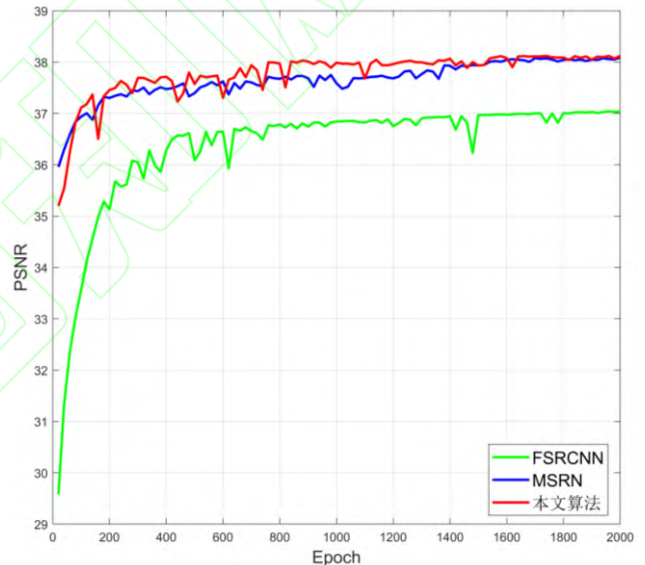


Fig.6 Convergence comparison of different algorithms

图 6 不同算法收敛性对比

### 3.4.3 实验三

为了测试本文算法在不同整数放大系数下的重建效果,对本文算法在 2、3、4 倍放大系数下对各测试集进行重建,并从 PSNR 和 SSIM 角度进行评判。表 6 所示为本文算法在 DIV2K 数据集下与对比算法的 PSNR 和 SSIM 平均值。

由表 6 的实验结果可以看出,由于局部通道注意力的使用,网络可以自适应地强化与特征相关的通道,并抑制无效信息。多个多尺度特征提取模块获得了不同尺度下的特征信息,并通过全局特征融合模块有效地抑制了梯度消失,防止后面几层发生特征消失的现象。因此,本文算法优于当前主流的基于卷积神经网络的超分辨率算法。相较于 MRFN,



本文算法在以上数据集下 PSNR 指标平均提升了 0.34db, SSIM 平均提升了 0.0391。在对图像整体和细节方面的修复获得了一定的提升。

Table 6 PSNR|SSIM performance comparison of various algorithms with different amplification factors

表 6 不同放大系数下各算法的 PSNR|SSIM 性能对比

数据集	放大系数	双三次插值	FSRCNN	VDSR	MSRN	SRFBN-S	MRFN	本文算法
Set5	2	33.66 0.9299	37.06 0.9554	37.53 0.9587	38.08 0.9605	37.78 0.9597	37.98 0.9611	<b>38.12 0.9610</b>
Set14	2	30.24 0.8688	32.76 0.9078	33.03 0.9124	33.74 0.9170	33.35 0.9156	33.41 0.9159	<b>33.76 0.9186</b>
BSD100	2	29.56 0.8431	31.53 0.8912	31.90 0.8960	<b>32.23 0.9013</b>	32.00 0.8970	32.14 0.8997	32.21 0.9001
Urban100	2	26.88 0.8438	29.88 0.9024	30.79 0.9157	32.22 0.9326	31.41 0.9207	31.45 0.9221	<b>32.25 0.9334</b>
Set5	3	30.39 0.8682	33.20 0.9149	33.66 0.9213	34.38 0.9262	34.20 0.9255	34.21 0.9267	<b>34.46 0.9274</b>
Set14	3	27.55 0.7742	29.54 0.8277	29.77 0.8314	<b>30.34 0.8395</b>	30.10 0.8372	30.03 0.8363	30.32 0.8412
BSD100	3	27.21 0.7385	28.55 0.7945	28.82 0.7976	29.08 0.8041	28.96 0.8010	28.99 0.8029	<b>29.09 0.8054</b>
Urban100	3	24.46 0.7411	26.48 0.8175	27.15 0.8315	28.08 0.8554	27.66 0.8415	27.53 0.8389	<b>28.12 0.8576</b>
Set5	4	28.42 0.8104	30.73 0.8601	31.35 0.8838	32.07 0.8903	31.98 0.8923	31.90 0.8916	<b>32.22 0.8952</b>
Set14	4	26.00 0.7027	27.71 0.7488	28.01 0.7674	28.60 0.7751	28.45 0.7779	28.31 0.7746	<b>28.62 0.7816</b>
BSD100	4	25.96 0.6675	26.98 0.7029	27.29 0.7251	27.52 0.7273	27.44 0.7313	27.43 0.7309	<b>27.57 0.7368</b>
Urban100	4	23.14 0.6599	24.62 0.7272	25.18 0.7543	26.04 0.7896	25.71 0.7719	25.46 0.7654	<b>26.12 0.7936</b>

图 7 所示为在 2 倍放大系数下双三次插值、VDSR、MSRN 和本文算法对测试集 Urban100 中的 img\_098 的重建效果；图 8 为以上 4 种方法在 3 倍放大系数下对 Urban100 中 img\_051 的重建效果；图 9 为以上 4 种方法在 4 倍放大系数下对 Urban100 中 img\_060 的重建效果。由于 VDSR 和 MSRN 只训练 YCbCr 通道中的 Y 通道,为了生成 RGB 图像,对 LR 的 Cr、Cb 通道进行双三次插值放大,并与网络生成的 Y 通道合并成为 RGB 图像。

从图 7、图 8 和图 9 可以看出,双三次插值的重建图像较为模糊;VDSR 的重建图像在低放大系数下效果良好,但是在较高放大系数下其图像边缘模糊,且对于图像纹理的重建并不理想;MSRN 在 2、3 倍放大系数下虽然有不错的效果,但在 4 倍放大系数下的重建图像的细节处仍有瑕疵;本文算法的重建图像边缘清晰,且对于图像纹理的重建较为准确,在整体和细节上均具有良好的重建效果。

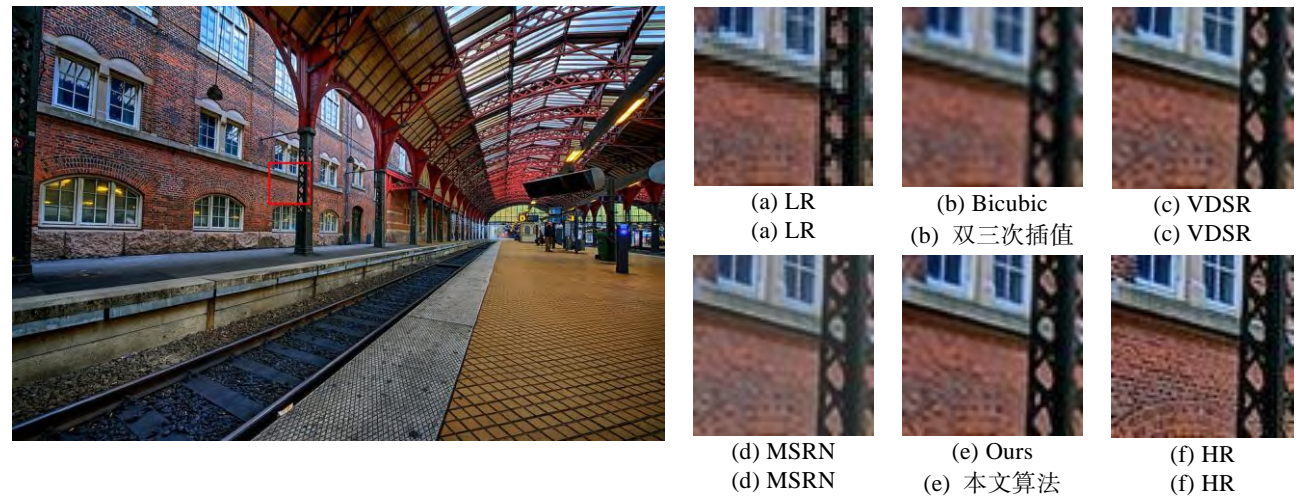


Fig.7 Img\_098 at 2x magnification reconstruction effect comparison



图 7 2 倍放大系数下 img\_098 重建效果对比

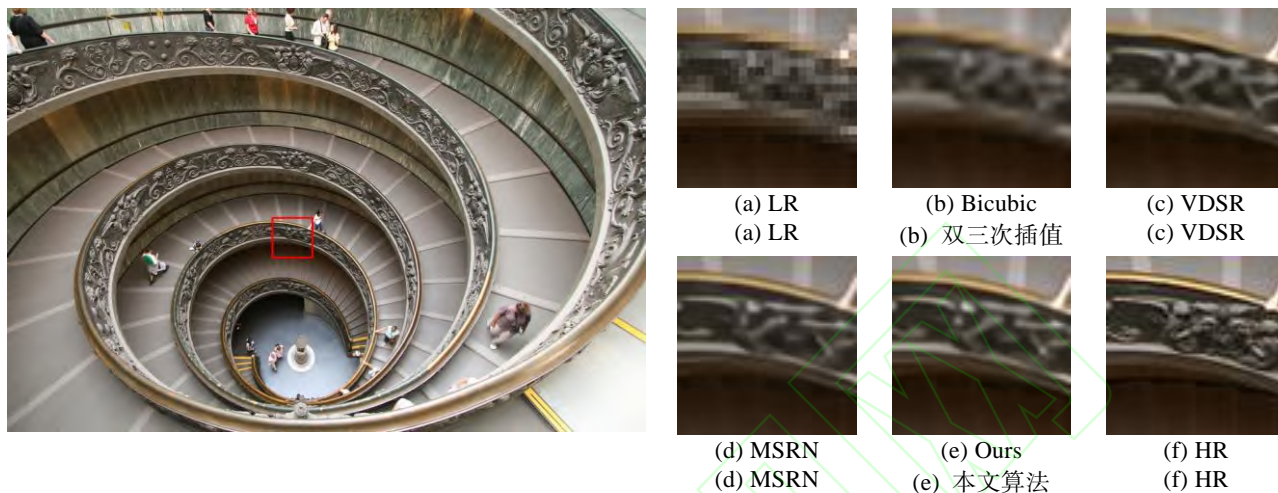


Fig.8 Img\_051 at 3x magnification reconstruction effect comparison

图 8 3 倍放大系数下 img\_051 重建效果对比

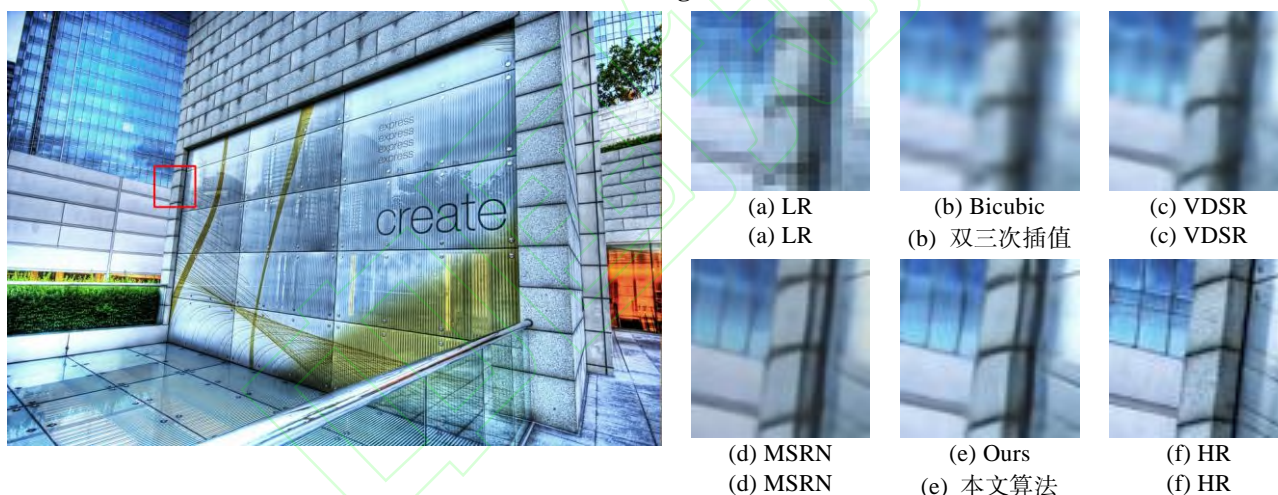


Fig.9 Img\_060 at 4x magnification reconstruction effect comparison

图 9 4 倍放大系数下 img\_060 重建效果对比

此外, 本文算法在不同放大系数下的参数量与不同尺寸输入图像的计算量如表 7 所示。

从表 7 的实验结果看出, 因为自适应上采样模块中包含了 2 个不同放大系数的亚像素层, 因此本文算法的参数量不随放大系数的改变而改变; 在不同放大系数下相同尺寸的输入的计算量有很小的差距, 但随着输入尺寸的增加, 计算量也随之增大, 重建时间也随之变长。为了测试本文算法的实时性, 对 100 张  $176 \times 144$  的图像在 2、3、4 倍放大系数下执行重建, 其每张图像的平均处理时间为 0.059 秒、0.061 秒和 0.064 秒, 即每秒能够处理 16.94 张、16.39 张和 15.63 张图像, 在该分辨率下能够满足监

控等每秒 15 帧视频的实时性要求。

Table 7 The amount of parameters and computation of our algorithm under different size of inputs

表 7 不同输入下本文算法的参数量与计算量

输入尺寸	放大系数	模型参数量(M)	计算量(GFlops)	重建时间(s)
64x64	2	10.23	40.47	0.019
128x128	2	10.23	161.86	0.042
256x256	2	10.23	647.45	0.137
512x512	2	10.23	2589.80	0.527
64x64	3	10.23	41.26	0.019
128x128	3	10.23	165.03	0.043
256x256	3	10.23	660.12	0.141
512x512	3	10.23	2640.47	0.540
64x64	4	10.23	43.58	0.020

128x128	4	10.23	174.30	0.045
256x256	4	10.23	697.21	0.151
512x512	4	10.23	2788.84	0.581

### 3.4.4 实验四

为了验证自适应上采样模块中不同 ratio 值对算法精度的影响,将算法上采样系数强制设置为 2、3 和 4,然后设置不同的 ratio 值实现不同倍数的上采样,使用 PSNR 评价指标对 BSD100 测试集进行测试,实验结果如表 8 所示。

Table 8 Effect of ratio value of adaptive upsampling module on PSNR performance

表 8 自适应上采样模块 ratio 值对 PSNR 性能影响

放大系数	ratio 值	实际放大系数	PSNR
2	0.5	1	34.50
2	0.75	1.5	32.48
3	0.333	1	33.03
3	0.5	1.5	31.61
3	0.667	2	31.77
3	0.833	2.5	28.96
4	0.25	1	32.30
4	0.375	1.5	30.97
4	0.5	2	31.43
4	0.625	2.5	28.55
4	0.75	3	28.99
4	0.875	3.5	28.19

从表 8 的实验结果可以看出,在实际放大系数相同的情况下,强制设置的放大系数越大, ratio 值越小,其 PSNR 性能越低。因为双线性插值在下采

样的过程中会对图像精度造成一定损失,即上采样模块中的 ratio 值越小对精度造成的损失越大,所以本文自适应上采样模块在 ratio 值较小时重建性能较差,随着 ratio 值的逐步提升,重建性能也随之提升。

为了测试本文算法在非整数放大系数下的重建性能,使用 BSD100 数据集分别测试了 1.5、1.8、2、2.2、2.5、2.8、3、3.2、3.5、3.8、4 倍放大系数,并使用只针对 2、3、4 倍放大系数训练的模型。使用 PSNR 作为其评价指标,并将其与双三次插值进行比较,如表 9 所示。

Table 9 PSNR performance comparison under non-integer amplification factor

表 9 非整数放大系数下的 PSNR 性能比较

放大系数	双三次插值	EDSR	本文算法
1.5	32.14	<b>35.30</b>	32.48
1.8	30.38	<b>33.09</b>	31.62
2	29.56	32.04	<b>32.21</b>
2.2	28.87	<b>31.17</b>	29.15
2.5	28.13	<b>30.24</b>	28.96
2.8	27.51	<b>29.47</b>	27.54
3	27.21	29.08	<b>29.09</b>
3.2	26.89	<b>28.69</b>	28.49
3.5	26.42	28.13	<b>28.19</b>
3.8	26.07	<b>27.73</b>	27.50
4	25.96	27.56	<b>27.57</b>

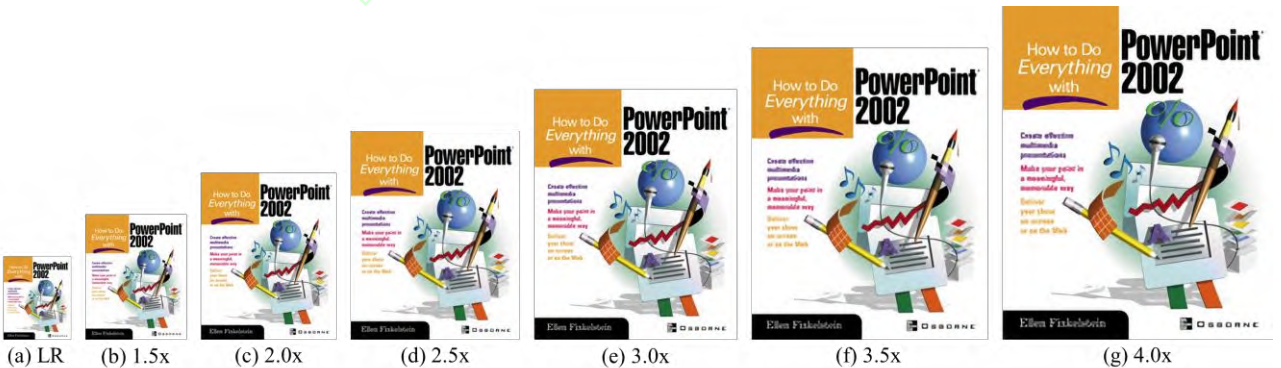


Fig.10 Image ppt3 reconstruction with different amplification factors

图 10 不同放大系数下 ppt3 重建效果

从表 9 的实验结果可以看出,本文算法采用了基于深度学习的模型对图像进行重建,因此,本文

算法的 PSNR 性能超越了基于传统方法的双三次插值。然而,在较小放大系数下自适应上采样的 ratio

值较小,对重建精度造成的损失越大。因此在较小的非整数放大系数下的重建性能低于 EDSR,且随着放大系数的增加其性能逐步追平 EDSR,但是在文献[12]中, Hu 等从 1.1 到 4.0 的放大系数以 0.1 为步长进行训练,需要针对 30 个放大系数进行训练。而本文算法只需要训练比当前非整数放大系数大的整数放大系数即可,即按上述训练策略本文算法只需要针对 3 个放大系数进行训练,是 EDSR 的十分之一,且本文算法在按照上述训练策略训练结束后可以针对 2~4 倍放大系数间的任意整数与非整数放大系数进行重建,极大地缩减了训练时间,提升了模型在非整数放大系数下的灵活性。图 10 展示了本文算法在不同放大系数下对测试集 Set14 中 ppt3 重建效果。

## 4 结束语

针对网络训练难以收敛、不同上采样系数和非整数放大系数不适配的问题,本文根据分而治之的思想提出了一种自适应上采样模块,结合多尺度通道注意力特征提取模块用于在不同尺度下对图像进行提取特征,提出了多尺度自适应上采样的超分辨率算法。本文算法能够充分利用多尺度特征和层次化特征从而获取精确的 SR 图像。使用自适应上采样模块,灵活地将 LR 放大至目标分辨率。在仿真实验中,本文算法在不使用任何初始化方法时仍然具有良好的收敛性。同时,与 VDSR、MSRN、SRFBN-S 等主流网络相比,本文算法不仅更好地实现了超分辨率图像重建效果,而且解决了对非整数放大系数的适配问题。对比双三次插值方法,本文算法在处理非整数放大系数的问题时具有更高的性能;对比 EDSR 算法,本文算法在非整数放大系数下不需要针对每一个放大系数进行训练,减少了

训练的复杂度,并且增加了模型的灵活性。未来将针对现有的上采样方法进行优化,进一步提升模型的灵活性和非整数放大系数下的重建性能,同时对特征提取模块继续优化,在提高超分辨率重建性能的同时使得网络更加轻量化。

## 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, 42(3): 1-22.
- [2] Wang R, Zhang Y H, et al. Image super-resolution re-construction based on CNN[J]. Computer Engineering and Design, 2019, 40 (6): 1654-1659.  
王容, 张永辉, 张健, 等. 基于 CNN 的图像超分辨率重建方法[J]. 计算机工程与设计, 2019, 40(6): 1654-1659.
- [3] Zhu S, Li Y. Single image super-resolution under multi-frame method[J]. Signal, Image and Video Processing, 2019, 13(2): 331-339.
- [4] Deng L, Zhou Z, Xu G, et al. TV2++: a novel spatial-temporal total variation for super resolution with exponential-type norm[J]. EURASIP Journal on Wireless Communications and Networking, 2020, 16(4): 223-240.
- [5] Khan A, Sohail A, Zahoor U, et al. A survey of the recent architectures of deep convolutional neural networks[J]. Artificial Intelligence Review, 2020, 53(8): 5455-5516.
- [6] Zhang K, Zuo W, Zhang L. Deep plug-and-play super-resolution for arbitrary blur kernels[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Los Alamitos: IEEE Computer Society Press, 2019: 1671-1681.
- [7] Xi Z H, Hou C Y, Yuan K P, et al. Super-resolution reconstruction of accelerated image based on deep residual network[J]. Acta Optica Sinica, 2019, 39(2): 89- 98.  
席志红, 侯彩燕, 袁昆鹏, 等. 基于深层残差网络的加速图像超分辨率重建[J]. 光学学报, 2019, 39(2): 89-98.
- [8] Fan P P, Dong X C, Li T, et al. Super-resolution reconstruction of depth map based on non-local means constraint[J]. Journal of Computer-Aided Design & Computer Graphics, 2020, 32(10): 1671-1678.

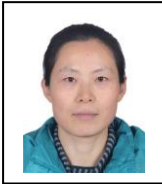


- 范佩佩, 董秀成, 李滔, 等. 基于非局部均值约束的深度图像超分辨率重建[J]. 计算机辅助设计与图形学学报, 2020, 32(10): 1671-1678.
- [9] Liu Y, Wang Y, Li N, et al. An attention-based approach for single image super resolution[C]//Proceedings of the 24th International Conference on Pattern Recognition, Beijing, China, Aug 20-24 2018. Los Alamitos: IEEE Computer Society Press, 2018: 2777-2784.
- [10] 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. Honolulu, Jul 21-26 2017. Los Alamitos: IEEE Computer Society Press, 2017: 4681-4690.
- [11] Li J, Fang F, Mei K, et al. Multi-scale residual network for im-age super-resolution[C]//Proceedings of the European Conference on Computer Vision. Munich, Germany, Sep 8-14 2018. Switzerland: Springer International Publishing Cham, 2018: 517-532.
- [12] Hu X, Mu H, Zhang X, et al. Meta-SR: A magnification-arbitrary network for super-resolution[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA, Jun 16-20 2019. Los Alamitos: IEEE Computer Society Press, 2019: 1575-1584.
- [13] Wang B, Liu Y, Tang S, et al. Pedestrian detection with fusion of multi-models and intra-frame information[J]. Journal of Computer-Aided Design & Computer Graphics, 2017, 29(3): 444-449.
- 王斌, 刘洋, 唐胜, 等. 融合多模型和帧间信息的行人检测算法[J]. 计算机辅助设计与图形学学报, 2017, 29(3): 444-449.
- [14] 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. Munich, Germany, Sep 8-14 2018. Switzerland: Springer International Publishing Cham, 2018: 286-301.
- [15] Han B, Niu G, Yu X, et al. Sigua: Forgetting may make learning with noisy labels more robust[C]//Proceedings of the 37th International Conference on Machine Learning. Vienna, Austria, Jul 13-18 2020. Cambridge: Proceedings of Machine Learning Research, 2020: 4006-4016.
- [16] Peng C, Zhang X, Yu G, et al. Large kernel matters--improve semantic segmentation by global convolutional network[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Honolulu, Jul 21-26 2017. Los Alamitos: IEEE Computer Society Press, 2017: 4353-4361.
- [17] 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. Honolulu, Jul 21-26 2017. Los Alamitos: IEEE Computer Society Press, 2017: 1132-1140.
- [18] Lai W S, Huang J B, Ahuja N, et al. Fast and accurate image super-resolution with deep laplacian pyramid networks[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2018, 41(11): 2599-2613.
- [19] Agustsson E, Timofte R. Ntire 2017 challenge on single image super-resolution: Dataset and study[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Honolulu, Jul 21-26 2017. Los Alamitos: IEEE Computer Society Press, 2017: 1110-1121.
- [20] Zhang Y, Fan Q, Bao F, et al. Single-image super-resolution based on rational fractal interpolation[J]. IEEE Transactions on Image Processing, 2018, 27(8): 3782-3797.
- [21] Li Z, Yang J, Liu Z, et al. Feedback network for image super-resolution[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA, Jun 16-20 2019. Los Alamitos: IEEE Computer Society Press, 2019: 3867-3876.
- [22] Li X G, Feng X X, Li J X. Single image super-resolution reconstruction based on multi-scale residual network [J]. Computer Engineering and Applications, 2021, 57(7): 215-221.
- 李现国, 冯欣欣, 李建雄. 多尺度残差网络的单幅图像超分辨率重建[J]. 计算机工程与应用, 2021, 57(7): 215-221.
- [23] Lugmayr A, Danelljan M, Timofte R. Ntire 2020 challenge on real-world image super-resolution: Methods and results[C]//Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops. Seattle, USA, Jun 16-21 2020. Los Alamitos: IEEE Computer Society Press, 2020: 494-495.
- [24] Park T, Liu M Y, Wang T C, et al. Semantic image synthesis with spatially-adaptive normalization[C]// Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. Long Beach, USA, Jun 16-20 2019. Los Alamitos: IEEE Computer Society Press, 2019: 2337-2346.



[25] He Z, Cao Y, Du L, et al. MRFN: Multi-receptive-field network for fast and accurate single image su-

per-resolution[J]. IEEE Transactions on Multimedia, 2020, 22(4): 1042-1054.



吕佳 (1978-), 女, 四川达州人, 博士, 教授, 主要研究方向为机器学习、数据挖掘。CCF 会员。

LYU Jia, was born in 1978, She received the Ph.D. degree from Inner Mongolia University in 2012. Now she is a professor and Ph.D. supervisor at Chongqing Normal University. Her research interests include machine learning, data mining, etc.



许鹏程 (1997-), 男, 江苏南京人, 重庆师范大学硕士研究生, 主要研究方向为深度学习、图像处理。CCF 学生会员。

XU Pengcheng was born in 1997, He is an M.S. candidate in Chongqing Normal University. His research interests include deep learning, image processing, etc.