

文章编号 1004-924X(2024)20-3047-12

面向显微视觉的端对端去模糊模型

徐 征, 何佳珩, 王彦琪, 王晓东*, 任同群
(大连理工大学 机械工程学院, 辽宁 大连 116081)

摘要:显微视觉测量在微装配等领域中应用广泛,受成像景深等因素的影响,图像会出现多重离焦模糊现象,影响后续准确测量,而显微自动对焦技术虽然可以缓解离焦问题,但耗时较长,难以适应高效生产要求。本文提出了将模糊度判别和多分支恢复相结合的端对端去模糊模型,建立了分块、判别、去模糊、融合的分而治之策略:首先将一幅图像切割成子图像组,同时送入判别器和恢复网络;在判别器中,通过傅里叶变换等获取频域分布,再利用 Vision-Transformer 网络从频域图中提取具有全局关联性的频域深层模糊特征,然后对模糊度进行判别输出。根据判别结果,由多分支恢复网络对不同模糊度的子图像进行定向恢复,最后融合拼接处理后的子图像,获得高清晰度的图像。实验结果表明,本文提出的模型能有效恢复多重模糊的显微图像,判别准确率达 0.94,而模糊图像经过多分支恢复网络处理后,PSNR 指标平均提升了 6.3。

关键词:微装配;显微视觉;模糊度判别;频域处理

中图分类号:TP394.1;TH691.9 **文献标识码:**A **doi:**10.37188/OPE.20243220.3047

End-to-end deblurring model for microscopic vision

XU Zheng, HE Jiaheng, WANG Yanqi, WANG Xiaodong*, REN Tongqun

(College of Mechanical Engineering, Dalian University of Technology, Dalian 116081, China)

* Corresponding author, E-mail: xdwang@dlut.edu.cn

Abstract: The measurement of microscopic vision is commonly used in micro-assembly and other fields. However, due to limitations such as depth of field in microscopic imaging, the image may appear blurred and affect the accuracy of measurement. Although the technology of auto-focusing in optical microscopy can alleviate defocusing problems, it will be too time-consuming to adapt to the requirements of efficient production. Herein, an end-to-end deblurring model that integrates blurring discrimination and multi-branch recovery was presented, in which a divide-and-conquer strategy of chunking, discrimination, deblurring, and fusion was established. Firstly, the image was divided into sub-images, which were then simultaneously processed by a discriminator and a recovery network. The discriminator employed the Fourier transform to obtain the frequency-domain map of the sub-images. From the frequency domain map, the Vision Transformer network extracted deep blur features with global correlation. The output of the blurring degree was then discriminated. The multi-branch recovery network was used to directionally recover sub-images with different blurring degrees based on the discriminative output. Finally, the spliced sub-im-

收稿日期:2024-03-25;修订日期:2024-07-25.

基金项目:国防基础科研计划资助项目(No. JCKY2022203B006);中央高校基本科研业务费资助项目(No. DUT24LAB112)

ages were fused to obtain high-resolution images. The experimental results indicate that the model can effectively restore multi-blurred microscopic images, with a discriminator accuracy reaching 0.94. Moreover, after undergoing processing by the multi-branch restoration network, the PSNR metric shows an average improvement of 6.3.

Key words: micro-assembly; microscopic vision; blur classification; frequency domain processing

1 引言

显微视觉在现代精密工程中应用广泛。以自动微装配为例,其典型过程是通过显微视觉精确获取待组装零件和夹持器的位置,然后引导机器人等进行零件拾取、放置、连接等作业,检测准确性对于装配的精度影响显著。而在显微成像中,出现图像模糊现象往往难以避免的,零件的制造偏差、受力变形以及视觉单元中光学元件的变化等因素均有可能导致图像模糊,有必要研究如何消除模糊以获得被观察对象的准确轮廓。消除模糊的过程可视为模糊的逆过程,简称去模糊。传统上使用具有固定核的“硬”去卷积方法去模糊^[1],例如维纳滤波^[2]、Richardson-Lucy 方法、盲去卷积法等^[3-6],然而,显微成像模糊是由多重因素作用引起的,很难准确估计出描述这类综合作用的模糊核,导致去模糊时容易产生重影、振铃等伪图像。

卷积神经网络法(Convolutional Neural Network, CNN)是将从模糊图像到目标图像(清晰图像)映射学习表现为多层次的卷积核自主堆叠,呈现出灵活、可扩展和可学习的“软”特性。值得注意的是,在这一领域存在“去模糊”和“超分”两个类似概念,前者主要针对去成像模糊,而后者则更重视提高图像尺寸分辨率,但二者主体目标是一致的,即从低质量的图像中获得高质量图像,为叙述方便,后文将不再专门区分。

早期用于去模糊的 CNN 多采用均方误差损失函数(Mean Squared Error, MSE),恢复的图像虽然具有较好的外观,但轮廓细节容易过度平滑。Christian 等^[7]针对这一问题,采用了生成对抗网络和具有残差结构的深度 CNN 来处理图像。在 BSD100 数据集上的实验表明,生成对抗网络获得的图像真实感更强,而具有残差结构的

深度卷积网络生成的图像信噪比更高。除 CNN 外,另一典型的架构是 Transformer,它克服了卷积视野有限的问题,能够学习到长距离的关联信息。使用 Transformer 的图像重建算法^[8-9]在纹理等细节恢复上效果出色。此外,针对一幅图像中存在多重模糊的情况,Liu 等提出了基于深度卷积神经网络的多焦点融合的思路^[10],可以将多幅同一场景下聚焦区域不同的图像,融合成一幅全清晰的图像,消除图像融合中常见的伪影问题^[11, 12]。

在其他需要使用显微视觉领域也会遇到相似的图像模糊问题。赵宇轩^[13]等针对主流光片存在不足限制了显微技术对亚细胞结构的高层析精度,无法对亚细胞结构实现更精细的观测的问题,提出了基于多段级联各向同性卷积神经网络的单图超分辨算法;朱威铭^[14]和陈浩^[15]等分别用 U-Net 网络和生成对抗网络增强图像。

此外,基于传统的爬山法的显微自动对焦技术虽然可以缓解离焦问题,但需要结合机械进给进行粗精往复搜索,耗时严重,难以适应高效生产要求,而且处理背景复杂的小目标效果不够理想^[16-18]。

以上研究更多关注图像的外观效果,而在显微视觉测量的领域,还需要考虑恢复图像中轮廓的准确性。另外,宏观成像模糊更多是由镜头晃动带来的运动模糊,而由聚焦不准和光学景深限制引起的模糊在显微视觉中更为突出,二者的成像机制和频域分布不同。

本文针对工业测量与控制中显微成像模糊问题开展研究,主要工作包括:

(1)将图像模糊度判别与图像恢复相集成,形成重视恢复精度的端对端去模糊模型。

(2)在模糊度判别中,通过变换获取频域图,

以挖掘出更能反映模糊度的深度特征。

(3) 针对显微场景存在多重离焦的情况,建立了分块、判别、恢复、融合的分而治之策略。

2 端对端去模糊模型网络结构

如图 1 所示,为解决多重模糊问题,我们将模糊度判别与去模糊相集成,形成了端对端的功能模型:首先将待恢复图像进行分块,然后把分割

的子图像依次送入模糊度判别器,再根据判别结果将子图像送入多分支恢复模块进行重建,最后将去模糊后的子图像融合拼接,获得完整的去模糊图像。

通过前期研究发现,采用单一的主体网络很难对多重模糊图像准确恢复,这与斯捷等的研究结论一致^[19]。为此,本文设计了多分支恢复网络,分别针对清晰、轻度、中度、重度的模糊情况进行处理。

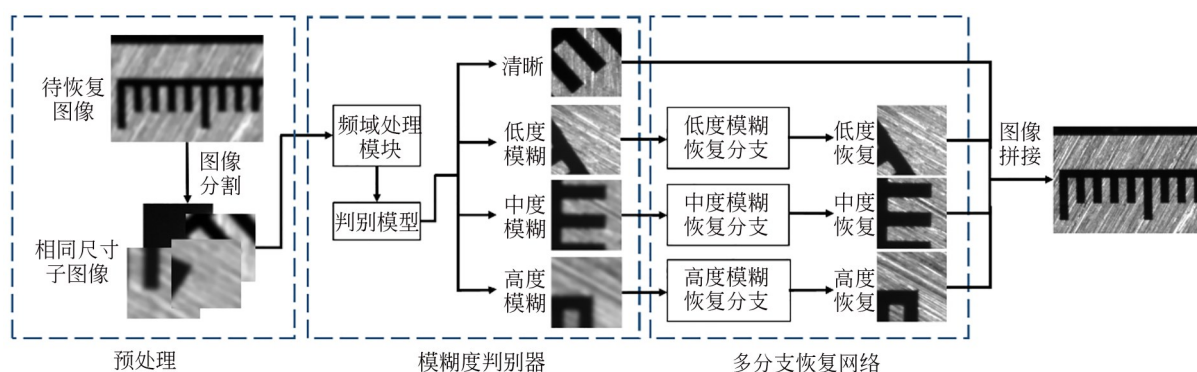


图 1 端对端去模糊模型整体流程

Fig. 1 Workflow of end-to-end deblurring model

2.1 模糊度判别器

判别器主要由频域处理模块和判别模型组成。构建判别器时,主要考虑两点:从频域信息中提取特征作为判别依据;考虑频域的整体分布,而不仅仅是根据局部峰值或某一分量去判别。判别器接收预处理模块分割出的子图像,再将子图像依次送入判别器中进行推理判别,获得每个子图像的模糊度。

2.1.1 频域处理模块

频域处理模块的步骤如图 2 所示:

(1) 傅里叶变换:利用式 1 对分块后的子图像进行变换,得到频域图(图 2 列 b)。

$$F(u, v) = \sum_{x=0}^{M-1} \sum_{y=0}^{N-1} f(x, y) e^{-2j\pi(ux/M + vy/N)}, \quad (1)$$

其中: $f(x, y)$ 表示一幅 $M \times N$ 像素的图像, x, y 用于表示空间域(也即图像原图)中的像素点位置, $F(u, v)$ 表示转换后得到的频域图, u, v 用于表示频域图中的像素点位置。

(2) 极坐标转换:将频域图进行极坐标转换,使高频部分处于左侧。从频域图可以看出,不同

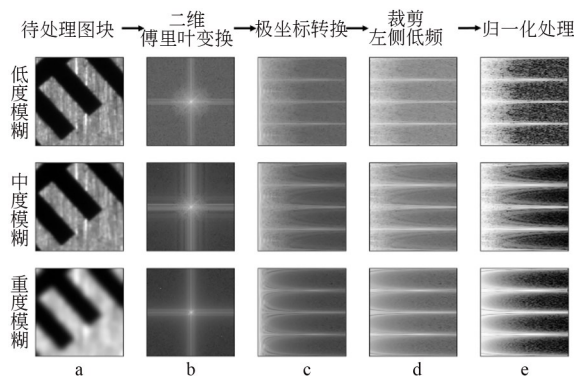


图 2 图像频域化处理步骤

Fig. 2 Process workflow of frequency domain

模糊度图像频域图的中心区域(低频部分)相对相似,而区别主要体现在高频部分。为了剔除低频部分,提升分类正确率,对频域图进行坐标转换,得到极坐标频域图(图 2 的列 c),此时低频部分在左侧。

(3) 裁剪低频部分:将极坐标转换后的频域图中,代表低频的左侧部分裁剪(图 2 列 d),由实验得出,裁剪量为总宽度的 10% 时效果最佳。

(4) 归一化:为突出频域图细节,对其进行归一化处理(图 2 列 e)。

频域处理模块中各步骤的消融实验见章节 4.1。

2.1.2 判别模型

传统的频域特征提取更重视某一方面的内容,例如特征频率、能量谱等,而考虑到显微成像的复杂性,有效捕捉信息中的长距离依赖关系或许能带来更有启发性的信息。相比 CNN,Transformer 网络通过自注意力机制处理图像中的长距离依赖性,能更好地捕获全局信息。因此本文选择 Google 提出的 Vision Transformer(VIT)作为判别模型基础,从频域图中提取全局关联信息,以更有效地判别模糊状态。VIT 模型由嵌入层、编码器和全连接头组成:

(1) 嵌入层(Embedding layer):首先将输入

图像分割为固定大小的图块(Patch),通过线性映射将每个图块映射展平为一维向量。之后通过位置编码嵌入向每个图块的特征表示中添加位置信息,以及通过添加特征向量 class token 集成所有图块信息,使网络在训练时保留空间信息和全局特征。

(2) 编码器(Encoder):VIT 包含多个编码器层,每个编码器层由多头自注意力机制和前馈神经网络组成,用于学习图块之间的关系。编码器通过处理包含特征向量 class token 的图块向量,挖掘与模糊度相关的深度信息。

(3) 全连接分类头(MLP Head):将编码器处理后的特征向量映射为判别概率分布,得到各模糊度的判别概率,将最大概率对应的模糊度作为判别结果。

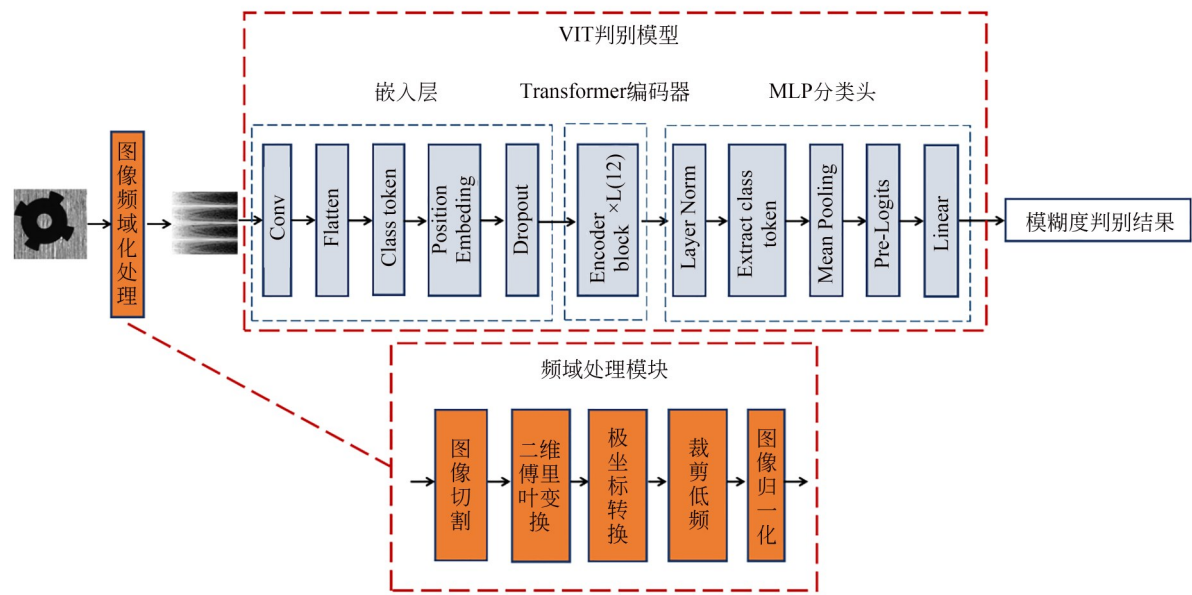


图3 模糊度判别器的结构

Fig. 3 Architecture of the blur discriminator.

2.2 多分支恢复网络

获得子图像模糊度判别结果后,使用多分支恢复网络对模糊子图像进行恢复处理。本文所使用的多分支网络如图 4 所示,三个分支分别对应三种模糊度的判别。将子图像按照判别结果,经由不同的分支进行恢复(清晰子图像无需恢复),每个分支网络的结构相同,均为包含 16 个残

差模块的 ResNet 网络。

深度残差模块运用跳链技术将输入连接到网络输出,使每层的映射关系为 $y=f(x)+x$,当深层网络中的某一层存在梯度消失问题时,也即 $f(x)$ 趋向于 0 时,可以直接“跳过”该层向更浅层传播,这样能有效缓解梯度消失问题,提高模型的整体性能。

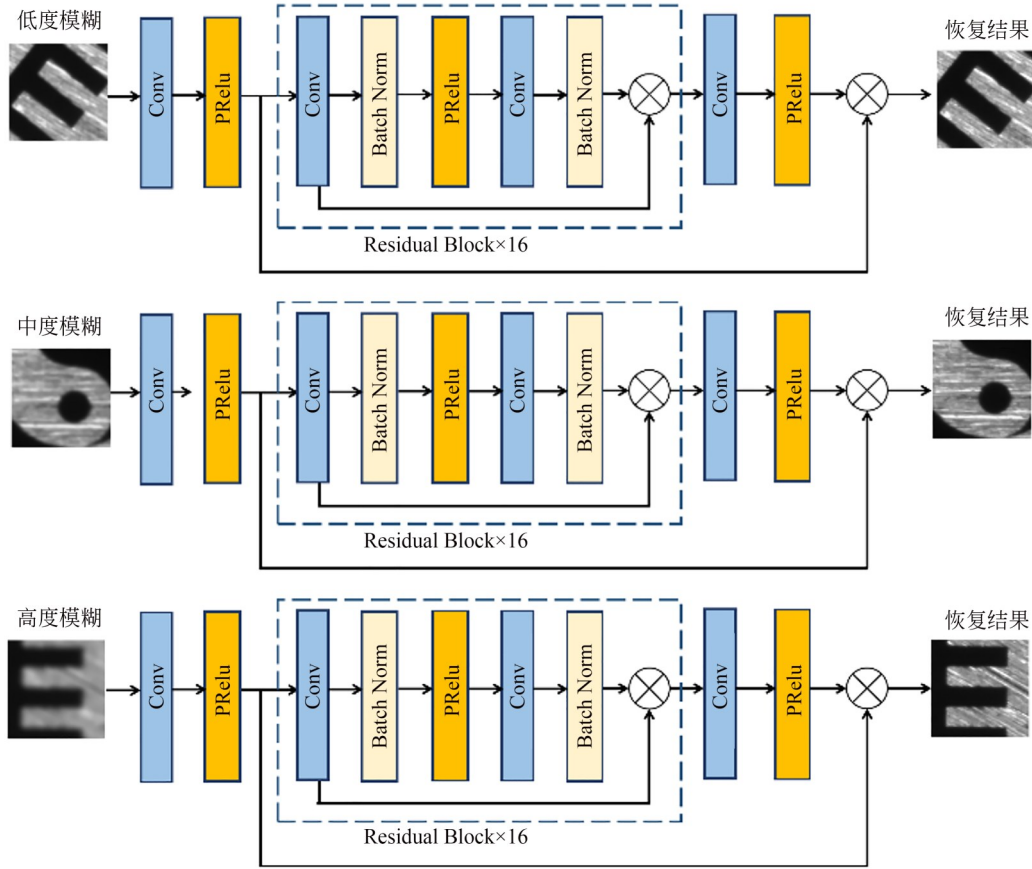


图 4 去模糊网络的结构

Fig. 4 Architecture of deblurring network

3 实验设置

3.1 数据集构建

训练模型需要大量的样本,而目前没有公开的显微视觉的数据集,因此本文采用了实拍和数据增强相结合的方法构建数据集。实验中所用的实际图像是通过一套显微成像单元采集(NA: 0.14, $5\times$ 放大倍率)获得的,显微成像单元搭载在三维精密运动平台上,用于调整观测位置。零件布置在下方的三维精密转台上,方便进行零件姿态的调整。

PSNR(峰值信噪比)和 SSIM(结构相似性)是两种衡量图像质量的指标,被广泛应用于图像恢复领域中。其中 PSNR 的计算公式如式(2):

$$\text{PSNR}(X, Y) = 10 \cdot \lg \left(\frac{\text{MaxValue}^2}{\text{MSE}(X, Y)} \right), \quad (2)$$

其中:MaxValue 为图像像素可取到的最大值, $\text{MSE}(X, Y)$ 为两张图像的均方误差。

SSIM 指标基于人眼提取图像中结构化信息的假设,更符合视觉感知,计算公式如式(3):

$$\text{SSIM}(X, Y) = [\mu(x, y)]^2 \cdot [c(x, y)]^2 \cdot [s(x, y)]^2, \quad (3)$$

其中: $\mu(x, y)$, $c(x, y)$, $s(x, y)$ 分别代表亮度、对比度和结构计算, $\alpha, \beta, \gamma > 0$, 用以调整三部分比重。

在显微成像中,光学失焦模糊是主要因素,失焦模糊的过程可以用高斯模糊近似。为了探究不同的物理离焦量与高斯模糊核之间的关系,对同一零件的同一位置,采集不同离焦量的清晰-光学失焦图像对,计算不同离焦量下的模糊图像与清晰图像的 PSNR 值与 SSIM 值,再对清晰图像进行不同程度的高斯模糊。高斯模糊核尺寸越大,模糊度也越高,其默认的具体关系为: $\sigma_x = \sigma_y = 0.3[0.5(k-1.0)-1.0]+0.8$, (4) 其中: σ_x 和 σ_y 分别为 X 和 Y 方向的高斯核标准偏差, k 为高斯核尺寸。

计算不同高斯模糊下的 PSNR 值与 SSIM 值得到图 5。从图中见,高斯模糊与离焦模糊具有相似的趋势,可以用于确定不同模糊度的光学模糊图像所对应的高斯模糊核。

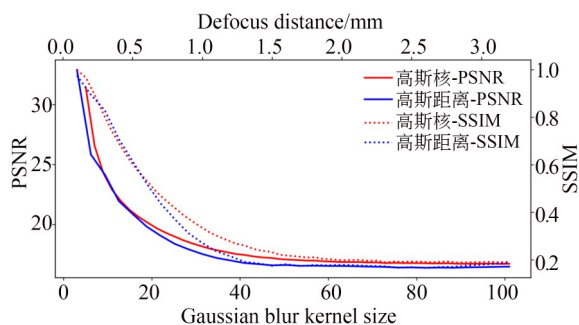


图 5 高斯模糊核尺寸与光学离焦距离效果对比

Fig. 5 Comparison between gaussian blur kernel size and optical defocus distance

在构建数据集、寻找与离焦模糊相对应的高斯模糊时,先计算出当前离焦距离下,模糊图像的 PSNR 或 SSIM 值,进而确定其在图 5 高斯模糊曲线上所对应的高斯核。构建的数据集中,如图 6 所示,包含清晰图像、轻度模糊图像(高斯核 $k=5$,离焦距离 $100\ \mu\text{m}$)、中度模糊图像(高斯核 $k=13$,离焦距离 $450\ \mu\text{m}$)、重度模糊图像($k=39$,离焦距离 $1\ 000\ \mu\text{m}$)。三种模糊度的 SSIM 指标分别约等于 0.9、0.6 和 0.2,同时在人眼观察效果上也有明显区分。

最终将得到的图像分为训练集、测试集和验证集三部分。其中,训练集用以进行模糊度判别器和多分支恢复网络的训练,验证集用于模型的调整和优化,测试集则用于评估最终选定模型的性能,既不参与训练过程,也不参与超参数和网络的优化。除此之外,额外制作了包含上述三种模糊度图像的光学模糊测试集,用以测试网络对光学模糊图像的恢复效果。

3.2 训练过程

在模糊度判别器进行训练时,按照图像模糊度标签训练;在多分支恢复网络进行训练时,使用清晰-轻度模糊图像、清晰-中度模糊图像和清晰-重度模糊图像制作三个子数据集,训练对应的网络分支。

如图 7 所示,分别对模糊度判别器和多分支

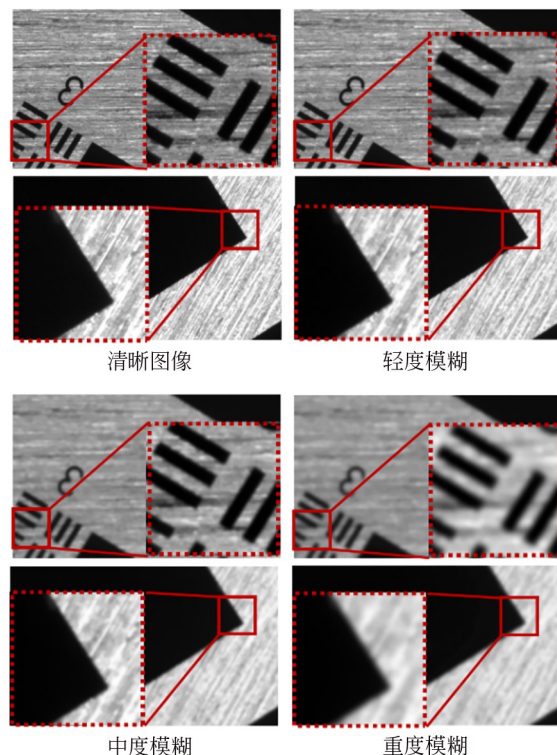


图 6 不同模糊度的图像

Fig. 6 Images with different blurriness degrees

恢复网络进行训练。其中,模糊度判别器训练时对于图像进行判别;多分支恢复网络内的三个分支会使用对应模糊度的子数据集进行恢复重建训练。

3.2.1 判别网络训练

在改进 VIT 网络的最后环节,特征向量 class token 被送入全连接的分类头,得到推理出的判别结果。VIT 中使用交叉熵损失函数:

$$Loss = -\frac{1}{N} \sum_{n=0}^{N-1} y_n \log(p_n) \quad (5)$$

其中: N 代表图像数量, y_n 代表第 n 个图像的标签,是值为 0 或 1 的标量, p_n 代表网络输出的此标签的判别概率, $Loss$ 代表得到的损失值。

在进行模糊度判别器的网络训练时,构筑的数据集中共包含 31 200 幅尺寸为 224×224 pixels 的图像,其中,训练集包含 21 600 幅图像、测试集包含 4 800 幅图像、验证集包含 4 800 幅图像。训练时 batch-size 设为 8,学习率 10^{-3} ,当训练轮次为 150 轮时,损失函数趋于稳定。

3.2.2 多分支恢复网络训练

与判别器采用频域信息挖掘模糊度深层特

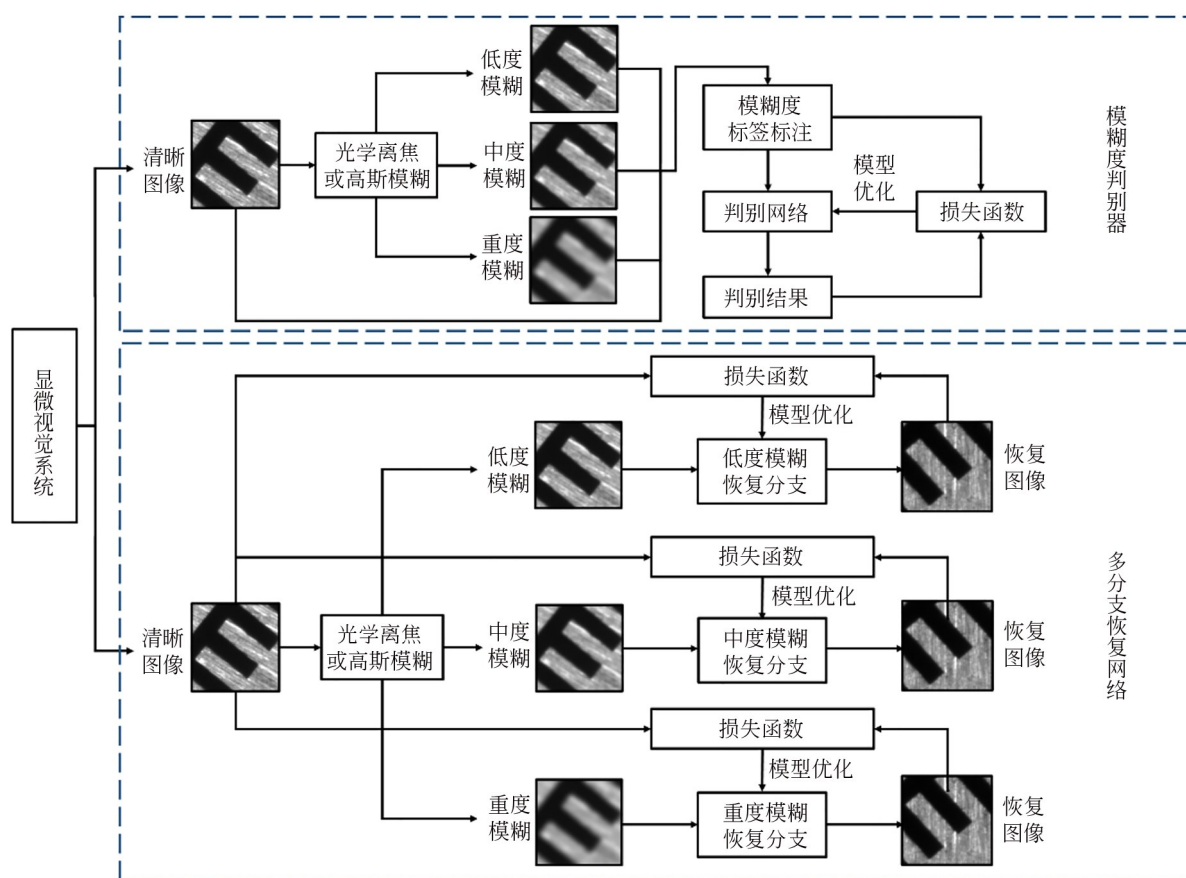


图 7 训练的整体流程

Fig. 7 Overall training process

征相似,在恢复网络的损失函数中,加入了能反映模糊度

变化的频域损失部分。损失函数表达式如下:

$$loss = MSE(f(\hat{\varphi}).re, f(\varphi).re) +$$

$$MSE(f(\hat{\varphi}).im, f(\varphi).im) + SE(\hat{\varphi}, \varphi), \quad (6)$$

其中: φ 为网络输入图像的灰度矩阵, $\hat{\varphi}$ 为网络输出图像的灰度矩阵, $f(\cdot)$ 为傅里叶变换, $f(\cdot).re$ 为傅里叶变换后的实部, $f(\cdot).im$ 为傅里叶变换后的虚部,MSE为均方损失函数。

在多分支恢复网络的训练时,构筑的3种模糊度的数据集中各有31 200幅尺寸为 96×96 图像,其中训练集含21 600幅图像,测试集含4 800幅图像,验证集包含4 800幅图像。训练batch-size为96,学习率为 10^{-4} ,训练轮次为100轮时,损失函数趋于稳定。

4 结果讨论

如前文所述,本文构建的端对端去模糊模型由模糊度判别器和多分支恢复网络构成,下面将分别对三者进行讨论。

4.1 模糊度判别效果

本节用包括准确率、精确率、召回率、F1值指标综合衡量模糊度判别效果。

准确率是指正确判别样本个数占总样本个数:

$$Accuracy = \frac{(TN + TP)}{(TN + TP + FP + FN)}, \quad (7)$$

其中:TP, TN, FP和FN分别表示真正例(预测为正例且实际为正例)、真负例(预测为负例且实际为负例)、假正例(预测为正例但实际为负例)和假负例(预测为负例但实际上是正例)。

精确率是指预测的正样本中,真正正样本所占的比例:

$$Precision = \frac{TP}{TP + FP} \tag{8}$$

召回率是指预测为正例数据, 占所有实际为正例数据的比例:

$$Recall = \frac{TP}{TP + FN} \tag{9}$$

F1 是基于精确率和召回率的调和平均:

$$F1 = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{10}$$

模糊度判别测试集中包含 4 种模糊度的图像各 1 200 幅, 实验结果见表 1。

表 1 模糊度判别效果

Tab. 1 Blurriness discrimination results

类别	精确率	召回率	F1 值	准确率
清晰	0.93	0.99	0.96	0.94
轻度模糊	0.99	0.99	0.99	
中度模糊	0.88	0.97	0.92	
重度模糊	0.97	0.84	0.90	

从表 1 可见, 测试集上的总体判别准确率为 0.94, 在不同模糊度上, 判别精确率、召回率、F1 值均大于 0.88。在轻度模糊组上判别结果最好, 各指标均不小于 0.98。

针对频域处理模块及其组成部分进行了消融实验, 结果见表 2。表中√表示包含该部分, 例如第一列表示无消融对照组; 第二列表示只消融归一化部分; 而最后一列表示将频域处理模块中的所有部分都完全消融。

从表 2 可见, 完全消融频域处理模块, 也即将图片原图直接送入 VIT 中进行训练, 判别效果最差, 总数据集准确率仅为 0.69; 分别消融低频剔除部分、图像归一化部分以及两者均消融后, 总数据集准确率均逊于未消融对照组, 其中, 消融归一化对判别准确率影响最小。可见, 频域处理模块的各部分均有效果。

4.2 多分支恢复效果

去模糊测试集由三部分组成: 轻度模糊图像 1 200 幅、中度模糊 1 200 幅、重度模糊 1 200 幅。使用对应的网络分支对三个数据集进行图像恢复, 最后使用 PSNR 和 SSIM 指标对图像恢复效果进行评估, 结果见表 3。

表 2 频域处理模块消融实验
Tab. 2 Frequency domain processing module ablation experiment

组成部分		选择				
傅里叶变换+极坐标转换		√	√	√	√	
低频剔除		√	√			
归一化		√		√		
清晰图片子数据集	精确率	0.93	1.00	0.98	0.99	0.77
	召回率	0.99	0.99	0.97	1.00	1.00
	F1 值	0.96	1.00	0.97	0.99	0.87
轻度模糊子数据集	精确率	1.00	0.97	0.99	0.90	0.51
	召回率	0.99	0.90	0.96	0.94	0.44
	F1 值	0.99	0.93	0.98	0.92	0.48
中度模糊子数据集	精确率	0.88	0.82	0.62	0.81	0.57
	召回率	0.97	0.84	0.99	0.83	0.84
	F1 值	0.92	0.83	0.76	0.82	0.68
重度模糊子数据集	精确率	0.97	0.85	0.97	0.91	0.96
	召回率	0.84	0.89	0.73	0.85	0.96
	F1 值	0.90	0.87	0.83	0.88	0.96
总数据集准确率		0.94	0.91	0.89	0.90	0.69

表 3 使用多分支网络的恢复效果
Tab. 3 Recovery result of corresponding branch networks for different degrees of blurriness

类别	SSIM		PSNR	
	恢复	模糊	恢复	模糊
轻度模糊	0.983	0.962	38.7	34.8
中度模糊	0.973	0.810	37.2	27.8
重度模糊	0.771	0.482	27.6	22.1
均值	0.909	0.751	34.5	28.2

可见, 在三种模糊度数据集上, 处理后的图像相较原始图像 SSIM 指标平均提升了 0.158, 对中度模糊和轻度模糊恢复后的 SSIM 指标能达到 0.983 和 0.973, 而在重度模糊测试集上, 模型虽也有恢复效果, 但改善程度相对较差, 这主要是因为重度模糊图像本身包含信息较少。

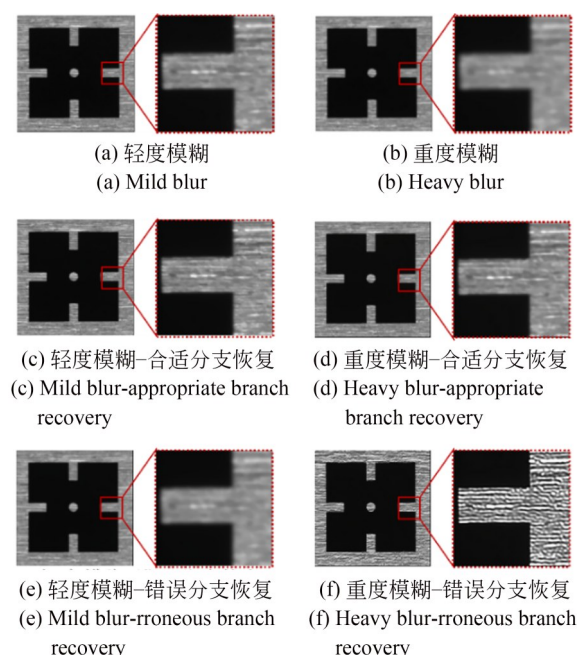


图 8 不同分支恢复效果对比

Fig. 8 Comparison of restoration results

如前文所述,采用单一的网络去处理多重失焦图像的精度效果是有限的。图 8 为使用不同分支网络对不同模糊度图像的处理结果。可见,使用合适的分支恢复图像,可以较好地生成清晰图像,而使用错误的分支恢复图像,则会导致效果欠佳(如图 8(e)为使用轻度模糊分支恢复重度模糊图像)或出现振铃现象(如图 8(f)为使用重度

模糊分支恢复轻度模糊图像)。这说明采用多分支恢复网络是合理的。

4.3 端对端去模糊的处理效果

本节将判别模型和多重模糊恢复网络联合,形成端对端模型,对测试集图像进行处理,如图 9 所示。仍以 SSIM 指标进行评判,指标越高则表示输出图像和无失真图像的差距越小,结果见表 4,可见,端对端的模型对图像中的三种模糊图块上能进行分而治之的有效处理。

为了测试模型鲁棒性,对模糊度进行微调,具体方法为,在保持高斯核尺寸不变,只调整 σ 值:

$$\sigma_{0.25} = 0.25 [0.5(ksize - 1.0) - 1.0] + 0.8, \quad (11)$$

$$\sigma_{0.35} = 0.35 [0.5(ksize - 1.0) - 1.0] + 0.8, \quad (12)$$

其中: $ksize$ 为高斯核尺寸,建立的两个微调数据集分别记为 $\sigma_{0.25}$ 和 $\sigma_{0.35}$,相较原数据集(记为 σ), $\sigma_{0.25}$ 数据集模糊度更低,而 $\sigma_{0.35}$ 数据集模糊度更高。

这两个数据集上恢复结果的 SSIM 指标如表 5 所示,在微调后的数据集上,恢复后的图像的 SSIM 值均高于恢复前。可见,端对端模型在一定模糊度的变化影响下,依然能对多重模糊图片进行有效处理。

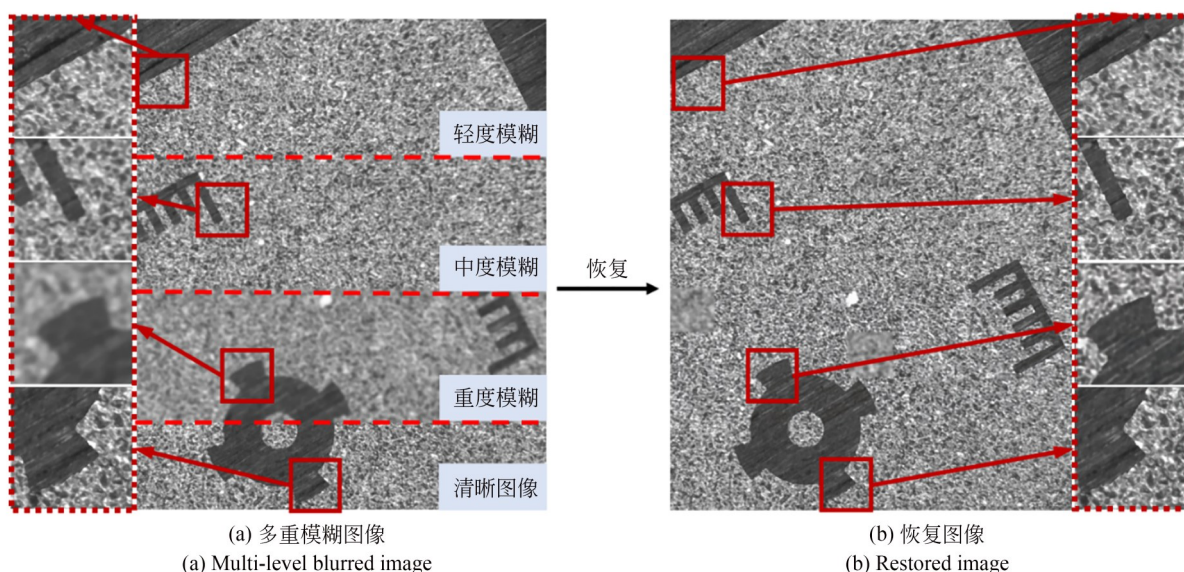


图 9 多重模糊度图像的恢复

Fig. 9 Recovery of image with multiple blurriness

表 4 端对端去模糊模型测试的 SSIM 评分

Tab. 4 SSIM scores of end-to-end deblurring model

	低度模糊 部分	中度模糊 部分	重度模糊 部分	整体图像
多重模糊	0.960	0.789	0.404	0.769
图像恢复	0.963	0.863	0.706	0.864

表 5 微调数据集上恢复效果

Tab. 5 Recovery result on fine-tuned datasets

数据集类别	图像	SSIM
σ	多重模糊	0.769
	图像恢复	0.864
$\sigma_{0.25}$	多重模糊	0.792
	图像恢复	0.852
$\sigma_{0.35}$	多重模糊	0.749
	图像恢复	0.780

4.4 与其他网络模型的对比实验

在验证了网络对模糊图像的恢复效果后,本节在光学多重模糊验证集上,将本文端对端去模糊网络与一些先进算法进行对比。

表 6 为分类器与其他分类网络的性能对比,可见本文分类器相比 ResNet18 和 ResNet50 网络^[20]均性能略优;相比 VGG 网络^[21]稍差,这可能是因为 VGG 的网络层次更深。

表 6 分类器对比实验

Tab. 6 Comparison experiment of discriminator

	精确率	召回率	F1
本文分类器	0.877	0.873	0.874
ResNet18	0.855	0.847	0.849
ResNet50	0.870	0.863	0.865
VGG16	0.928	0.929	0.928

表 7 为本文网络与其他去模糊网络的性能对比。测试数据集包含三种模糊度的光学模糊图像,最终得分为所有恢复图像得分的平均值。对比网络均为单一网络,因此在训练时使用清晰-中度模糊子数据集进行训练。通过表 7 中数据对比可以看出,本文的去模糊网络在恢复光学多重模糊上略优于其他算法。

然而对比 4.2 节发现,模型在显微镜实拍的测试集下的表现弱于合成测试集,这是因为实际

表 7 去模糊网络模型对比试验

Tab. 7 Comparison experiment of deblurring networks

	SSIM	PSNR
本文网络	0.520	23.1
SRGAN ^[7]	0.245	16.6
SRResNet ^[7]	0.415	20.5
VDSR ^[22]	0.291	17.6
SRCNN ^[23]	0.325	20.2

拍摄的离焦图像中,存在光照变化等随机因素的影响,同时高斯模糊虽然是离焦模糊的主因,但仍存在其他因素参与光学模糊的形成,因此对光学离焦图像恢复指标会稍逊于合成数据集。但如 4.5 节的实验所示,此时的恢复效果并不影响对零件特征的识别。

4.5 去模糊网络对检测精度的影响

正如前文所述,本文去模糊网络的最终目的是零件检测。图 10 为实验所使用的样板零件,该零件上设计了 A,B,C,D 四个台阶平面。零件通过精密加工技术获取,每个台阶均使用激光打标技术得到尺寸相同的矩形框。

表 8 检测精度对比实验

Tab. 8 Comparison experiment of detection accuracy μm
(μm)

		基准距离	平均误差	最大误差
轻度模糊组	模糊	847.46	6.66	9.42
	恢复		5.49	8.55
中度模糊组	模糊		12.45	12.69
	恢复		3.62	5.44
重度模糊组	模糊		12.43	16.02
	恢复		3.49	4.18

在实际拍摄时,首先将相机对平面 A 进行对焦,此时拍摄的平面 B、平面 C 和平面 D 分别对应轻度模糊、中度模糊和重度模糊图像,相机的像素当量为 $2.49\text{ }\mu\text{m}/\text{pixel}$ 。

在进行检测时,首先在清晰图像上测量得到矩形长边间的距离,作为检测基准;其次,在模糊图像上和恢复图像上测量矩形长边间的距离,与基准进行比较,结果如表 8 所示。其中在轻度模糊组、中度模糊组和重度模糊组上,用恢复图像测量得到的平均误差分别比恢复前降低了 $1.17\text{ }\mu\text{m}$, $8.83\text{ }\mu\text{m}$ 和 $8.94\text{ }\mu\text{m}$ 。可以看出,恢复图像上

距离的检测误差显著减少,本文的图像增强有助于提高边缘检测准确率。

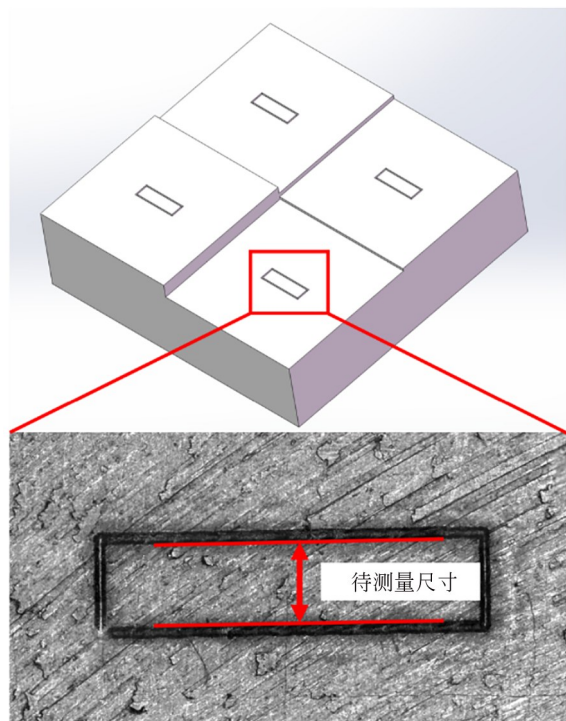


图 10 实验样件

Fig. 10 Experimental parts for precision measurement

参考文献:

- [1] SAGE D, DONATI L, SOULEZ F, *et al.* DeconvolutionLab2: an open-source software for deconvolution microscopy[J]. *Methods*, 2017, 115: 28-41.
- [2] LIN F, JIN C H. An improved Wiener deconvolution filter for high-resolution electron microscopy images[J]. *Micron*, 2013, 50: 1-6.
- [3] ZHANG M N, FANG Y Y, NI G X, *et al.* Pixel screening based intermediate correction for blind deblurring[C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. New Orleans, LA, USA. IEEE, 2022: 1-9.
- [4] XU Z H, CHEN H S, LI Z H. Fast blind deconvolution using a deeper sparse patch-wise maximum gradient prior[J]. *Signal Processing: Image Communication*, 2021, 90: 116050.
- [5] LEVIN A, WEISS Y, DURAND F, *et al.* Understanding and evaluating blind deconvolution algorithms[C]. 2009 *IEEE Conference on Computer Vision and Pattern Recognition*. Miami, FL, USA. IEEE, 2009: 1964-1971.
- [6] FERGUS R, SINGH B, HERTZMANN A, *et al.* Removing camera shake from a single photograph [J]. *ACM Transactions on Graphics*, 2006, 25 (3): 787-794.
- [7] 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: 105-114.
- [8] YANG F Z, YANG H, FU J L, *et al.* Learning texture transformer network for image super-resolution [C]. 2020 *IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*. Seattle, WA, USA. IEEE, 2020: 5790-5799.
- [9] LU Z S, LI J C, LIU H, *et al.* Transformer for single image super-resolution [C]. 2022 *IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW)*. New Orleans, LA, USA. IEEE, 2022: 456-465.
- [10] LIU Y, CHEN X, PENG H, *et al.* Multi-focus image fusion with a deep convolutional neural net-

5 结 论

本文针对显微视觉中存在的多重模糊问题,提出了一种将模糊度判别器和多分支恢复网络相结合的端到端去模糊模型。模糊度判别器使用 Vision-Transformer(VIT)模型与频域处理相结合,利用深度学习网络提取频域特征的方法,进行图像模糊度的识别,经过实验在测试集上表现良好,判别准确率为 0.94。多分支恢复网络具有各分支能够学习不同特征的优点,能够对不同模糊度的子图像进行针对恢复,提高图像整体恢复质量,在测试集上去模糊后的图像 SSIM 指标平均提升了 0.158。集成两者得到的端到端去模糊网络在对多重模糊图像的恢复上表现良好,并且能适应一定程度的模糊度变化,并且在实际光学模糊下的视觉检测中有效减少了因模糊带来的误差。此外,分块后若相邻块采用不同模糊度处理,拼接后图块的边缘会有不连续现象,不过对多数装配作业情况,单像素的边界物理线宽约 2~3 μm ,对装配中测量精度的影响是可以接受的。本文提出的方法为解决显微视觉中的多重模糊问题,提供了新思路。

- work[J]. *Information Fusion*, 2017, 36: 191-207.
- [11] WU P, JIANG L M, HUA Z, *et al.* Multi-focus image fusion: transformer and shallow feature attention matters[J]. *Displays*, 2023, 76: 102353.
- [12] WANG C, ZHAO Z Y, REN Q Q, *et al.* A novel multi-focus image fusion by combining simplified very deep convolutional networks and patch-based sequential reconstruction strategy[J]. *Applied Soft Computing*, 2020, 91: 106253.
- [13] 赵宇轩. 单图超分辨率光片显微成像技术及其在活细胞中的应用研究[D]. 武汉: 华中科技大学, 2022.
- ZHAO Y X. *Single-image Super-Resolution Light Sheet Microscopy and Its Application in Living Cells* [D]. Wuhan: Huazhong University of Science and Technology, 2022. (in Chinese)
- [14] 朱威铭. 基于生成对抗学习的煤岩显微图像超分辨率重建[D]. 江苏: 中国矿业大学, 2023.
- ZHU W M. *Super-resolution Reconstruction of Coal and Rock Microscopic Images Based on Generative Adversarial Learning* [D]. Jiangsu: China University of Mining and Technology, 2023. (in Chinese)
- [15] 陈浩. 基于视觉特征的显微图像预处理研究[D]. 江苏: 中国矿业大学, 2020.
- CHEN H. *Microscopic Image Preprocessing Based on Visual Features* [D]. Jiangsu: China University of Mining and Technology, 2020. (in Chinese)
- [16] 徐征, 陈聿夫, 孙谦, 等. 机器视觉精密测量中的显微光学聚焦[J]. *光学 精密工程*, 2016, 24(9): 2095-2100.
- XU Z, CHEN Y F, SUN Q, *et al.* Auto-focusing in optical microscopy for machine-vision-based precise measurement [J]. *Opt. Precision Eng.*, 2016, 24(9): 2095-2100. (in Chinese)
- [17] 周丽平, 孙志峻, 张泉. 显微视觉系统的自动聚焦及控制[J]. *光学 精密工程*, 2013, 21(3): 807-812.
- ZHOU L P, SUN Z J, ZHANG Q. Auto-focusing and control of micro-vision system [J]. *Opt. Precision Eng.*, 2013, 21(3): 807-812. (in Chinese)
- [18] MAIER A. Efficient focus assessment for a computer vision-based Vickers hardness measurement system[J]. *Journal of Electronic Imaging*, 2012, 21(2): 021114.
- [19] 斯捷, 肖雄, 李泾, 等. 基于生成对抗网络的多幅离焦图像超分辨率重建算法[J]. *计算机工程*, 2021, 47(9): 266-273.
- SI J, XIAO X, LI J, *et al.* Super-resolution reconstruction algorithm with multi-frame defocused images based on generative adversarial network [J]. *Computer Engineering*, 2021, 47(9): 266-273. (in Chinese)
- [20] HE K M, ZHANG X Y, REN S Q, *et al.* Deep residual learning for image recognition [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA. IEEE, 2016: 770-778.
- [21] SIMONYAN K, ZISSERMAN A. Very deep convolutional networks for large-scale image recognition [EB/OL]. 2014: *arXiv*: 1409.1556. <http://arxiv.org/abs/1409.1556>
- [22] KIM J, LEE J K, LEE K M. Accurate image super-resolution using very deep convolutional networks [C]. 2016 *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. Las Vegas, NV, USA. IEEE, 2016: 1646-1654.
- [23] DONG C, LOY C C, HE K M, *et al.* Image super-resolution using deep convolutional networks [J]. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 2016, 38(2): 295-307.

作者简介:



徐 征(1973—),男,河南郑州人,博士,副研究员,1997年和2000年于吉林工业大学分别获得学士和硕士学位,2004年于大连理工大学获得博士学位,主要从事显微视觉、智能微装配、微纳流体等方向的研究工作。E-mail: xuzheng@dlut.edu.cn

通讯作者:



王晓东(1967—),男,黑龙江哈尔滨人,教授,博士生导师,1989年于南京航空学院获得学士学位,1992年于哈尔滨船舶工程学院获得硕士学位,1995年于哈尔滨工业大学获得博士学位。主要从事微装配技术与系统、精密仪器设计与制造等方面的研究。E-mail: xdwang@dlut.edu.cn