

文章编号:1007-2780(2022)11-1476-12

## 基于多尺度特征融合网络的傅里叶叠层成像

宋东翰<sup>1,2</sup>, 王 斌<sup>1\*</sup>, 朱友强<sup>1</sup>, 刘 鑫<sup>3</sup>

(1. 中国科学院 长春光学精密机械与物理研究所, 吉林 长春 130033;

2. 中国科学院大学, 北京 100049;

3. 拉彭兰塔-拉赫蒂理工大学 工程科学学院 计算机视觉与模式识别实验室,  
芬兰 拉赫蒂 15210)

**摘要:**傅里叶叠层成像是一种实现光学系统高分辨率、大视场成像的技术。传统FP方法的高分辨率重建过程需要较高的孔径重叠率,导致采集图像数量较多,采样效率低。此外,FP重建算法的复杂度高,重建时间长。针对以上问题,本文结合深度学习,提出一种基于多尺度特征融合网络的傅里叶叠层成像算法,通过改进的特征金字塔卷积神经网络,能够从稀疏采样的低分辨率振幅图像中提取特征信息并进行融合,实现超分辨率的复图像重建。实验结果表明,在相同采样条件下,与传统方法相比,本文提出的深度学习算法提高了图像重建的质量,减少了90%以上的重建时间,并且对高斯噪声的鲁棒性较高。所提出的方法能够将相邻频域子孔径间的重叠率从50%降低至25%,减少50%的采集图像数量,大幅提高采样效率。

**关 键 词:**计算成像;傅里叶叠层成像;特征金字塔;稠密连接;通道注意力

**中图分类号:**O438;TP391 **文献标识码:**A **doi:**10.37188/CJLCD.2022-0094

## Fourier ptychography based on multi-scale feature fusion network

SONG Dong-han<sup>1,2</sup>, WANG Bin<sup>1\*</sup>, ZHU You-qiang<sup>1</sup>, LIU Xin<sup>3</sup>

(1. *Changchun Institute of Optics, Fine Mechanics and Physics, Chinese Academy of Sciences, Changchun 130033, China;*

2. *University of Chinese Academy of Sciences, Beijing 100049, China;*

3. *Computer Vision and Pattern Recognition Laboratory, School of Engineering Science, Lappeenranta-Lahti University of Technology, Lahti 15210, Finland)*

**Abstract:** Fourier Ptychography (FP) is a technology of achieving high-resolution, large field-of-view imaging of optical system. However, the high-resolution reconstruction based on traditional FP methods requires a high aperture overlap ratio, resulting in a large number of captured images and low sampling efficiency. In addition, the FP reconstruction algorithm has high complexity and long reconstruction time. Aiming at solving these problems of the FP, this paper proposes a deep learning algorithm based on multi-scale feature fusion network. Through the improved feature pyramid module, the feature information can be extracted from multiple low-resolution images captured by the FP imaging system, and the information is

收稿日期:2022-03-23;修订日期:2022-04-20.

基金项目:国家重点研发计划(No. 2020YFA0714102)

Supported by National Key R&D Program of China (No. 2020YFA0714102)

\*通信联系人, E-mail:175969722@qq.com

fused to achieve super-resolution reconstruction. Experimental results show that compared with traditional methods, the deep learning algorithm proposed in this paper improves the quality of image reconstruction, reduces the reconstruction time by 90%, and is more robust to Gaussian noise. In addition, the proposed method can reduce the overlap ratio between sub-apertures from 50% to 25% in frequency domain, and reduce the number of captured images by 50%, greatly improving the sampling efficiency.

**Key words:** computational imaging; fourier ptychography; feature pyramid; dense connectivity; channel attention

## 1 引言

在光学成像系统中,空间带宽积(Space Band width Product, SBP)<sup>[1]</sup>表示系统传递信息的能力。然而高分辨率和大视场在光学成像系统中相互制约<sup>[2]</sup>,使得高空间带宽积的成像需求难以满足。传统的解决方法是通过机械设备移动高数值孔径(Numerical Aperture, NA)的透镜,获取物体不同区域的高分辨率信息<sup>[3]</sup>,并利用算法进行图像拼接,实现高空间带宽积的成像。但此方法对设备的精度要求高、成本高昂、成像速度慢。傅里叶叠层成像(Fourier Ptychography, FP)<sup>[4]</sup>是一种兼顾高分辨率和大视场的计算成像技术。在傅里叶叠层显微成像(Fourier Ptychography Microscope, FPM)<sup>[5]</sup>中,LED阵列提供角度变化的平面波对物体进行照明,将超过光学系统衍射极限的频域信息转移至显微物镜内,得到一系列低分辨率图像。最后通过相位恢复算法,将低分辨率图像包含的信息在频域中拼接、融合,重建出物体的高分辨率、大视场复图像。该技术已在数字成像、生物显微等领域被广泛研究与应用<sup>[6-8]</sup>。

尽管FP技术已经发展了近10年,但是依然面临两个问题:一是采样效率低。在FP的重建过程中,样本频谱上相邻子孔径间需要存在一定程度的重叠,才能保证最终的频谱收敛。如果不存在重叠,则采集的图像间毫无联系,最终会导致相位恢复问题的固有模糊性<sup>[9]</sup>。因此,高重叠率的低分辨率图像集是FP高质量重建的基础。二是重建效率低。FP的重建算法是在空间域和频域内对样本的振幅和相位信息进行交替更新,收敛速度较慢,算法迭代时间长。为了弥补以上缺陷,目前多项研究成果被提出。在提升采样效率方面,Jason等<sup>[10]</sup>将相机阵列引入宏观FP成像系统中,可以同时采集所有低分辨率样本,但成

像系统过于复杂,设计困难。Tian等<sup>[11]</sup>对FPM成像系统的LED点光源进行编码控制,每次采样时随机使用多个LED照明,大幅度降低了采样次数。然而其重建算法是基于二阶梯度的序列更新并且需要对采集图像进行低分辨率图像解耦,复杂度高、运行时间长。Kappeler等<sup>[12]</sup>结合深度学习的方法,尝试降低FP采样中的孔径重叠率。但由于提出的模型只包含3层卷积,因此难以提取精细的特征信息,重建效果不理想。此外,在加快重建速度方面,Bian等<sup>[13]</sup>设计的Wirtinger flow优化法,在重建过程中舍弃部分高频信息以提升算法运行效率。Zhang等<sup>[14]</sup>提出的多尺度深度残差网络,采用端对端的模型实现快速重建。Jiang等<sup>[15]</sup>利用神经网络的反向传播机制求解样本的高分辨率复振幅,加速了传统算法的迭代过程。虽然这些方法都从不同程度上减少了重建时间,但都需要采集大量的样本,采样效率低。

随着近年来人工智能科学的快速发展,神经网络算法已经在光学成像的多个领域被广泛应用<sup>[16-18]</sup>。在一定的成像条件下,神经网络模型可以通过大量的训练数据,学习输入和输出图像之间的映射关系,能够快速且准确地处理图像信息。因此,本文提出基于多尺度特征融合的FP成像算法(Fourier Ptychography Based on Multi-feature Fusion Network, FPMFN),利用深度学习中的特征金字塔模块<sup>[19]</sup>,从FP采样的多张低分辨率图像中提取不同尺度的特征信息。为了确保特征信息的准确性和完整性,加入稠密连接<sup>[20]</sup>以增强信息在网络间的流动,并且引入通道注意力机制<sup>[21]</sup>,使网络更好地分配高低频信息的训练权重。特征金字塔模块的解码器部分能够将不同尺度的特征信息进行融合,采用亚像素卷积<sup>[22]</sup>完成图像从低分辨率到高分辨率的转变。通过不同孔径重叠率下的对比实验,验证了在孔径重叠

率相同时, FPMFN 比传统算法有更好的重建效果, 而且重建速度更快。FPMFN 在 25% 重叠率时的重建结果能够媲美传统算法在 50% 重叠率时的结果。实验结果表明, 本文算法能有效减少重建所需的样本数, 提升采样的效率, 并大幅减少重建所需的时间。

## 2 傅里叶叠层成像原理

FP 计算成像方法由成像过程和重建过程组成。典型的 FP 成像模块如图 1 所示, 样品与 LED 阵列的距离足够远, 因此不同角度的入射波可以近似地视为平面波。样品的出射场经过透镜传输, 被图像传感器 (Charge Coupled Device, CCD) 接收, 输出低分辨率强度图像。最后经过重建过程, 将低分辨率强度图像恢复为高分辨率复图像。

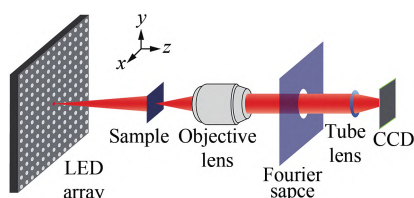


图 1 FP 的成像模块装置图

Fig. 1 Imaging module diagram of FP

### 2.1 FP 的成像过程

在成像过程中, 使用复振幅分布  $o(r)$  表示样品, 其中  $r=(x, y)$  是空间域的二维坐标。在第  $m$  个 LED 点光源照明的条件下, 样品的出射场为  $o(r) \cdot \exp(ik_m r)$ 。其中“ $\cdot$ ”表示点乘,  $k=(k_x^m, k_y^m)$  表示照明光源的入射波矢。经过傅里叶变换, 能够获得样品出射场的频谱, 此过程表示为:

$$O(k - k_m) = F\{o(r) \cdot \exp(ik_m r)\}, \quad (1)$$

式中:  $F$  为傅里叶变换运算符,  $k=(k_x, k_y)$  为傅里叶域的二维坐标。经过物镜时, 透镜会对样品频谱进行低通滤波, 最终 CCD 接收的低分辨率图像可以表示为:

$$I_m(r) = |F^{-1}\{O(k - k_m) \cdot P(k)\}|^2, \quad (2)$$

式中:  $I_m(r)$  为 CCD 平面接收到的强度信息,  $F^{-1}$  为傅里叶逆变换运算符,  $P(k)$  为物镜的相干传递函数 (Coherent Transfer Function, CTF), 可以表示为:

$$P(k) = \begin{cases} 1, & k_x^2 + k_y^2 < \left(\frac{2\pi}{\lambda} \text{NA}\right)^2 \\ 0, & k_x^2 + k_y^2 \geq \left(\frac{2\pi}{\lambda} \text{NA}\right)^2 \end{cases}, \quad (3)$$

式中:  $\frac{2\pi}{\lambda} \text{NA}$  为系统的截止频率, NA 为物镜的数值孔径,  $\lambda$  为 LED 照明光波的波长。

### 2.2 FP 的图像重建过程

交替投影法 (Alternate Project, AP) 是经典的 FP 重建算法, 图 2 为算法的重建流程。

以 AP 算法为基准的 FP 图像重建流程如下:

(1) 使用中心位置的 LED 作为照明光源, 采集低分辨率图像并进行上采样处理。处理后的图像经过傅里叶变换后得到高分辨率频谱  $O_0(k)$ , 作为初始估计解。当迭代进行至第  $j$  轮, 高分辨率频谱表示为  $O_j(k)$ 。

(2) 选择第  $m$  个 LED 照明样品, 得到此时低分辨率的傅里叶频谱为:

$$O_{m,j}(k) = P(k) \cdot O_j(k - k_m). \quad (4)$$

$O_{m,j}(k)$  经过逆傅里叶变换得到空间域的低分辨率图像  $o_{m,j}(r) = F^{-1}\{O_{m,j}(k)\}$ 。

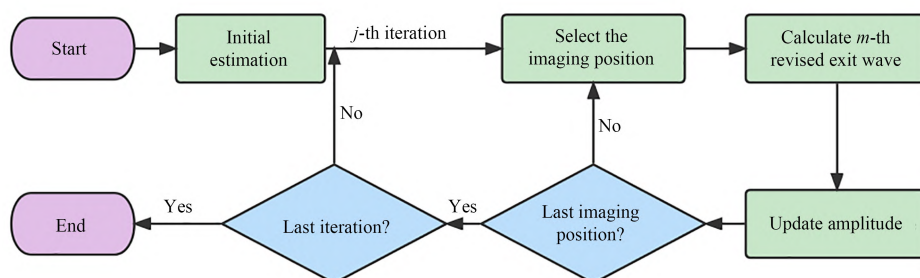


图 2 FP 重建流程图

Fig. 2 Flowchart of FP reconstruction

(3)不改变  $O_{m,j}(k)$  的相位信息,用 CCD 采集的低分辨率强度图像  $I_m(r)$ ,更新  $O_{m,j}(k)$  的振幅,表示为:

$$o'_{m,j}(r) = \sqrt{I_m(r)} \frac{o_{m,j}(r)}{|o_{m,j}(r)|}. \quad (5)$$

(4)振幅更新后,将当前照明条件下子孔径区域内的频谱替换为

$$o'_{m,j}(k) = F[o'_{m,j}(r)], \quad (6)$$

(5)重复上述(2)~(4)的过程,直到 LED 阵列中所有光源完成照明。

(6)重复上述(2)~(5)的过程,直到完成所有迭代轮次,将最后更新的频谱逆傅里叶变换至空域获取高分辨率复图像。

### 3 多尺度特征融合的特征神经网络

现有的 FP 神经网络大多通过单层的特征信息变换实现输入与输出图像之间的映射求解。这种网络结构对高频信息的关注度较低,需要从大量的低分辨率图像中提取充足的高频信息。

因此,本文基于深度学习中的特征金字塔网络,结合稠密连接、通道注意力(Channel Attention, CA)模块和亚像素卷积(Sub-Pixel Convolution, SP-Conv),提出一种多尺度特征融合的特征神经网络。网络为串行的跳层连接结构,利用卷积层在不同的感受野下提取并融合多尺度的特征信息。神经网络的输入为多张低分辨率强度图像,最终输出是单张高分辨率复图像。

#### 3.1 FP 神经网络的框架

为了获取网络的输入,需要先在训练集中随机选择样本的振幅图像  $I_G$  和相位图像  $\varphi_G$ ,合成复图像  $C_G$ 。经过 FP 的成像模块,得到低分辨率图像集  $I_{L,m}$ 。如图 3 所示,网络整体由浅层卷积模块、特征金字塔模块、上采样模块以及重建模块组成。网络结构如下:

首先将  $I_{L,m}$  输入浅层卷积模块,而浅层卷积模块中包含两个卷积层:第一个  $1 \times 1$  卷积层将输入数据的通道数进行压缩,减少神经网络的参数量和后续卷积核的尺寸,提高神经网络的训练效率;第二个  $3 \times 3$  卷积层可以从压缩后的数据中生成含有粗略信息的特征图。

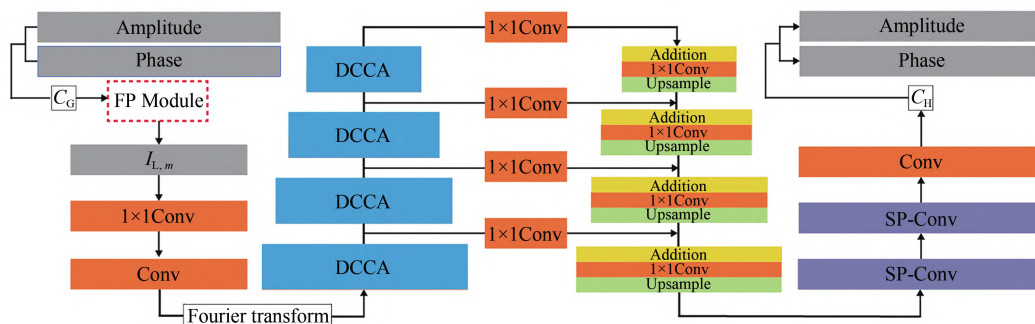


图 3 本文提出的网络模型结构图

Fig. 3 Structure diagram of the proposed network model

从浅层卷积模块的输出中提取振幅和相位信息,输入至特征金字塔模块。该模块分为编码器和解码器。在编码器中,多层的稠密连接通道注意力(Densely connected with channel attention, DCCA)模块进行叠加,自底向顶地提取不同尺度的特征图,采样尺度因子为 2,每层特征图的行列分辨率都是前一层的 1/2。顶层的特征图经过了更深层的卷积神经网络,因此包含丰富的语义信息,但位置信息是不准确的;相反,准确的位置信息需要从底层的特征图中获取。为了尽可能

完整地利用特征信息,编码器中每一层的特征图都会经过  $1 \times 1$  卷积。在解码器中,按位相加进行融合,不同特征图的尺度对齐通过双三次插值上采样实现。

解码后的全局特征图需要进行超分辨处理,达到与  $C_G$  相同的尺寸。常见的上采样选择有插值上采样、反卷积上采样等,为了实现更有效的超分辨,参考高效亚像素卷积神经网络(Efficient Sub-Pixel Convolutional Neural Network, ESPCNN)<sup>[22]</sup>。本文选用计算量更小且超分性能更佳的 SP-Conv 对

全局特征图进行上采样,通过2次亚像素卷积,最终达到4倍上采样的效果。

最后使用一个 $3\times 3$ 卷积层作为重建模块,完成特征图到高分辨率复图像 $C_H$ 的转变。整体的神经网络以输出的 $C_H$ 尽可能接近 $C_G$ 为目标进行训练。

### 3.2 稠密连接通道注意力模块

常见的卷积神经网络通常选择多个卷积层叠加的设计,以此提升网络提取特征的能力。这种方式需要依赖大量的数据进行训练,且模型的参数量庞大。因此,本文提出DCCA模块,通过加强卷积层之间的联系,提高特征信息的利用率,获取高质量的重建效果。模块包含两个部分:稠密连接层和CA模块,如图4所示。

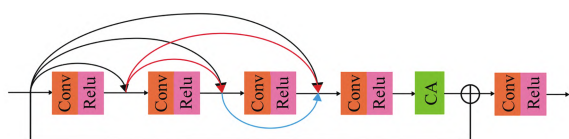


图4 稠密连接通道注意力模块

Fig. 4 Densely connected channel attention module

稠密连接层能够使特征和梯度信息在卷积层之间更加有效地传递。局部特征信息通过卷积和非线性激活函数Relu后,与之前所有层提取的特征信息进行融合作为当前卷积层的输入,有效提高网络的泛化能力,同时避免了梯度消失的问题。

假设 $Q_n$ 表示第 $n$ 层的卷积和激活函数运算,该层的输出为:

$$y_n = Q_n([y_0, y_1, \dots, y_{n-1}]) \quad (7)$$

本文通过消融实验确定神经网络对4种不同尺度的特征图进行提取和融合时的重建效果最好,因此设置特征金字塔模块的编码器部分由4层DCCA组成。考虑到深层的稠密连接会造成计算复杂度过高,不利于整个神经网络的训练与测试,本文选用5个卷积层和1个CA模块共同组成单个DCCA模块。DCCA的稠密连接层能从粗略的特征图中提取更详细的特征信息,不同的高低频分量信息被储存在不同的通道。由于低频特征更容易被神经网络表达,因此均匀分配特征通道的权重会降低神经网络对高频特征的重建质量。为了增强对特征通道权重的分配能力,在

稠密连接层之后设置了CA模块,使神经网络在训练的过程中自适应地调整不同特征通道的权重,从而提升对高频特征信息的关注。模块整体采用残差结构对输入和CA模块的输出进行连接,提升信息在模型内部的流动,改善网络的映射能力,同时可以确保在多个卷积叠加的过程中,网络不会因为复杂的结构而产生退化。最后经过一个 $3\times 3$ 卷积和激活函数,得到DCCA的输出。

CA模块包含5部分:全局池化层、特征通道压缩层、特征通道扩张层、Relu非线性激活层和Sigmoid非线性激活层,模块结构如图5所示。形状为 $H\times W\times C$ 的特征张量首先经过全局池化层,在 $H$ 和 $W$ 的维度上被压缩为 $1\times 1\times C$ 的向量,这种形式有利于在通道维度上计算权重。接着用 $1\times 1$ 的卷积在通道维度上进行压缩,假设压缩系数为 $t(t<1)$ ,压缩后的向量形状变为 $1\times 1\times tC$ 。经过激活函数Relu后,再通过第二个 $1\times 1$ 卷积进行通道扩张,形状恢复为 $1\times 1\times C$ 。然后通过激活函数Sigmoid,将不同通道的权重系数用 $(0,1)$ 的数值表示。最后对原始的特征张量按权重系数进行通道维度的加权分配。

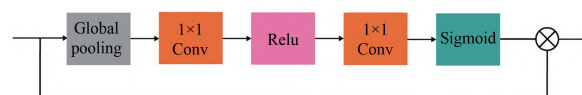


图5 通道注意力模块

Fig. 5 Channel attention module

### 3.3 网络参数及损失函数

对于本文提出的神经网络, $1\times 1$ 尺寸的卷积核应用于浅层卷积模块中输入特征的通道压缩、特征金字塔模块中编解码器的横向连接以及CA模块中的通道压缩和扩张。网络中的其他卷积核的尺寸都为 $3\times 3$ 。浅层卷积模块、DCCA模块的稠密链接层、重建模块在进行 $3\times 3$ 卷积计算前,对输入做0数值填充处理,以此维持特征图的尺寸。网络的输入和输出都是单通道的灰度图像,但同样适用于彩色图像的重建任务。网络的损失函数使用平均绝对值误差(L1损失),假设本文的算法为 $R$ ,高分辨率复图像的像素总数为 $N$ ,训练的优化目标可以表示为

$$\text{Loss} = \frac{1}{N} \sum_{i=1}^N \|R(I_{L,m}) - C_G\|_1 \quad (8)$$

## 4 实验设计

### 4.1 实验环境

本研究基于深度学习开源框架 Pytorch 1.10 搭建神经网络模型,训练及实验的硬件环境为: Intel(R) Core(TM) i7-8700 CPU, 32G RAM 及 NVIDIA GeForce 1080 GPU。对于不同条件下获得的实验结果,采用图像结构相似度(SSIM)和峰值信噪比(PSNR)作为图像重建质量的评价指标,使用 Matlab 进行测量。其中 PSNR 与图像的均方误差(Mean Square Error, MSE)成反比, SSIM 与 PSNR 越大,重建结果越接近真值图像。

### 4.2 数据准备与实验设置

神经网络的训练需要大规模的数据集,直接通过光学成像系统获得是困难的。因此本文基于文献[18]提供的 DIV2K 数据集制作仿真训练集。该数据集共包含 1 000 张不同场景下的高分辨率图像,其中前 800 张为训练图像,后 200 张为验证及测试图像。为了获得真值复图像集  $C_G$ ,首先将 800 张训练图像批量处理为 8 位的单通道格式,然后对 2k 分辨率的高清灰度图像按照尺寸  $128 \times 128$  进行分块裁剪,并随机挑选两张不同的图像合成一张复图像。利用 2.1 节中叙述的 FP 成像原理,将  $C_G$  作为仿真成像系统的输入,获得低分辨率图像集  $I_{L,m}$ ,其中每一张高分辨率复图像与相对应的  $m$  张低分辨率图像构成一组训练对。在测试集方面,本文使用 4 个基准数据集: Set5<sup>[23]</sup>、Set14<sup>[24]</sup>、B100<sup>[25]</sup> 和 Urban100<sup>[26]</sup>,共含有 219 张高清图像,与训练集的处理过程相同。训练集和测试集的不同来源提升了实验结果的准确性。

在对 FP 成像系统的设置中,LED 光源为 SMD3528,发光尺寸为  $3.5 \text{ mm} \times 2.8 \text{ mm}$ ,波长  $\lambda$  为  $630 \text{ nm}$ ,与样品相距  $80 \text{ mm}$ 。系统的放大倍率为 4 倍,数值孔径 NA 为 0.1,采样的相机为 Lumenra 公司的 Infinity-4 ( $2\,650 \text{ 像素} \times 4\,600 \text{ 像素}$ ),像元尺寸  $9 \mu\text{m}$ 。为了系统地比较不同方法中 FP 的孔径重叠率对重建结果的影响以及探索在稀疏采样条件下的重建效果,相邻 LED 点光源之间距离分别为 10, 5, 3, 2 mm。根据 Dong 等<sup>[27]</sup>的研究结果,孔径重叠率需要保证在 35% 以上才能实现传统 FP 方法的有效重建。在实际低分辨率图像的

采样过程中,考虑到重建算法的恢复效果,通常会将孔径重叠率设置在 60%~75% 的区间内<sup>[28]</sup>。不同重叠率下的孔径排列如图 6 所示,每一个圆形代表相应 LED 在频谱上对应的子孔径范围,其中亮度高的部分为相邻子孔径的重叠区域。图 6 表示 LED 数量分别为 25, 49, 81, 121 时孔径的重叠情况,4 种采样方式的孔径重叠率<sup>[28]</sup>分别为 25%, 50%, 62.5%, 70%,最终的合成孔径  $\text{NA}_{\text{syn}}$  都为 0.5。

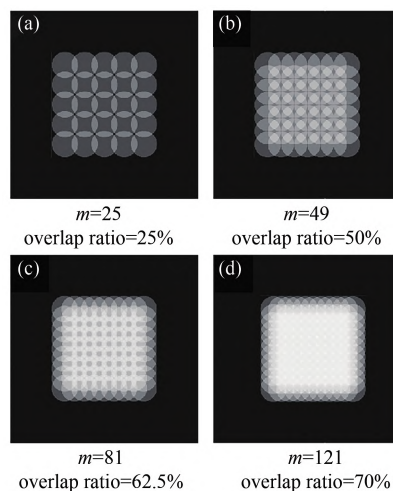


图 6 4 种不同孔径重叠率的示意图

Fig. 6 Schematic of four different aperture overlap ratio

神经网络训练过程中使用 Adam 优化器对损失函数进行更新,其中  $\beta_1$  设置为 0.9,  $\beta_2$  设置为 0.999。网络整体的初始学习率设置为  $10^{-4}$ ,权重衰减系数设置为 0,批尺寸(Batchsize)设置为 8。每经过 100 个反向传播周期学习率减半,最后经过 400 个周期完成训练。

## 5 结果与讨论

本文通过消融实验,探究了多尺度特征图数量对重建效果的影响。并且分别对稠密连接模块、通道注意力模块以及 DCCA 模块的重建效果进行探究。在改变孔径重叠率的条件下,使用本文方法(FPMFN)、AP 算法、PtychNet 和 Jiang 等学者提出的迭代算法(FP Neural Network, FPN)进行重建测试,对图像重建质量和算法运行时间做对比实验分析,同样也在实际的数据上进行重建实验。本文还将探究高斯噪声对不同方法重建结果的影响。

5.1 DCCA 的消融实验

表 1 为设置不同数量的多尺度特征图时,使用相同的训练集和训练周期,FPMFN 重建 50% 重叠率测试集的平均 PSNR 和 SSIM。当多尺度特征图数量少于 4 个时,神经网络的深度有限,特征金字塔模块的编码器对顶层的语义信息提取不充分,导致重建结果不佳。而多尺度特征图数量过多时,神经网络的参数量随之增加,在训练过程中容易导致过拟合现象,影响最终的重建效果。因此,本文选用重建效果最佳的 4 层特征结构搭建网络。

表 1 5 种特征图数量的网络在测试集上的平均 PSNR 和 SSIM

Tab. 1 Average PSNR and SSIM on the test datasets for networks with 5 feature map numbers

| Number of feature map | Max feature map size | Min feature map size | PSNR/dB      | SSIM           |
|-----------------------|----------------------|----------------------|--------------|----------------|
| 2                     | 128×128              | 64×64                | 23.70        | 0.781 3        |
| 3                     | 128×128              | 32×32                | 25.13        | 0.816 9        |
| 4                     | 128×128              | 16×16                | <b>28.87</b> | <b>0.847 1</b> |
| 5                     | 128×128              | 8×8                  | 27.39        | 0.832 2        |
| 6                     | 128×128              | 4×4                  | 27.21        | 0.817 3        |

注:粗体数字为最优值

对分别使用 DCCA 模块与 3×3 卷积层的神经网络比较了训练过程中的损失收敛情况。对比组将 DCCA 模块替换为 5 个标准卷积加激活函数模块,在 50% 的孔径重叠率下,对 49 张低分辨率图像进行重建。如图 7 所示,曲线的横坐标表示训练周期,纵坐标表示损失函数返回值(Loss)。与标准卷积相比,DCCA 模块的 Loss 曲线在 100 个训练周期附近达到了良好的收敛效

果,而使用标准卷积则需要大约 150 个训练周期,而且 DCCA 模块在训练中始终维持更低的 Loss 值,因此具有更好的收敛性能。

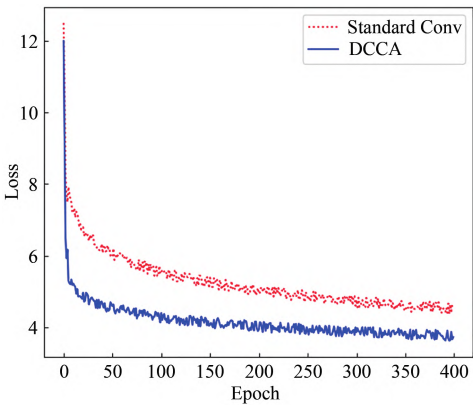


图 7 DCCA 模块与标准卷积在训练中的收敛性  
Fig. 7 Convergence of DCCA module and standard convolution in training

表 2 为使用不同模块在 4 个基准测试集上得到振幅图像的 PSNR 和 SSIM 平均值。“ConvCA”代表在使用标准卷积的基础上,增加了如 3.2 节中描述的通道注意力机制,而 DCCA 模块则进一步增加了卷积之间的稠密连接。由表 2 可以看出,DCCA 模块在不同的基准数据集上都取得了最好的重建效果,稠密连接提高了信息在神经网络中的利用率,有利于特征金字塔模块充分提取有效的特征信息。与标准卷积模块相比,DCCA 在 4 个测试数据集上的平均 PSNR 分别提升了 0.62,1.12,0.24,0.55 dB。同时,在多层卷积操作之后增加通道注意力机制,有效地分配了通道间的权重信息,使神经网络对不同频率的特征信息进一步细化,提升了最终的重建质量。

表 2 不同模块在测试数据集上的 PSNR 和 SSIM 测量值  
Tab. 2 PSNR and SSIM of different module on test datasets

| Module | Set5         |                | Set14        |                | B100         |                | Urban100     |                |
|--------|--------------|----------------|--------------|----------------|--------------|----------------|--------------|----------------|
|        | PSNR/dB      | SSIM           | PSNR/dB      | SSIM           | PSNR/dB      | SSIM           | PSNR/dB      | SSIM           |
| Conv   | 32.84        | 0.938 1        | 28.93        | 0.863 8        | 28.18        | 0.839 5        | 28.38        | 0.840 4        |
| ConvCA | 33.17        | 0.940 3        | 29.12        | 0.864 9        | 28.35        | 0.833 4        | 28.81        | 0.842 9        |
| DCCA   | <b>33.46</b> | <b>0.941 5</b> | <b>30.05</b> | <b>0.867 3</b> | <b>28.42</b> | <b>0.839 4</b> | <b>28.93</b> | <b>0.847 1</b> |

注:粗体数字为最优值

5.2 不同孔径重叠率的实验对比

为了研究低分率图像样本数与重建效果的

关系,以及进一步探索所需样本数目更低的采样方式,本文在 4 种孔径重叠率条件下,对不同方法

重建的振幅图像进行对比,孔径重叠率的选择如 4.2 节叙述,固定合成孔径  $NA_{syn}$  尺寸,分别使用  $5\times 5,7\times 7,9\times 9,11\times 11$  的 LED 阵列实施照明采样。对于 FPMFN 和 PtychNet,随着低分辨率样本数目改变,需要通过调整仿真成像系统参数获取相应的训练数据集,并训练不同的神经网络模型。

不同方法在不同孔径重叠率下,对重建振幅图像的测量结果如表 3 所示。通过 PSNR 的变化可以得知,所有方法的重建误差都会随着重叠率的增加而减小,说明足够的重叠率是实现有效重建所不可或缺的。PtychNet 的 PSNR 平均值比 AP 算法低 4.88 dB,重建的误差是由于网络使用了 3 层的卷积和激活函数,其结构比较简单,对低分辨率图像的特征信息提取不充分。FPNN 的 PSNR

平均值与 AP 算法接近,相差 1.03 dB。FPNN 使用神经网络中的反向传播完成重建过程,但是针对不同的图像都需要重新训练,泛化能力差,导致与 AP 算法的测量结果间存在误差。而 FPMFN 在 4 种孔径重叠率下,测量值都优于传统算法,相比于 PtychNet、AP 算法和 FPNN,PSNR 的平均值分别提升了 7.07,2.19,3.32 dB,与真值图像之间的误差更小。此外,FPMFN 模型在 25% 重叠率下的测试结果与 AP 算法在 50% 重叠率下的结果相差无几,证明 FPMFN 能够有效减少样本数目,实现更高效率的采样。这是因为 FPMFN 能够利用 DCCA 模块对高频部分的频谱信息进行有效提取,从而减小重建误差,在更低的重叠率下实现精细的重建。

表 3 不同重叠率下各个方法的 PSNR 和 SSIM

Tab. 3 PSNR and SSIM of each method with different overlap ratio

| method   | Overlap ratio/% | Set5         |                | Set14        |                | B100         |                | Urban100     |                |
|----------|-----------------|--------------|----------------|--------------|----------------|--------------|----------------|--------------|----------------|
|          |                 | PSNR/dB      | SSIM           | PSNR/dB      | SSIM           | PSNR/dB      | SSIM           | PSNR/dB      | SSIM           |
| PtychNet | 25              | 19.42        | 0.612 3        | 16.99        | 0.546 2        | 16.21        | 0.474 9        | 15.90        | 0.431 6        |
| AP       | 25              | 25.95        | 0.820 2        | 23.43        | 0.771 4        | 21.96        | 0.728 6        | 21.20        | 0.701 1        |
| FPNN     | 25              | 23.62        | 0.735 8        | 22.71        | 0.699 7        | 22.72        | 0.714 7        | 20.79        | 0.675 0        |
| FPMFN    | 25              | <b>29.07</b> | <b>0.891 7</b> | <b>27.91</b> | <b>0.833 1</b> | <b>27.48</b> | <b>0.835 7</b> | <b>26.55</b> | <b>0.828 4</b> |
| PtychNet | 50              | 25.49        | 0.805 1        | 24.29        | 0.764 9        | 23.35        | 0.638 4        | 22.81        | 0.644 1        |
| AP       | 50              | 30.92        | 0.915 8        | 28.01        | 0.841 3        | 27.39        | 0.829 5        | 26.53        | 0.824 9        |
| FPNN     | 50              | 28.92        | 0.914 3        | 28.78        | 0.849 7        | 24.49        | 0.801 3        | 24.75        | 0.796 2        |
| FPMFN    | 50              | <b>33.46</b> | <b>0.941 5</b> | <b>29.75</b> | <b>0.867 3</b> | <b>28.42</b> | <b>0.839 4</b> | <b>28.93</b> | <b>0.847 1</b> |
| PtychNet | 62.5            | 27.03        | 0.836 7        | 26.31        | 0.818 7        | 23.18        | 0.748 7        | 22.14        | 0.721 3        |
| AP       | 62.5            | 32.25        | 0.952 3        | <b>29.40</b> | 0.880 7        | 28.57        | 0.862 1        | 27.72        | 0.832 3        |
| FPNN     | 62.5            | 31.40        | 0.929 6        | 28.26        | 0.855 2        | 26.64        | 0.820 3        | 26.24        | 0.831 4        |
| FPMFN    | 62.5            | <b>35.18</b> | <b>0.968 4</b> | <b>29.39</b> | <b>0.894 7</b> | <b>28.71</b> | <b>0.864 2</b> | <b>29.38</b> | <b>0.873 9</b> |
| PtychNet | 70              | 27.88        | 0.855 8        | 26.26        | 0.836 8        | 25.04        | 0.782 4        | 24.58        | 0.747 5        |
| AP       | 70              | 33.92        | 0.958 8        | 31.94        | 0.907 4        | 28.36        | 0.887 4        | 30.41        | 0.905 0        |
| FPNN     | 70              | 33.69        | 0.941 2        | 29.34        | 0.880 2        | 27.69        | 0.848 3        | 29.68        | 0.883 8        |
| FPMFN    | 70%             | <b>36.23</b> | <b>0.963 3</b> | <b>32.06</b> | <b>0.931 9</b> | <b>29.35</b> | <b>0.904 7</b> | <b>31.07</b> | <b>0.912 8</b> |

注:粗体数字为最优值

表 4 记录了重建  $128\times 128$  灰度图像所消耗的时间。孔径重叠率和所需的低分辨率图像数量成正比,随着输入图像的增多,各个算法的数据量与计算量也随之增加,延长了重建所需时间。PtychNet 和 FPMFN 同为端对端模型,虽然训练过程需要消耗大量的时间,但是极大地减少了重建的时间成本,两种算法在 70% 的重叠率下也能在 0.1 s 内完成重建。由于 FPMFN 的模型相对复杂,所以重建速度稍慢一些。AP 算法与 FPNN 都是基于迭代求解的方法,FPNN 在

每次重建前都需要对网络重新进行训练。当输入图像个数为 121 时,FPNN 重建所需时间达到了 28.59 s,AP 算法则需要 2.51 s。端对端形式的 FPMFN 与基于迭代的 AP 算法相比,减少了 90% 以上的重建消耗时间,将重建速度提升了 20 倍以上。4 种算法在相同的硬件设备上运行,因此重建速度更快的 FPMFN 和 PtychNet 在算法功耗方面同样具有优势。相比于传统重建算法,FPMFN 不仅能有效降低 25% 的孔径重叠率和 50% 的样本数目,并且在相同的孔径重叠率下,

表 4 不同方法重建 128×128 图像的时间

Tab. 4 Time consumption of reconstructing 128×128 image in different methods (ms)

| Overlap ratio/% | PtychNet | AP      | FPNN     | FPMFN |
|-----------------|----------|---------|----------|-------|
| 25              | 17.3     | 532.2   | 2 659.7  | 19.4  |
| 50              | 32.4     | 1 181.5 | 8 419.9  | 35.1  |
| 62.5            | 52.7     | 2 057.1 | 17 374.2 | 58.3  |
| 70              | 83.2     | 2 506.4 | 28 587.4 | 90.5  |

注:粗体数字为最优值

该算法所需功耗更低,运行速度更快、重建质量更优。

图 8 为不同重建方法在 25% 和 50% 孔径重叠率下对仿真数据重建的结果展示,其中第一列

为原始的振幅和相位图像。在 PtychNet 的方法中没有引入相位信息,模型的重建结果出现较多的模糊和扭曲,图像的细节信息丢失过多。AP 算法相比于 PtychNet 和 FPNN 在图像细节上有更好的修复效果,但重建后的轮廓信息仍有不同程度的缺失。对比其他的方法, FPMFN 的重建有最佳的视觉效果,图像细节处的边缘锐度和清晰度都有较大提升。本文提出的特征金字塔模块能够提取不同感受野范围内的多层特征,并通过卷积神经网络的自适应融合,强化了对细节和轮廓信息的重建,使 FPMFN 能更加充分利用图像的低频与高频信息,在不同的场景下也有更高质量的重建结果。

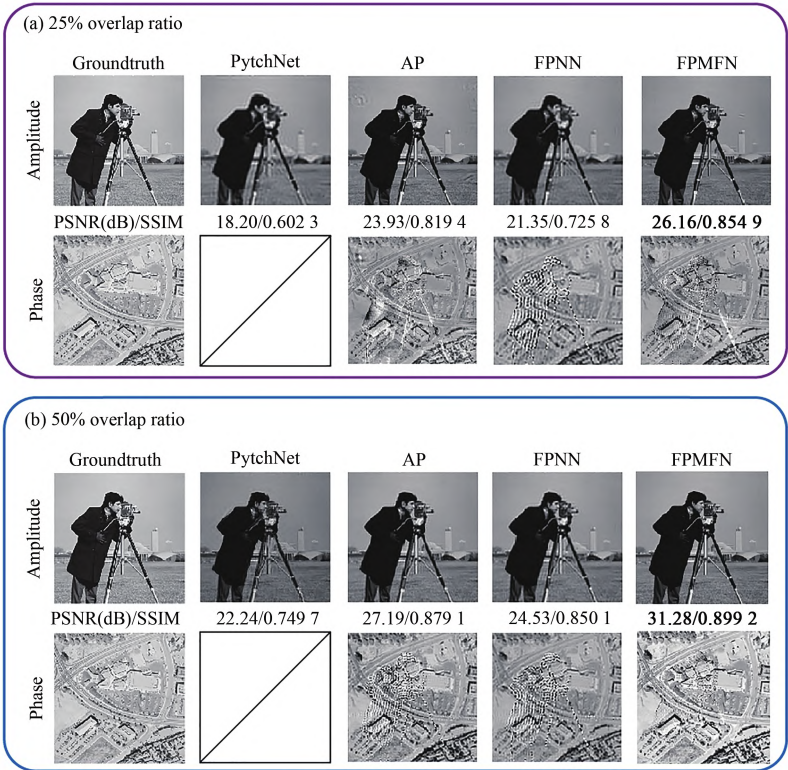


图 8 各个方法在不同重叠率时对仿真数据集的重建结果。(a) 25%;(b) 50%。

Fig. 8 Reconstruction on simulation datasets of each method with different overlap ratio. (a) 25%;(b) 50%.

此外,为了探究各算法在不同噪声条件下的重建效果,本文将均值为 0 的高斯噪声加入测试数据集中,噪声的标准差在 $[2 \times 10^{-5}, 3 \times 10^{-4}]$ 范围内变化,图 9(a)和图 9(b)的横坐标为高斯噪声的标准差,纵坐标分别为重建后振幅图像的平均 PSNR 和 SSIM。

FPMFN 的重建效果受噪声的影响最小,随

着高斯噪声的标准差增加,重建后的 PSNR 在测试范围内波动最大值为 3.32 dB,SSIM 的波动最大值为 0.082 5,而其他方法受噪声的影响较大。端对端形式的 FPMFN 具有更出色的普适性,因此能够有效抵抗噪声的干扰。

在上述对仿真数据的重建结果中,不同方法在 25% 和 50% 重叠率时的差异更加明显,且

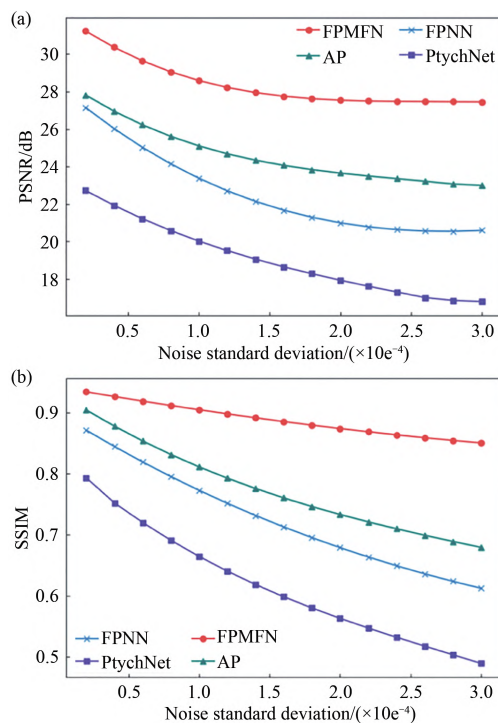


图9 不同方法高斯噪声的鲁棒性分析。(a) PSNR;(b) SSIM。

Fig. 9 Robustness analysis of different methods to Gaussian noise. (a) PSNR;(b) SSIM.

FPMFN与AP算法的重建效果最好。因此在实际系统的采样中,设置LED间距为10 mm和5 mm,并使用FPMFN和AP算法进行重建。图10为两种算法分别在25%和50%重叠率下对实际获取的25张和49张USAF低分辨率图像的重建结果,其中最右侧的LR Image为中心处LED照明时采集到的低分辨率图像。FPMFN对振幅和相位图像的重建结果都比AP算法更清晰。通过放大的细节部分可以看出,FPMFN对噪声有更好的抑制能力,并且对于条纹细节和轮廓处的重

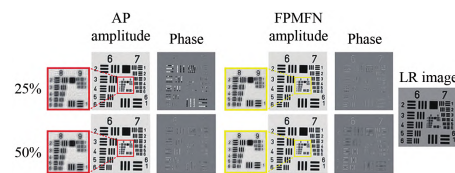


图10 AP与FPMFN算法在实际USAF数据集上的重建结果

Fig. 10 Reconstruction on actual USAF datasets by AP and FPMFN

建效果更佳。

## 6 结 论

本文提出一种基于多尺度特征融合网络的傅里叶叠层成像算法。该算法利用深度学习的卷积神经网络搭建模型FPMFN,通过稠密连接和通道注意力机制组成核心模块DCCA,4个DCCA共同构成了特征金字塔模块的编码器,能够高效地从低分辨率图像中提取高频和低频特征信息,并经过解码器进行融合。使用亚像素卷积对经过多尺度融合后的特征图进行上采样,完成FPMFN的高分辨率重建。DCCA模块有利于加强信息在网络间的流动,提升模型的收敛效率以及抗噪性能。为了验证网络同样适用于低样本数目的重建任务,本文在不同孔径重叠率的条件下进行实验,结果表明,FPMFN能有效降低FP中25%的孔径重叠率,减少50%的采样数量,提升重建与采样速度,同时实现效果更出色的重建。但是,在实际应用中,孔径扫描式的样本采集过程较为复杂,因此利用神经网络进一步降低孔径重叠率,使用更高效的采集方式仍然需要深入探索。

## 参 考 文 献:

- [1] LOHMANN A W, DORSCH R G, MENDLOVIC D, *et al.* Space-bandwidth product of optical signals and systems [J]. *JOSA A*, 1996, 13(3): 470-473.
- [2] 孙佳嵩,张玉珍,陈钱,等. 傅里叶叠层显微成像技术:理论、发展和应用[J]. 光学学报,2016,36(10):1011005.  
SUN J S, ZHANG Y Z, CHEN Q, *et al.* Fourier ptychographic microscopy: theory, advances, and applications [J]. *Acta Optica Sinica*, 2016, 36(10): 1011005. (in Chinese)
- [3] 张韶辉,周国城,崔柏岐,等. 傅里叶叠层显微成像模型、算法及系统研究综述[J]. 激光与光电子学进展,2021,58(14): 1400001.  
ZHANG S H, ZHOU G C, CUI B Q, *et al.* Review of Fourier ptychographic microscopy: models, algorithms, and systems [J]. *Laser & Optoelectronics Progress*, 2021, 58(14): 1400001. (in Chinese)

- [4] ZHENG G A, HORSTMAYER R, YANG C. Wide-field, high-resolution Fourier ptychographic microscopy [J]. *Nature Photonics*, 2013, 7(9): 739-745.
- [5] ZHENG G A. *Fourier Ptychographic Imaging: A MATLAB Tutorial* [M]. San Rafael: Morgan & Claypool Publishers, 2016: 10-22.
- [6] CHUNG J, OU X Z, KULKARNI R P, *et al.* Counting white blood cells from a blood smear using Fourier ptychographic microscopy [J]. *PLoS One*, 2015, 10(7): e0133489.
- [7] WILLIAMS A J, CHUNG J, OU X Z, *et al.* Fourier ptychographic microscopy for filtration-based circulating tumor cell enumeration and analysis [J]. *Journal of Biomedical Optics*, 2014, 19(6): 066007.
- [8] HORSTMAYER R, OU X Z, ZHENG G A, *et al.* Digital pathology with Fourier ptychography [J]. *Computerized Medical Imaging and Graphics*, 2015, 42: 38-43.
- [9] ZHENG G A, SHEN C, JIANG S W, *et al.* Concept, implementations and applications of Fourier ptychography [J]. *Nature Reviews Physics*, 2021, 3(3): 207-223.
- [10] HOLLOWAY J, ASIF M S, SHARMA M K, *et al.* Toward long-distance subdiffraction imaging using coherent camera arrays [J]. *IEEE Transactions on Computational Imaging*, 2016, 2(3): 251-265.
- [11] TIAN L, LI X, RAMCHANDRAN K, *et al.* Multiplexed coded illumination for Fourier ptychography with an LED array microscope [J]. *Biomedical Optics Express*, 2014, 5(7): 2376-2389.
- [12] KAPPELER A, GHOSH S, HOLLOWAY J, *et al.* Ptychnet: CNN based Fourier ptychography [C]//2017 *IEEE International Conference on Image Processing (ICIP)*. Beijing, China: IEEE, 2017: 1712-1716.
- [13] BIAN L H, SUO J L, ZHENG G A, *et al.* Fourier ptychographic reconstruction using Wirtinger flow optimization [J]. *Optics Express*, 2015, 23(4): 4856-4866.
- [14] ZHANG J Z, XU T F, SHEN Z Y, *et al.* Fourier ptychographic microscopy reconstruction with multiscale deep residual network [J]. *Optics Express*, 2019, 27(6): 8612-8625.
- [15] JIANG S W, GUO K K, LIAO J, *et al.* Solving Fourier ptychographic imaging problems *via* neural network modeling and TensorFlow [J]. *Biomedical Optics Express*, 2018, 9(7): 3306-3319.
- [16] 陈献明, 王阿川, 王春艳. 基于深度学习的木材表面缺陷图像检测[J]. 液晶与显示, 2019, 34(9): 879-887.  
CHEN X M, WANG A C, WANG C Y. Image detection of wood surface defects based on deep learning [J]. *Chinese Journal of Liquid Crystals and Displays*, 2019, 34(9): 879-887. (in Chinese)
- [17] 刘杨帆, 曹立华, 李宁, 等. 基于YOLOv4的空间红外弱目标检测[J]. 液晶与显示, 2021, 36(4): 615-623.  
LIU Y F, CAO L H, LI N, *et al.* Detection of space infrared weak target based on YOLOv4 [J]. *Chinese Journal of Liquid Crystals and Displays*, 2021, 36(4): 615-623. (in Chinese)
- [18] 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)*. Honolulu, HI, USA: IEEE, 2017: 4681-4690.
- [19] LIN T Y, DOLLÁR P, GIRSHICK R, *et al.* Feature pyramid networks for object detection [C]//2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA: IEEE, 2017: 2117-2125.
- [20] HUANG G, LIU Z, VAN DER MAATEN L, *et al.* Densely connected convolutional networks [C]//2017 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Honolulu, HI, USA: IEEE, 2017: 4700-4708.
- [21] HU J, SHEN L, SUN G. Squeeze-and-excitation networks [C]//2018 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Salt Lake City, UT, USA: IEEE, 2018: 7132-7141.
- [22] SHI W Z, CABALLERO J, HUSZÁR F, *et al.* Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA: IEEE, 2016: 1874-1883.
- [23] BEVILACQUA M, ROUMY A, GUILLEMOT C, *et al.* Low-complexity single-image super-resolution based on nonnegative neighbor embedding [C]//*British Machine Vision Conference*. Surrey, UK: BMVC, 2012.
- [24] ZEYDE R, ELAD M, PROTTER M. On single image scale-up using sparse-representations [C]//*Proceedings of the 7th International Conference on Curves and Surfaces*. Avignon, France: Springer, 2010: 711-730.
- [25] MARTIN D, FOWLKES C, TAL D, *et al.* A database of human segmented natural images and its application to

- evaluating segmentation algorithms and measuring ecological statistics [C]//*Proceedings Eighth IEEE International Conference on Computer Vision*. Vancouver, BC, Canada: IEEE, 2001: 416-423.
- [26] HUANG J B, SINGH A, AHUJA N. Single image super-resolution from transformed self-exemplars [C]//2015 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Boston, MA, USA: IEEE, 2015: 5197-5206.
- [27] DONG S Y, BIAN Z C, SHIRADKAR R, *et al.* Sparsely sampled Fourier ptychography [J]. *Optics Express*, 2014, 22(5): 5455-5464.
- [28] 沙浩,刘阳哲,张永兵. 基于深度学习的傅里叶叠层成像技术[J]. *激光与光电子学进展*, 2021, 58(18): 1811020.
- SHA H, LIU Y Z, ZHANG Y B. Fourier ptychography based on deep learning [J]. *Laser & Optoelectronics Progress*, 2021, 58(18): 1811020. (in Chinese)

## 作者简介:



宋东翰(1995—),男,河南洛阳人,硕士研究生,2018年于华中科技大学获得学士学位,主要从事深度学习、计算机视觉等方面的研究。E-mail: songdonghan@qq.com



王 斌(1980—),男,吉林省吉林市人,博士,研究员,2018年于东北师范大学获得博士学位,主要从事工程光学、计算光学等方面的研究。E-mail: 175969722@qq.com