基于轻量级多尺度聚合网络的红外图像电子变倍

刘 馨[1]，张 斌[2]
（1. 兰州工业学院 电子信息工程学院，甘肃 兰州 730030；
2. 甘肃省无线电管理委员会，甘肃 兰州 730030）

摘要：为了解决光电领域中低分辨率红外图像影响观瞄的问题，构建一种轻量级多尺度聚合网络算法来增强电子变倍时中心区域图像分辨率，该算法首先使用不同大小的尺度核从图像中提取特征信息，并利用浅层残差结构将局部多尺度残差特征有效聚合在一起，以获得更强大的特征表示能力；然后采用基于对比度感知的通道注意层来聚合更多尺度特征信息，最终重构出具有丰富细节而清晰的高分辨率红外图像。仿真实验结果表明，本文所提出的电子变倍方法在不引入额外参数的情况下能够提取出精细的多尺度特征信息，获得清晰的重建效果。

关键词：光电系统；红外图像；电子变倍；多尺度聚合；对比度感知；轻量级
中图分类号：TP37 **文献标识码：**A **文章编号：**1001-8891(2025)04-0445-08

Electronic Zooming of Infrared Image Based on Lightweight Multi-scale Aggregation Network

LIU Xin[1]，ZHANG Bin[2]
(1. *School of Electronic Information Engineering*, *Lanzhou Institute of Technology*, *Lanzhou* 730030, *China*;
2. *Gansu Provincial Radio Regulatory Commission*, *Lanzhou* 730030, *China*)

Abstract: To solve the problem of low-resolution infrared images affecting viewing and aiming in the photoelectric field, a lightweight multi-scale aggregation network is proposed to enhance the resolution of the central region when the IR image is zoomed. First, the algorithm uses scale kernels of different sizes to extract feature information and employs a shallow residual structure to effectively aggregate local multi-scale residual features, thereby obtaining stronger feature representation capability. Then, a channel attention layer based on contrast perception is used to aggregate more multi-scale feature information. Finally, a high-resolution infrared image with rich detail and clarity is reconstructed. Simulation results show that the zooming method can extract fine multi-scale feature information without introducing additional parameters and can produce clear reconstruction results.

Keywords: photoelectric system, infrared image, electronic zooming, multi-scale aggregation, contrast-aware, lightweight

0 引言

变倍技术是光电探测设备的重要功能，有助于提升目标清晰度与观瞄精度[1]。随着光电技术的发展，目前有光学变倍与电子变倍两种方式对目标区域进行放大，前者通过开窗或变焦方式提升中心区域的分辨率，后者则是通过信息处理技术把区域放大。在实际观瞄过程中，虽然改变焦距缩小视场有助于提升目标的探测细节，但也降低了操作手对周围环境的感知，尤其是跟踪过程切换视场可能导致跟踪丢失[2-3]。因此，大多数光电系统是采用电子变倍方式对局部区域进行观测或跟踪点锁定。

电子变倍是通过图像处理算法来实现局部放大，不需要改变光电系统的硬件设施，无需增加制造成本，而光学变倍则需要复杂的变焦系统，也大大降低了系统的可靠性[4]。虽然电子变倍技术具有明显的优势，但

收稿日期：2023-03-02；**修订日期：**2023-08-23.
作者简介：刘馨（1983-），女，甘肃兰州人，硕士，副教授，主要研究方向：模式识别，深度学习，图像处理。E-mail:79697234@qq.com。
基金项目：2025年甘肃省高校教师创新基金项目（2025A-235）；中国高校产学研创新基金资助课题（2024HY033）；甘肃省重点人才项目（2025RCXM014）；2023年甘肃省高校教师创新基金项目（2023B-239）；2018年兰州工业学院青年科技创新项目（2018K-019）。

该技术不会改变图像的信噪比，却增加了视觉噪声[5]。目前，电子变倍技术大都采用最邻近点插值、双线性插值、双立方插值等传统算法[6]来提高变倍后的图像分辨率，但是放大图像仍然是模糊不清，存在不光滑或者失真等情况。虽然一些学者提出反演算法[7]可放大图像同时弥补图像的细节损失，在一定程度上可以降低局部放大对图像质量的影响。这是由于该类方法利用有限的像素信息区推算未知像素信息，其效果往往非常有限。

近年来，随着深度卷积网络的出现，国内外学者提出了大量基于深度学习的单帧图像超分辨重建方法，并应用于电子变倍，已经取得了显著的放大重建效果。Dong 等人[8]首次将卷积神经网络用于超分辨率图像重建，提出了一个三层卷积网络来模拟高/低分辨率图像间的映射函数，其重建性能超越了传统插值算法，带动了深度网络在图像重建领域的发展。杨明等人[9]提出了基于多级深度网络的高清晰度红外电子变倍算法，能够生成具有丰富细节而清晰的高分辨率红外图像，增强了目标锁定的精度。为进一步提高超分辨率算法的重建性能，加快网络快速收敛效率。Chen 等人[10]认为 SwinIR 结构拥有更强的局部表征能力，能够使用更少的信息来达到更高的性能，设计了一种混合注意力 transformer 模型，结合自注意力、通道注意力和设计的重叠交叉注意力，激活更多的像素以更好地重建，有效地改善了视觉效果。Liang 等人[11]利用 SR 的结果和真实图像之间的残差局部方差区分伪纹理和真实细节的有效特征，设计出一个深度框架来规范和稳定模型训练过程，有助于特征信息在网络中流动，惩罚伪纹理，而不影响实现细节。西南技术物理研究所唐中和团队是国内首次在光电系统中采用深度学习实现电子变倍工程应用，提升了目标锁定的精度[12]。

尽管单帧超分辨率重建算法已经取得了很大进展，但目前网络模型主要关注如何调整网络结构的深度和宽度以达到更高的性能，忽略了多尺度特征间的潜在耦合关系，导致特征表征能力有限。与此同时，随着网络深度与宽度的增加，模型的复杂度与数据量也呈现指数型增加[13]。为了解决上述问题，本文提出了一种轻量级多尺度信息聚合网络，该网络首先使用不同大小的尺度核从图像中提取特征信息，并利用浅层残差结构将局部多尺度残差特征有效聚合在一起，以获得更强大的特征表示能力；然后采用基于对比度感知的通道注意层来聚合更多尺度特征信息。仿真实验结果表明，本文所提出的电子变倍方法在不引入额外参数的情况下能够提取出精细的多尺度特征信息，最终重构出具有丰富细节而清晰的高分辨率红外图像。

1 相关技术

国内外学者已经提到了许多单帧超分辨网络及其网络结构，本章将对本文采用的相关技术进行集中讨论。

1.1 多分支模块

为了改善数据流并使网络训练变得容易，学者们设计出由多分支模块组成的网络结构。残差网络[14]是通过堆叠包含残差映射分支和恒等映射分支的模块；GoogLeNet[15]的 Inception 模块由具有不同尺度核的并行卷积分支组成，然后在模块末端连接这些分支以增加宽度和信息融合。尽管这些重建方法在网络拓扑和训练过程中有所不同，但它们都有一个关键特征：在每个模块中包含多个分支。另一方面，多分支网络可以被视为具有不同深度的多个子网络的融合，从而提高重建性能。考虑到多分支模块的优点，本文也利用不同大小的尺度核从图像中提取特征信息，通过在多个并行分支上提取并融合不同尺度的信息，以实现微弱细节的重建。

1.2 级联结构

当变倍因子较大时，超分辨过程中重建高频细节变得非常困难。为了简化网络结构与降低重建难度，学者们提出了基于模块级联的网络结构[16]。级联结构有两种方式：逐阶段优化和逐阶段放大[16-18]。前者将前一级的输出作为输入，并将基准结构作为每个级的优化目标，最终目的是使每个阶段的预测误差最小化，以逐渐接近最优目标。Chen 等人[17]提出了一种级联线性回归框架，以从粗到细的方式重新确定预测的特征块，然后合并所有预测的块以生成高分辨率图像。后者结构每一级将用较小的比例因子放大低分辨率图像，其输出被进一步馈送到下一级，直到获取所需的变倍尺度。本文设计了一种具备 $2^{\times}$、$3^{\times}$和 $4^{\times}$的网络架构，采用逐阶段放大以获得相应的变倍结果。

2 轻量级多尺度聚合网络

本章将详细描述本文所采用的网络架构，尤其是所改进的多尺度特征聚合模块。

2.1 网络架构

如图 1 所示，本文提出的电子变倍网络模块由卷积模块（Cov），多尺度聚合模块（multi-scale aggregation module, MAM），特征融合模块（feature fusion module, FFM）和图像重建模块（image reconstruction module, IRM）组成，其基本流程如图 1

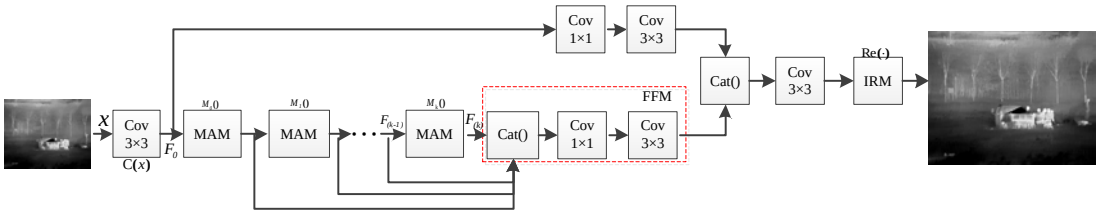


图 1 本文改进的电子变倍网络

Fig.1 The improved electronic zooming network

所示：首先，将非制冷探测器输出的低分辨图像输入尺寸为 3×3 的卷积模块，提取红外图像的初始特征信息；然后将多个多尺度聚合模块堆叠起来进行特征融合，提取变倍区域中最本质的特征信息；最后采用图像重建模块实现电子变倍。

给定输入红外图像 x，卷积模块的数学模型可以表述为 $F_0=C(x)$，其中 $C(x)$表示特征提取函数，F_0 表示提取的卷积特征。多尺度聚合模块是采用多个不同尺度倍率分支来提取多尺度信息，通过将提取到不同尺度的卷积特征以链式连接的方式堆叠在一起，逐步细化特征的表征能力，其过程可以表示为 $F_k=M_k(F_{(k-1)})$，其中 F_{k-1} 与 F_k 分别表示第 k 个模块的输入和输出向量。初始特征信息经过多个堆叠的多尺度聚合模块提取特征后，所有中间特征被合并到一起，然后通过卷积核为 1×1 卷积模块进行融合，其多尺度融合过程可以表述为 $F=C_{3\times3}[C_{1\times1}(\mathrm{Cat}(F_1, F_2,\cdots,F_n))]$，其中 Cat(·)是所有特征沿通道的级联操作，$C_{1\times1}$ 和 $C_{3\times3}$ 分别表示 1×1 卷积融合算子和 3×3 卷积平滑算子；最后，使用 3×3 卷积层对聚合后的多尺度信息特征进行平滑处理，得到变倍区域的放大图像，重构过程可以描述为 $y=\mathrm{Re}[F+C_{3\times3}(C_{1\times1}(F_0))]$，其中 Re(·)表示重构函数，$y$ 表示本文所提网络输出的变倍图像。可以看出，变倍重建过程只使用了 3×3 卷积与非参数的子像素卷积，其中 3×3 卷积输出的特征个数需要设置成变倍倍数 r 的平方，并与训练数据中高分辨率图像尺寸一致，而子像素卷积是将像素进行重新排列得到电子变倍后的放大图。

2.2 多尺度特征聚合模块

受 IMDB 网络启发[18]，本文引入了多尺度残差学习，通过逐层次提取有用特征以提高多尺度聚合网络[19]的特征表征能力，而无需引入其他附加参数。图 2 展示了多尺度聚合模块的结构，其中浅层残差结构提取到特征被分成两部分，一部分输入到不同卷积核大小(3×3 和 5×5)的卷积层中提取不同尺度的特征；另一部分直接输出，并与不同尺度特征叠加在一起，再输入到 ReLU 层；然后通过级联操作将获得的主干网络特征与 1×1 卷积特征进行融合，最终得到不同像素区域的多尺度特征。

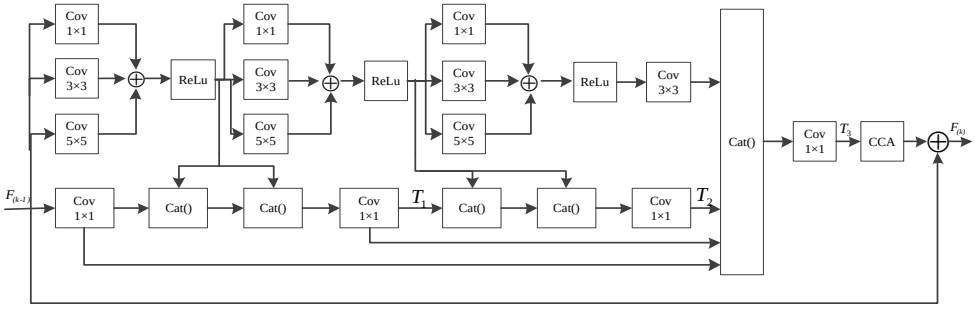


图 2 多尺度聚合模块

Fig. 2 Multi-scale aggregation module

给定输入到浅层残差结构的特征为 F_{i-1}，浅层残差结构的特征提取过程可描述为如下等式：

$$F_i=\mathrm{ReLU}[F_{i-1}+C_{3\times3}(F_{i-1})+C_{5\times5}(F_{i-1})],\quad i\in1,2,\cdots,n \tag{1}$$

式中：F_i 表示第 i 个浅层残差结构提取的特征。可以看出，输入到该模块的特征经过多尺度特征学习后又经过等式(1)进行处理，去除卷积结果中的负值，保留正值不变，最大限度降低梯度消失现象。为了便于将不同级联结构的特征聚合，设计了一个多尺度聚合模块，该过程可描述为如下等式：

$$T_i=\mathrm{Cat}[\mathrm{Cat}(F_i, C_{1\times1}(T_{i-1})), F_i],\quad i\in1,2,\cdots,n-1 \tag{2}$$

$$T_n=\mathrm{Cat}[C_{1\times1}(F_0), T_1, \cdots,T_n, C_{3\times3}(F_n)] \tag{3}$$

2.3 轻量化参数设计

多重相关特征映射与多通道数据级联将增大参数量，使得网络模型遭受巨大的内存消耗，同时也增加

优化的难度。另一方面，由于光电设备功耗限制，只能在有限的硬件资源下实现高效的电子变倍。高效的轻量级网络设计的关键是如何充分利用有限的特征来生成更具有代表性的特征。目前，常见的注意机制包含基于对比度的通道注意力（channel-wise context aggregation，CCA）[20]，增强空间注意力（efficient spatial attention，ESA）[21]等，这些都是多分支拓扑。尽管平均池化/最大池化能捕获高层次视觉问题中的全局特征信息，并提高图像的峰值信噪比（PSNR）值，但它缺少与图像细节相关的结构、纹理和边缘的信息，不足以提高图像的结构相似性（SSIM）。因此，为了提高变倍性能，并加快推理速度和降低内存消耗，本文采用基于对比度的通道注意机制来增强特征表示，如图 3 所示。另一方面，顺序操作能够避免频繁地访问之前的状态信息和额外的节点，有利于减少内存消耗和资源开销。因此，本文选用了 1×1 卷积层对提取的多尺度特征 T_n 进行降维，并将得到的特征输入到基于对比度感知的通道注意层来聚合更多多尺度特征信息，在不引入额外参数的情况下，有助于增强图像细节。

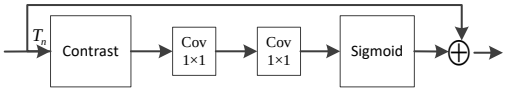


图 3 基于对比度的通道注意力

Fig.3 Contrast-aware channel attention

2.4 图像重构及模型优化

多尺度聚合模块获取的图像特征经过基于对比度感知的通道注意层处理后最终得到图像的深度特征，通过将其与原始特征 F_0 叠加，有助于提取出更加精细的多尺度特征信息。

$$F=P(C_{1\times1}(T_n))+F_0 \tag{4}$$

式中：$P(\cdot)$表示本文选择的基于对比度感知的通道注意层（CCA）。通过叠加多个多尺度特征聚合模块，并按照一定的融合策略，最终得到低分辨率图像的特征信息，并在重构函数 Re(·)的优化下得到变倍后的高分辨率图像 y。为了便于对整个网络的性能进行优化，采用 L_1 损失函数进行网络训练与性能优化。给定训练集中包含 N 个高/低分辨率图形组成的样本对 $\{I_i^{LR},I_i^{HR}\}_{i=1}^{N}$，损失函数可以表示为如下等式：

$$L(\theta)=\frac{1}{N}\sum_{i=1}^{N}\left|\mathrm{Re}(x)-I_{i=1}^{\mathrm{HR}}\right| \tag{5}$$

式中：θ为训练过程的模型超参数；I_i^{LR} 与 I_i^{HR} 分别表示低分辨率图像及其对应的高分辨率图像。

3 实验结果对比与分析

3.1 数据集和评价指标

深度学习是一种依赖于海量数据解决现实问题的数据驱动方法，通过对训练数据进行整合和提炼形成自动化的学习模型。与现有其他超分辨网络一样，本文仍然从基准数据集与自建数据集角度进行定性定量分析。值得注意的是，国内外学术机构并未开源出红外超分辨率数据集。本文所提出的改进网络并不仅仅针对红外图像。为了便于定性定量对比分析，实验选用 DIV2K 作为基准数据集[22]。DIV2K 数据集是一个高质量基准数据集，包含 1000 张具有不同场景的图像，并按照 8:1:1 的比例进行训练、验证与测试。本实验将实验数据集全部转成灰度图像，并裁剪成高低分辨率对应的子图对（sub-image-pair），其中高分辨率图像大小为 640×512。自建数据集则是利用项目组成员采集的大小为 640×512 的红外图像，经双三次插值得到不同分辨率的低分辨率图像组成的图像对。

实验采用定性定量方法评价电子变倍后放大图像的质量，定性评价是一种主观评价方法，通过视觉感知来判别红外图像的清晰度、纹理结构等；定量评价则是通过量化指标来客观评价图像质量，本文采用图像处理领域常用的峰值信噪比（PSNR）和结构相似性（SSIM）进行分析[23]。

3.2 参数设置

在给定高分辨率训练图像的情况下，训练使用的低分辨率图像都是通过双三次插值法生成。所有低分辨图像随机裁剪出大小为 64×64 的图像块作为训练样本，并采用文献[24]提出的方法进行数据扩增。本文所提模型采用 Adam 优化器进行优化，具有收敛速度快、调参容易等优点，其中动量参数β_1＝0.9 和β_2＝0.999；初始学习速率设置为 2×10^{-4}，每 2×10^{5} 次迭代参数减半。训练过程中，批量大小设置为 16。堆叠的多尺度特征聚合模块数量设置为 6。本文实验硬件环境为 4.2 GHz Intel i7-7700K CPU，64 G RAM 和 NVIDIA 2080Ti GPU；软件环境则使用 PyTorch 深度框架进行网络模型构建。

3.3 轻量级消融分析

本文设计的网络首先使用不同大小的尺度核从图像中提取特征信息，并利用浅层残差结构将局部多尺度残差特征有效聚合在一起，以获得更强大的特征表示能力。为了验证该思路的有效性，设计了多尺度特征聚合模块的数量对变倍性能影响的消融分析。图 4 显示了当 k 从 1 增加到 8 时，本文网络所获得的 PSNR/SSIM，其中 k＝1 表明单尺度下提取目标特征。可以看出，随着 k 从 1 增加到 4 时，峰值信噪比（PSNR）的值在逐渐增加，且网络参数量也在增加。基于网络

参数和变倍性能之间的权衡，堆叠的多尺度特征聚合模块数量被设置为 4。

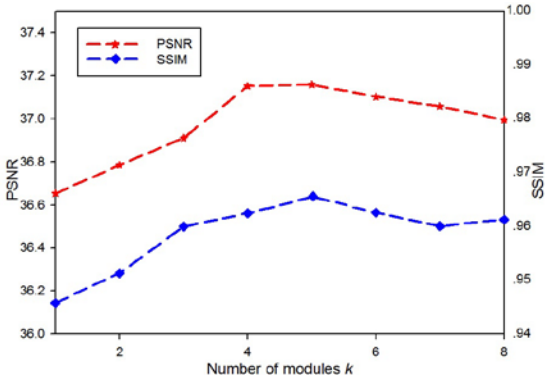


图 4 不同多尺度聚合模块数量的性能对比

Fig.4 Performance comparison for different number of multi-scale aggregation modules

对比实验也去掉了 1×1 卷积层，直接将提取的多尺度特征输入到基于对比度感知的通道注意层来聚合单尺度的特征信息。结果表明，采用 1×1 卷积能够避免频繁地访问之前的状态信息和额外的节点，有利于减少内存消耗和资源开销。

3.4 定性和定量比较

在超分辨率重建算法中，大多数方法使用 Bicubic 退化模型来模拟低分辨率图像，即将高分辨率图像进行双三次插值，缩小一定倍数，生成对应的低分辨率图像。本文采用的高分辨率图像为高德 EYAS 系列制冷红外机芯获取的 1280×1024 图像，并裁剪成 640×512，然后进行双三次插值处理，缩小到对应倍数，构建出高/低分辨率图像对。为了评估所设计的深度超分辨网络重建性能，实验选用了 Bicubic，SRCNN[8]，VDSR[25]，EDSR[26]，GAN-SR[27]算法进行对比，且对比算法都能在 Github 获取源码，其超分辨倍数选用 $2^{\times}$、$3^{\times}$和 $4^{\times}$。由于部分算法处理 RGB 数据，实验将灰度图像 RGB 化，不改变图像灰度以及对比网络模型。

表 1 展示了典型图像及其测试数据集的超分辨率重建结果对比（PSNR(dB)/SSIM），其中 img_01、img_02、img_03 来自自建数据集的红外图像，而 DIV2K、Self-built 来自 DIV2K 数据集。可以发现，本文提出的网络具有较好的重建优势。在 $2^{\times}$变倍模式下，所改进的模型在 DIV2K 测试数据上具有较好的性能，其中 PSNR 与 SSIM 分别 35.35/0.9641，而 Bicubic 的结果分别是 31.24/0.9387。值得注意的是，光电变倍主要采用 Bicubic 插值的方法来放大图像，一般只是 $2^{\times}$变倍。本文所采用的方法比 Bicubic 提升了 4.11 dB/0.0254，大大优于 Bicubic 方法。在 $3^{\times}$变倍模式下，本文所提模型在 DIV2K 测试数据上与最优 GAN-SR 结果相当，但在自建数据集上优于所有对比算法。在 $4^{\times}$变倍模式下，本文所提模型也有较好的性能，但提升效果并不明显，部分结果不如 GAN-SR。尤其是对于自建数据集中部分图像细节并不明显的图像，其变倍尺度越大，效果反而越差。从变倍结果也能看出，所提模型将平滑区域放大出较多纹理。另外，部分城市背景下的红外图像具有较多纹理细节，所提网络的重建结果较好，定量指标也有较大改善，这主要是由于该网络可以有效地学习图像的细节轮廓信息，因此

表 1 典型图像及其测试数据集的超分辨重建结果对比（PSNR(dB)/SSIM），其中黑色加粗表示最优结果

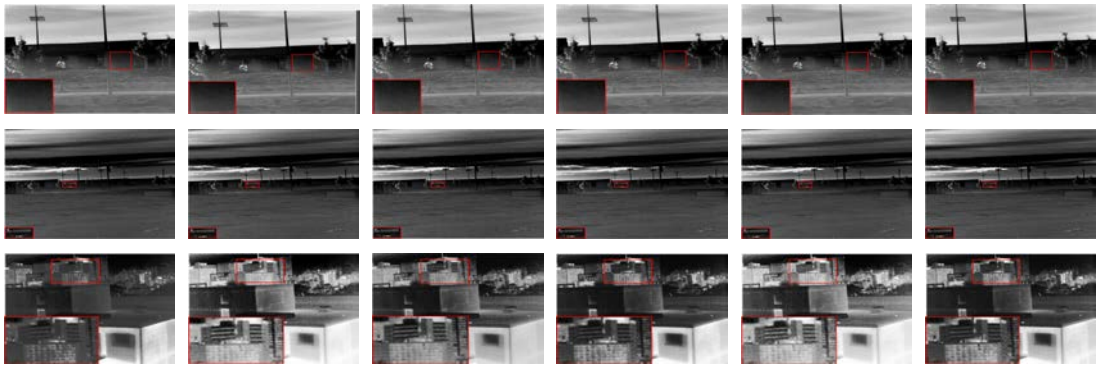
Table 1 Comparison of super-resolution reconstruction results for typical images and their test-data sets, where black represents the best result

Dataset	Scale	Bicubic	SRCNN	VDSR	EDSR	GAN-SR	Proposed
img_01	2×	33.12/0.9219	35.73/0.9544	37.52/0.9522	**38.17**/0.9600	38.08/95.35	**38.17/0.9631**
	3×	30.38/0.8682	32.57/0.9087	33.67/0.9215	34.56/0.9248	34.77/**0.9301**	**34.79**/0.9287
	4×	28.41/0.8102	30.28/0.8709	31.35/0.8838	**32.15**/0.8917	31.82/0.8991	32.11/**0.9011**
img_02	2×	29.56/0.8432	32.35/0.9045	33.03/0.9124	33.87/0.9153	33.85/0.9199	**33.91/0.9201**
	3×	27.20/0.7384	30.110.8726	29.78/0.8425	30.25/**0.8954**	30.45/0.8452	**30.53**/0.8467
	4×	25.92/0.6679	27.48/0.7427	28.11/0.7701	28.75/0.7846	28.71/0.7911	**28.91/0.7912**
img_03	2×	26.87/0.8403	29.12/0.8965	30.76/0.9144	32.49/**0.9384**	32.84/0.9125	**33.02**/0.9362
	3×	24.69/0.7439	25.99/0.8075	27.01/0.8257	28.75/0.8647	28.78/0.8655	**28.95/0.8702**
	4×	23.28/0.6544	24.81/0.7204	25.22/0.7528	26.62/**0.8034**	26.68/0.8017	**26.71/0.8034**
DIV2K	2×	31.24/0.9387	33.24/0.9451	33.38/0.9614	34.95/0.9648	35.25/**0.9684**	**35.35**/0.9641
	3×	28.12/0.8907	29.64/0.9127	30.14/0.9215	31.26/0.9341	31.29/**0.9354**	**31.42/0.9354**
	4×	26.58/0.8486	27.74/0.8705	28.14/0.8803	29.35/0.9014	**29.38**/0.9011	29.24/**0.9024**
Self-built	2×	30.15/0.9217	31.53/0.9197	31.82/0.9201	32.07/0.9297	32.21/**0.9311**	**32.32**/0.9308
	3×	27.15/0.8295	28.24/0.8336	29.11/0.9012	29.25/0.8924	29.29/0.8174	**29.67/0.8947**
	4×	26.64/0.8074	27.14/0.8391	27.32/0.8765	27.71/0.8607	**27.86**/0.8653	27.84/**0.8724**

定量指标较优。

图 5 展示了部分代表性图像在 3×变倍模式重建结果，其中从上到下图像分别是 img_01、img_02 与 img_03。img_01 中具有一条明显的直线，大多数对比算法在重建直线时会产生模糊伪影。Bicubic 作为一种常用的插值方法，放大目标区域的同时也平滑了边缘细节信息。相比之下，本文重建网络可以稍微改善此现象并恢复更多细节信息。对于“img_02”，大多数对比算法无法很好地重建细节信息。一些算法可以重建出错误的纹理细节，而其他方法可以重建出微弱的纹理区域。基于 GAN 的超分辨率重建生成细节的同时容易产生伪影，抑制伪影的同时会破坏真实细节，对于纹理丰富区域具有较好效果，但平滑区域伪影较多。相比之下，本文所提方法不仅可以正确地重建图像的纹理方向，而且与基准图具有较好一致性。img_03 的变倍结果中，大多数比较方法都会产生模糊的伪影。GAN-SR 可以重建所有的规则纹理轮廓，但图像中的一些地方会出现轮廓失真。SRCNN，EDSR 与 VDSR 的变倍结果要么无法重建部分轮廓，要么重建的图像模糊。本文所设计的网络不仅可以重建轮廓方向，而且重建的图像更清晰。定性结果对比表明多个多尺度聚合模块堆叠起来进行特征融合，可以提取变倍区域中最本质的特征信息。

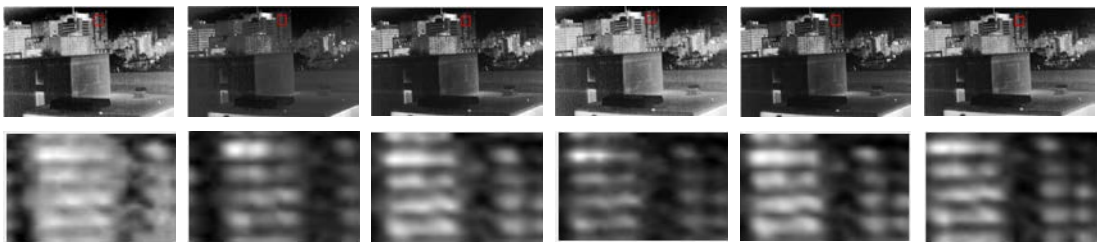
图 6 与图 7 分别是不同对比算法在 3×与 4×模式下的重建结果，其中 Bicubic 的结果非常模糊，几乎不能看出目标边缘，尤其是重构的图像比较模糊；SRCNN 与 VDSR 能看到部分细节信息，但有较重伪影。放大倍数越大，GAN-SR 生成的细节越多，但与实际细节不一致。部分区域甚至改变了图像细节信息。虽然该算法能一定程度上获得较好的视觉效果，但不利于观瞄应用。本文所提算法对平滑区域也产生较多伪影，尤其是红外图像不均匀噪声越强，伪影越严重，



(a) Bicubic (b) SRCNN (c) VDSR (d) EDSR (e) GAN-SR (f) Proposed

图 5 2×变倍模式下不同模型的重建结果

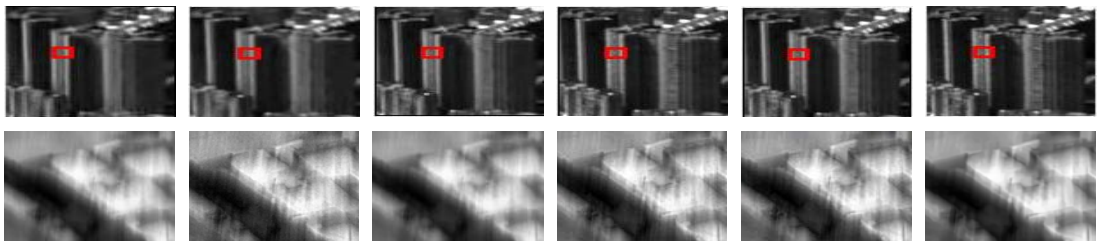
Fig.5 Visual results for a scale factor of 2×on the different models



(a) Bicubic (b) SRCNN (c) VDSR (d) EDSR (e) GAN-SR (f) Proposed

图 6 3×变倍模式下不同模型的重建结果

Fig. 6 Visual results for a scale factor of 3× on the different models



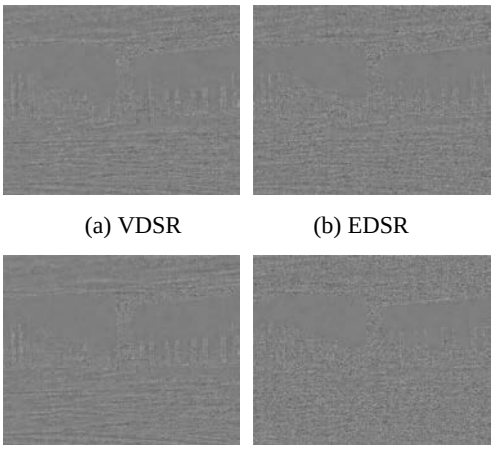
(a) Bicubic (b) SRCNN (c) VDSR (d) EDSR (e) GAN-SR (f) Proposed

图 7 4×变倍模式下不同模型的重建结果

Fig. 7 Visual results for a scale factor of 4×on the different models

但对于强纹理区域的变倍效果就好很多。

为了便于分析重建纹理的细节特征，图 8 展示了 $4^{\times}$变倍模式下重建结果与基准结果的差值图。理论上，理想的差值图应该近似为高斯噪声。然而，低分辨率图像电子变倍过程会预测大量未知像素，尤其是复杂结构区域的重建结果并不准确，呈现一些结构特征。因此，差值图中结构特征越明显，重建效果越差。图 8 中，GAN-SR 具有大量明显结构特征，尤其是天空中区域大量划痕，使得该区域不再平滑。值得注意的是，GAN-SR 对规则区域的重建结果较好，使得背景复杂的红外图像重建的 PSNR/SSIM 定量指标高。EDSR 的结果也存在较多纹理信息，但大部分位于强边缘附近，平滑区域纹理细节不明显。本文方法重建结果中也存在纹理细节，但视觉上比 EDSR 弱一些。总的说来，多尺度特征聚合模块比其他方法恢复了更多的纹理细节信息，这也进一步证明了该方法更关注图像的高频信息。



(a) VDSR　(b) EDSR

(c) GAN-SR　(d) Proposed

图 8　不同算法的重建差值图对比

Fig. 8　Comparison of reconstruction difference maps of different models

4　结论

本文提出了一种多尺度特征聚合模块来实现高精度的超分辨率重建，不仅使用不同卷积核大小的卷积层来提取多尺度图像特征，而且将提取的多尺度特征有效地聚合在一起以获得精细的图像特征。同时，本文也设计了一种浅层残差结构以进一步提高了超分辨率重建的性能。仿真使用结果表明，本文所提网络在定量和定性评估方面都优于最优对比方法。下一步，将继续探索将该方法应用于光电系统电子变倍功能的研制，并实现任意倍数的放大。

参考文献：

[1] 邓伟, 陈建飞, 张胜. 深层跳线残差网络热红外图像超分辨重建[J]. **电光与控制**, 2023, **30**(3): 27-32.
DENG Wei, CHEN Jianfei, ZHANG Sheng. Super-resolution reconstruction of thermal infrared images based on deep skip connection residual network[J]. *Electronics Optics and Control*, 2023, **30**(3): 27-32.

[2] 唐中和, 王升哲, 霍建亮, 等. 基于局部结构变换域稀疏外观模型的目标跟踪[J]. **电视技术**, 2017, **41**(Z2): 140-146.
TANG Zhonghe, WANG Shengzhe, HUO Jianliang, et al. The target tracking based on sparse appearance model of local structure transformed domain[J]. *The TV Engineering*, 2017, **41**(Z2): 140-146.

[3] 左岑, 杨秀杰, 张捷, 等. 基于轻量级金字塔密集残差网络的红外图像超分辨增强[J]. **红外技术**, 2021, **43**(3): 251-257.
ZUO Cen, YANG Xiujie, ZHANG Jie, et al. The super-resolution enhancement of the infrared images based on the lightweight and pyramid of dense residual network[J]. *Infrared Technology*, 2021, **43**(3): 251-257.

[4] ZHANG X, LI C, MENG Q, et al. Infrared image super resolution by combining compressive sensing and deep learning[J]. *Sensors*, 2018, **18**(8): 2587.

[5] ZOU Y, ZHANG L, LIU C, et al. Super-resolution reconstruction of infrared images based on a convolutional neural network with skip connections[J]. *Optics and Lasers in Engineering*, 2021, **146**: 106717.

[6] HUANG Y, JIANG Z, LAN R, et al. Infrared image super-resolution via transfer learning and PSRGAN[J]. *IEEE Signal Processing Letters*, 2021, **28**: 982-986.

[7] Mallat S, YU G. Super-resolution with sparse mixing estimators[J]. *IEEE Transactions on Image Processing*, 2010, **19**(11): 2889-2900.

[8] 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*, 2016, **38**(2): 295-307.

[9] 杨明, 王璇, 高宏伟. 基于多级深度网络的高清晰度红外电子变倍算法[J]. **弹箭与制导学报**, 2020, **40**(4): 14-19.
YANG Ming, WANG Xuan, GAO Hongwei. The high-definition infrared electronic zoom algorithm based on multi-level deep network[J]. *Journal of Projectiles, Rockets, Missiles and Guidance*, 2020, **40**(4): 14-19.

[10] CHEN X, WANG X, ZHOU J, et al. Activating more pixels in image super-resolution transformer[C]// *IEEE Conf. Comput. Vis. Pattern Recognit*.(CVPR), 2023: 236-248.

[11] LIANG J, ZENG H, ZHANG L. Details or artifacts: a locally discriminative learning approach to realistic image super-resolution[C]//2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition* (CVPR), 2022: 5647-5656.

[12] HE Z, TANG S, YANG J, et al. Cascaded deep networks with multiple receptive fields for infrared image super-resolution[J]. *IEEE Transactions on Circuits and Systems for Video Technology*, 2018, **29**(8): 2310-2322.

[13] 左芝勇. 采用分裂 Bregman 的空间邻近目标红外超分辨算法[J]. **电讯**

技术, 2020, **60**(7): 778-784.
ZUO Zhiyong. An infrared super-resolution algorithm using split bregman for closely spaced objects[J]. *Telecommunication Engineering*, 2020, **60**(7): 778-784.

[14] QIN J, HUANGY, WEN W. Multi-scale feature fusion residual network for single image super-resolution[J]. *Neuro Computing*, 2020, **379**: 334-342.

[15] HAI H, LI P, ZOU N, et al. End-to-end super-resolution for remote-sensing images using an improved multi-scale residual network[J]. *Remote Sensing*, 2021, **13**(4): 666-672.

[16] Szegedy C, Liu W, Jia Y, et al. Going deeper with convolutions[C]//*Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2015: 1-9.

[17] WANG Yingqian, YING Xinyi, WANG Longguang, et al. Light field super-resolution via adaptive feature remixing[J]. *European Conference on Computer Visio*n, 2022: 230-247.

[18] CHEN C, GONG D, WANG H, et al. Learning spatial attention for face super-resolution[J]. *IEEE Transactions on Image Processing*, 2021, **30**: 1219-1231.

[19] HU Jie, SHEN Li, SUN Gang. Deep convolutional neural networks with merge-and-run mappings[C]//I*nternational Journal of Computer Vision*, 2018: 1379-1398.

[20] LIU Zhiwei, ZHOU Lei, GAO Yuting et al. Large-scale object detection in the wild from imbalanced multi-labels[C]//*IEEE Transactions on Image Processing*, 2022: 1234-1247.

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

[22] LI Y, CAO J, LI Z, et al. Lightweight single image super-resolution with dense connection distillation network[J]. *ACM Transactions on Multimedia Computing, Communications, and Applications* (TOMM), 2021, **17**(1s): 1-17.

[23] 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 Workshop*s, 2017: 126-135.

[24] YANG W, ZHANG X, TIAN Y, et al. Deep learning for single image super-resolution: a brief review[J]. *IEEE Transactions on Multimedia*, 2019, **21**(12): 3106-3121.

[25] ZHANG Kai, LIANG Jingyun, Van Gool Luc, et al. Training real-world blind super-resolution with pure synthetic data[C]//*Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 2022: 1401-1411.

[26] Kim J, Lee J K, Lee K M. Accurate image super-resolution using very deep convolutional networks[C]//*IEEE Conf. Comput. Vis. Pattern Recognit*.(CVPR), 2016: 1646-1654.

[27] Lim B, Son S, Kim H, et al. Enhanced deep residual networks for single image super-resolution[C]//*IEEE Conf. Comput. Vis. Pattern Recogni*t.(CVPR), 2017: 1025-1032.