

# 基于深度学习的非视域成像

于亭义<sup>1,2</sup>, 乔木<sup>1,2</sup>, 刘红林<sup>1\*</sup>, 韩申生<sup>1</sup>

<sup>1</sup>中国科学院上海光学精密机械研究所量子光学重点实验室, 上海 201800;

<sup>2</sup>中国科学院大学材料与光电研究中心, 北京 100049

**摘要** 针对非相干光照明下的非视域成像问题, 提出一种基于深度学习的解决方法。结合计算机视觉领域中经典的语义分割及残差模型, 构造一种 URNet 网络结构, 并改进了经典瓶颈层结构。实验结果表明, 改进的网络可以恢复更多的图像细节, 并具有一定泛化性, 相比于基于非相干光照明的散斑自相关成像技术, 该网络恢复性能有较大提升。

**关键词** 成像系统; 非视域成像; 深度学习; 语义分割; 残差模型

中图分类号 O438

文献标识码 A

doi: 10.3788/AOS201939.0711002

## Non-Line-of-Sight Imaging Through Deep Learning

Yu Tingyi<sup>1,2</sup>, Qiao Mu<sup>1,2</sup>, Liu Honglin<sup>1\*</sup>, Han Shensheng<sup>1</sup>

<sup>1</sup>Key Laboratory for Quantum Optics, Shanghai Institute of Optics and Fine Mechanics,

Chinese Academy of Sciences, Shanghai 201800, China;

<sup>2</sup>Center of Materials Science and Optoelectronics Engineering, University of Chinese Academy of Sciences,

Beijing 100049, China

**Abstract** Aiming at the problem of non-line-of-sight imaging under incoherent illumination, we propose a solution based on deep learning. Combining the classical semantic segmentation and residual model in the field of computer vision, a URNet network structure is constructed and the classical bottleneck layer structure is improved. The experimental results show that the improved model has more details of recovery images and generalization ability. Compared with speckle autocorrelation imaging method under incoherent illumination, the recovery performance of this method is greatly improved.

**Key words** imaging systems; non-line-of-sight imaging; deep learning; semantic segmentation; residual model

**OCIS codes** 110.3000; 100.4996; 100.2000

## 1 引 言

传统非视域成像技术采用激光主动照明, 光信号经过中介面反射或散射, 之后照射到物体上, 经过物体反射的光再次通过中介面, 最后被接收器接收。目前, 该技术主要有几种形式:

1) 基于激光距离的选通成像技术<sup>[1-2]</sup>。该技术根据物体与接收器之间的距离估计光信号传播时间, 在两次反射期间选通门打开, 其余时间关闭, 从而消除其他一次或多次反射的干扰。

2) 基于超快激光的瞬态成像技术<sup>[3-5]</sup>。该技术将超快激光照射到中介面上的不同位置, 利用高分辨的条纹相机接收多个图像, 根据图像反演算法, 最

终构建出物体的三维图像<sup>[6-9]</sup>。

3) 基于光子计数的探测成像技术<sup>[10]</sup>。该技术利用单光子雪崩二极管 (SPAD) 获取物体反射回的单光子信号并构建光子计数直方图, 通过反投影算法进行图像重构。

4) 基于数字全息成像技术<sup>[11-12]</sup>。该技术将激光分为参考光和物光, 物光照射到物体后被反射, 与参考光在毛玻璃表面发生干涉, 干涉结果通过毛玻璃并经过透镜成像到接收面, 最后采用全息成像技术将接收到的信号进行恢复。

以上方法均是基于相干光源照明采集弹道光信号, 过滤散射光信号。在散射较强的情况下, 弹道光衰减迅速, 无法成像。利用非相干光照明的波前整

收稿日期: 2019-01-29; 修回日期: 2019-02-28; 录用日期: 2019-03-22

基金项目: 国家重点研发计划(2016YFC0100600)

\* E-mail: hlliu4@hotmail.com

形<sup>[13-15]</sup>和基于散射介质记忆效应的散斑自相关成像技术<sup>[16-17]</sup>,克服了弹道光成像技术的缺陷,并可以实现散射光成像。但波前整形需要目标先验信息,调节速度慢,过程复杂;基于散射介质记忆效应的成像方法受限于记忆效应范围,成像视场角小,之后虽然出现了扩大记忆效应范围的方法<sup>[18-19]</sup>,却仍然无法满足实际需求。

深度学习(DL)方法被应用到非视域成像技术中<sup>[20-22]</sup>。该技术采用数据驱动方法而非建立确定的模型,通过向数据集学习的方式解决问题。此方法有较好的适应性与灵活性,性能往往能够超越基于模型的方法。Li等<sup>[20-21]</sup>利用神经网络实现对单一散射体的非视域成像;Li等<sup>[22]</sup>对同一网络进行多次训练,实现了对不同散射体的成像。但以上方法均采用相干光源照明、透射方式成像,而在实际应用中,存在许多仅适用于非相干光照明的场景。针对该问题,提出一种基于深度学习的非相干光照明的非视域成像技术,利用计算机视觉领域中的经典模型,构造出一种新的卷积神经网络结构,并能改进传统瓶颈层结构。结果表明,该方法可实现端对端的图像恢复,恢复性能相比于传统的散斑自相关成像技术有较大提升。

## 2 理论分析

### 2.1 传播理论

在非相干光源照明下的非视域成像模型中,物体反射的非相干光经空间自由传播,之后被中介面散射或反射,再经空间自由传播后被探测器接收。与相干光传播方式不同,非相干光传播采用互强度传播公式。如图1所示,物面 $\sigma$ 经过自由空间传播,之后得到中介面 $\tau$ ,根据范希特-泽尼克定理,有

$$J(p_1, p_2) = \langle E^*(p_1)E(p_2) \rangle = \int_{\sigma} I(s) \frac{\exp[ik(R_1 - R_2)]}{R_1 R_2} ds, \quad (1)$$

式中: $p_1$ 和 $p_2$ 为中介面 $\tau$ 上两点; $E(p_2)$ 和 $E^*(p_1)$ 分别为 $p_2$ 和 $p_1$ 的光场和光场共轭; $I(s)$ 为物面 $\sigma$ 中 $s$ 点的光强; $k$ 为波数; $R_1$ 和 $R_2$ 为 $s$ 到 $p_1$ 和 $p_2$ 的距离; $J(p_1, p_2)$ 为被中介面调制前的互强度。

物光经过空间自由传播到中介面 $\tau$ 后,被中介面调制。由于中介面为一个粗糙面,实验中采用毛玻璃代替中介面,用调制函数 $T(\cdot)$ 表示该作用,由此推出调制后的光场为

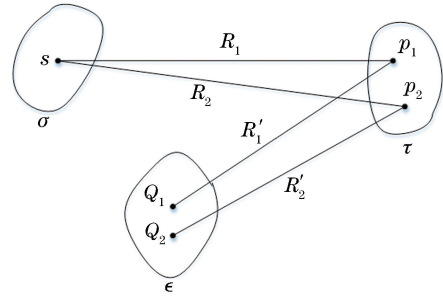


图1 非相干光传播图

Fig. 1 Incoherent light propagation

$$E_{\text{mod}}(p) = T(p)E(p), \quad (2)$$

式中: $E_{\text{mod}}(p)$ 为经过中介面调制后的光场。由(1)式可推出调制后的互强度为

$$J_{\text{mod}}(p_1, p_2) = \langle E_{\text{mod}}^*(p_1)E_{\text{mod}}(p_2) \rangle = \langle E^*(p_1)T^*(p_1)E(p_2)T(p_2) \rangle, \quad (3)$$

式中: $J_{\text{mod}}(p_1, p_2)$ 为经过调制后的互强度。光场经中介面调制后,最终经过空间自由传播被接收。利用范希特-泽尼克定理可推出接收像面 $\epsilon$ 中的点 $Q_1$ 和点 $Q_2$ 的互强度 $J(Q_1, Q_2)$ ,可得

$$J(Q_1, Q_2) = \frac{1}{\lambda^2} \iint_{\tau} J_{\text{mod}}(p_1, p_2) \cdot \frac{\exp[-ik(R'_1 - R'_2)]}{R'_1 R'_2} dp_1 dp_2, \quad (4)$$

式中: $\lambda$ 为光波长; $R'_1$ 和 $R'_2$ 分别为 $p_1$ 到 $Q_1$ 以及 $p_2$ 到 $Q_2$ 的距离。接收像面 $\epsilon$ 上任意一点 $Q$ 的光强 $I(Q) = J(Q, Q)$ ,由此得

$$I(Q) = \frac{1}{\lambda^2} \iint_{\tau} \langle E^*(p_1)T^*(p_1)E(p_2)T(p_2) \rangle \cdot \frac{\exp[-ik(R'_1 - R'_2)]}{R'_1 R'_2} dp_1 dp_2. \quad (5)$$

由于计算像面的公式过于复杂,用一个点扩展函数 $\mathcal{F}$ 表示该系统响应,即

$$I(x, y) = \mathcal{F}\{O(x_0, y_0)\}, \quad (6)$$

式中: $O(x_0, y_0)$ 为物面 $\sigma$ 上的物体信息; $I(x, y)$ 为接收像面 $\epsilon$ 上的接收信息。对于非视域成像技术,通过 $I(x, y)$ 反解出 $O(x_0, y_0)$ :

$$O(x_0, y_0) = \mathcal{R}\{I(x, y)\}, \quad (7)$$

式中: $\mathcal{R}$ 为点扩展函数 $\mathcal{F}$ 的逆函数。理论上无法解出 $\mathcal{R}$ 的解析表达式,因此采用深度学习中的卷积神经网络,通过学习的方式求得该函数的近似解。

卷积神经网络使用包含物体 $O_n$ 及对应像 $I_n$ 的数据集,其中 $n=1, 2, \dots, N$ 。采用最小化损失函数的方法,得到反向映射函数的近似解 $\mathcal{R}_{\text{learn}}$ 为

$$\mathcal{R}_{\text{learn}} = \underset{\mathcal{R}_\theta, \theta \in \Theta}{\text{argmin}} \sum_{n=1}^N f(O_n, \mathcal{R}_\theta\{I_n\}) + \Psi(\theta), \quad (8)$$

式中： $\Theta$  为网络的参数集合； $\theta$  为集合  $\Theta$  中的元素； $f$  为损失函数； $\Psi(\theta)$  为防止过拟合的正则惩罚。

### 2.2 实验结构

实验装置如图 2 所示。准直 LED 光源 (M530L3, Thorlabs, America) 和透镜  $L_1$  (焦距  $f_1 = 5\text{ cm}$ )、 $L_2$  (焦距  $f_2 = 5\text{ cm}$ ) 构成科勒照明系统, 产生均匀非相干光。由科勒照明系统出射的光被数字微反射镜器件 (DMD, DLP Discovery 4100, Texas Instruments, America) 反射, DMD 反射的光经过空间自由传播后被中介面反射, 反射后的光场最终被相机 (Stingray F-504B, Allied Vision Technologies, Germany) 接收。DMD 与中介面的距离为  $z$ 。当 DMD 上加载灰度图像时, 可将 DMD 看作是一个反射型的物体。该 DMD 的像素数为  $1024\text{ pixel} \times 768\text{ pixel}$ , 像元尺寸为  $13.6\text{ }\mu\text{m} \times 13.6\text{ }\mu\text{m}$ , 物体与中介面距离  $z = 15\text{ cm}$ , 中介面采用粗糙的毛玻璃片, 相机的像素数为  $2452\text{ pixel} \times$

$2056\text{ pixel}$ , 相机的像元尺寸为  $3.45\text{ }\mu\text{m} \times 3.45\text{ }\mu\text{m}$ 。DMD 上的图像加载和相机快门通过信号源产生的方波脉冲信号同步控制。

### 2.3 URNet 网络结构

为解决非相干光照明下的非视域成像问题, 提出一种 URNet 网络, 该网络结构主要包含 U-Net 网络结构<sup>[23]</sup> 和残差神经 (ResNet) 网络结构<sup>[24]</sup>。U-Net 网络结构利用下采样提取特征和上采样恢复图像, 实现端对端的语义分割; ResNet 网络结构通过跳跃连接解决反向传播时的梯度消失问题, 提高训练速度。

URNet 网络结构如图 3 所示。输入图像首先经过 4 个瓶颈层, 每个瓶颈层后有一个最大池化层, 之后经过一个瓶颈层和 4 个反卷积层, 每个反卷积层后有一个瓶颈层, 最后通过一个普通卷积层输出。为了最大限度地保留信息, 在网络结构中添加跳跃连接, 连接的块具有相同的图像尺寸。

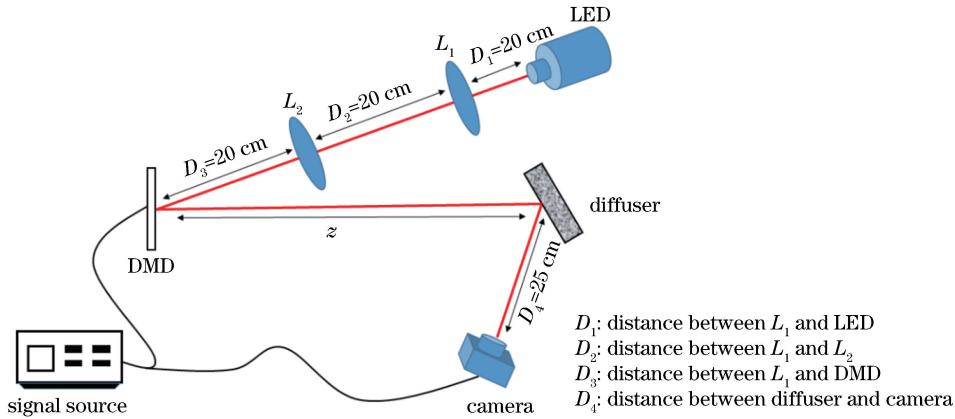


图 2 非视域成像实验结构示意图

Fig. 2 Experiment structure of non-line-of-sight imaging

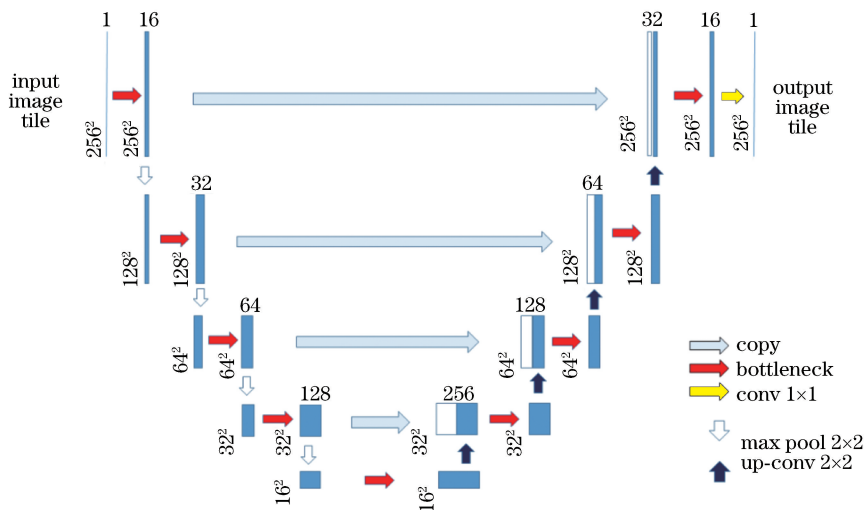


图 3 URNet 网络结构图

Fig. 3 Structural diagram of URNet network

对于经典的瓶颈层结构<sup>[24]</sup>,针对该应用场景进行了改进。如图4所示,图4(a)为经典的3层瓶颈层结构,包含两个 $1\times 1$ 卷积核、两个非线性激活函数和一个 $3\times 3$ 卷积核。图4(b)为改进的瓶颈层结构,在该结构中采用较大的 $7\times 7$ 卷积核代替较小的

$3\times 3$ 卷积核,用 $3\times 3$ 的卷积核代替 $1\times 1$ 的卷积核,相比于较小的卷积核,较大的卷积核可以恢复更多图像细节信息<sup>[25]</sup>。在激活函数后添加批量归一化(BN)层,可以防止过拟合。

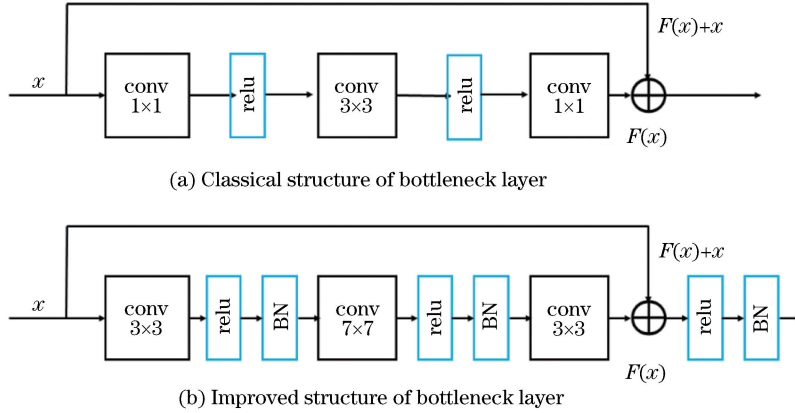


图4 瓶颈层结构图。(a)经典瓶颈层结构;(b)改进的瓶颈层结构

Fig. 4 Structures of bottleneck layer. (a) Classical structure of bottleneck layer. (b) improved structure of bottleneck layer

训练过程中的均方误差损失函数定义为

$$E_{MS} = \frac{1}{wh} \sum_{i=1}^w \sum_{j=1}^h (y(i, j) - \bar{y}(i, j))^2, \quad (9)$$

式中: $w$ 和 $h$ 分别为图像的宽度和高度; $y(i, j)$ 为原始物体图像; $\bar{y}(i, j)$ 为系统恢复的物体图像。

### 3 结果与分析

#### 3.1 模型训练

URNet网络恢复模型主要包含两个部分:训练部分和验证部分。在训练部分,将原始图像和相机采集图像调整大小后作为网络输入。原始图像尺寸为 $28 \text{ pixel} \times 28 \text{ pixel}$ ,考虑到DMD加载图像尺寸要求为 $1024 \text{ pixel} \times 768 \text{ pixel}$ ,将原始图片缩放为 $512 \text{ pixel} \times 512 \text{ pixel}$ ,并将其放在中心位置,对空白部分填充0,相机采集图像也选取中心 $512 \text{ pixel} \times 512 \text{ pixel}$ 部分。为了提高训练速度,将两种输入图像尺寸缩放为 $256 \text{ pixel} \times 256 \text{ pixel}$ 。考虑到信噪比和相机的帧频限制,输入信号频率为2 Hz,即DMD每秒加载两张图像,相机对应地采集两张图像。数据集选自经典人工智能数据集MNIST,约6万张手写阿拉伯数字图像,将数据集分为训练集和验证集两部分,二者比例为9:1,两部分的数据完全分开,没有重合,训练时长为15个Epoch,每个Epoch表示将所有数据训练一遍,使用装有Intel i7CPU和GTX 1070显卡的计算机进行数据处理,一次训练时间为6~8 h。

#### 3.2 模型恢复结果及分析

经典瓶颈层结构与改进的瓶颈层结构模型恢复结果对比图如图5所示。图5(a)为原始物体图像,图5(b)为经过中介面反射后被相机接收的图像,图5(c)为经典瓶颈层模型恢复结果,图5(d)为改进的瓶颈层模型恢复结果。两个模型的训练时间分别为

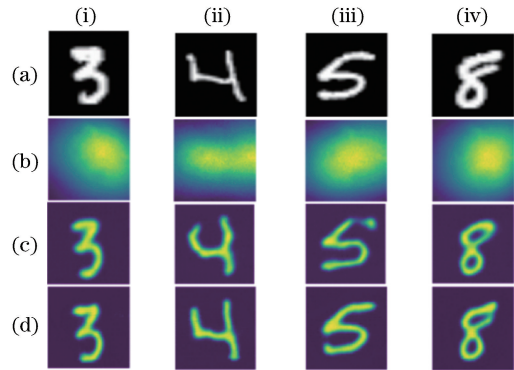


图5 不同瓶颈层模型恢复结果。(a)原始物体图像;(b)相机接收的光斑信号;(c)经典瓶颈层模型恢复结果;(d)改进的瓶颈层模型恢复结果。(i)~(iv)分别代表不同手写字符3,4,5,8

Fig. 5 Retrieving results of different bottleneck layer models. (a) Ground truths; (b) speckle images; (c) retrieving results of classical bottleneck layer model; (d) retrieving results of improved bottleneck layer model; (i)-(iv) represent different handwriting characters 3, 4, 5, and 8, respectively



6.7 h 和 8.4 h, 改进的瓶颈层模型训练时间略长。对比两组恢复结果, 改进的瓶颈层模型具有较好的恢复结果, 尤其在图像细节方面, 恢复结果更加丰富。采用平均余弦距离度量恢复结果与真实结果之间的相似度, 得到的结果分别为 0.8138 和 0.8397。该结果同样说明, 相比于经典瓶颈层模型, 改进的瓶颈层模型具有更好的恢复效果。

为了检验模型的泛化性, 对非 MNIST 数据集中的图像进行恢复。图 6 为手写英文字母的恢复结果, 图 6(a) 为原始图像; 图 6(b) 为相机采集到的光斑信号; 图 6(c) 为 URNet 网络恢复结果。对比原始图像与恢复图像可知, URNet 网络对手写英文字母具有较好的恢复结果。此结果进一步说明, 当恢复目标与训练集中的图像具有相似的结构时, URNet 网络对该目标适用。

为了进一步检验模型的可靠性, 对模型采集信号受到外界影响的情况进行讨论。实验中对采集到的光斑进行部分遮挡, 被遮挡的部分灰度值置为 0, 再将光斑信号输入到模型中。结果如图 7 所示, 图 7(a) 为被遮挡后的光斑信号图; 图 7(b) 为模型恢复结果; 图像尺寸为 256 pixel × 256 pixel。分析结果发现, 当被遮挡部分占比为 0.0625 和 0.1250 时, 恢

复图像开始发生变形, 但基本形状保持完整; 当被遮挡部分占比为 0.1250 和 0.1875 时, 恢复图像发生较大变形, 但仍然可以分辨出该图形为 9; 当被遮挡部分占比为 0.3125 时, 恢复图像已无法辨认。由上述结果可得, 当物体被遮挡的像素在一定范围内时, 模型依然有效。

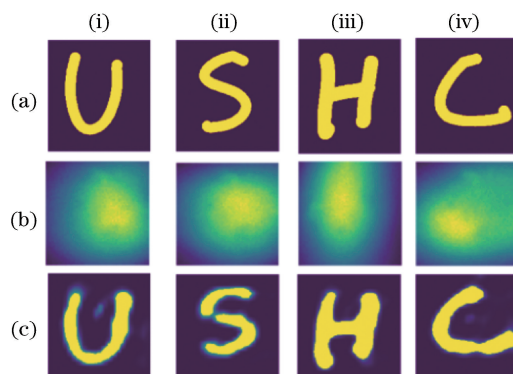


图 6 手写英文字母恢复结果。(a) 原始物体图像; (b) 相机接收的光斑信号; (c) URNet 模型恢复结果。(i)~(iv) 分别代表不同手写英文字母 U, S, H, C  
Fig. 6 Retrieving results of handwriting English letters. (a) Ground truths; (b) speckle images; (c) retrieving results of URNet model; (i)-(iv) represent different handwriting English letters U, S, H, and C, respectively

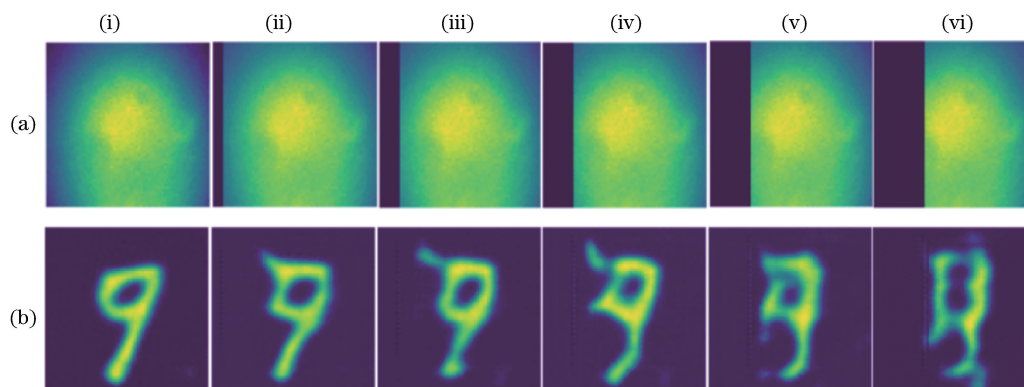


图 7 受外界影响下的模型恢复结果。(a) 光斑信号被遮挡结果; (b) URNet 模型恢复结果。(i)~(vi) 分别代表遮挡比例 0, 0.0625, 0.1250, 0.1875, 0.2500, 0.3125

Fig. 7 Retrieving results under external interference. (a) Results of occlusion of speckle signal; (b) retrieving results of URNet model. (i)-(vi) represent occlusion proportion of 0, 0.0625, 0.1250, 0.1875, 0.2500, and 0.3125, respectively

散斑自相关成像技术利用散射介质记忆效应, 可以在非相干光照明情况下实现非视域成像。图 8 所示为 URNet 网络与散斑自相关恢复结果对比图, 图 8(a) 为原始图像; 图 8(b) 为 URNet 网络恢复结果; 图 8(c) 为散斑自相关恢复结果; (i)~(iv) 对应不同的手写字符。由于散斑自相关成像技术在实验中无法对较大目标成像, 因此采用仿真模拟结果,

在仿真过程中添加了 10 dB 的高斯随机噪声, 并对 URNet 网络也添加了相同噪声, 计算所得两种结果的平均相似度分别为 0.7455 和 0.8366。由两种结果的对比可得, 相比于散斑自相关成像技术, URNet 网络恢复结果具有较大提升, 说明该网络对噪声干扰更加稳健。

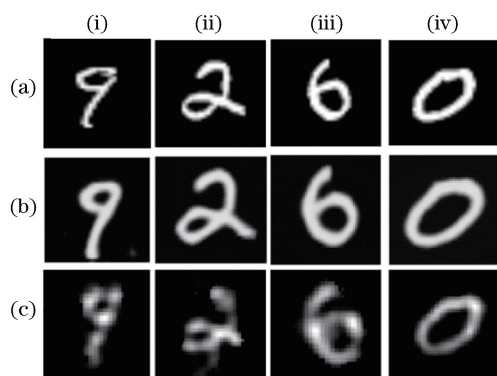


图 8 不同模型恢复结果。(a)原始物体图像;(b) URNet 网络恢复结果;(c)散斑自相关恢复结果。(i)~(iv)代表不同手写字符 9,2,6,0

Fig. 8 Retrieving results of different models. (a) Ground truths; (b) retrieving results of URNet model; (c) retrieving results of speckle autocorrelation model; (i)-(iv) different handwriting characters 9, 2, 6, and 0, respectively

## 4 结 论

针对非相干光照明情况下的非视域成像问题,提出了基于深度学习的解决方案。结合经典的 U-Net 网络和 ResNet 网络特点,构建出 URNet 网络。在此基础上,对经典的瓶颈层结构进行改进,对比改进前后的网络恢复结果发现,改进的网络可以恢复更多的图像细节。利用手写英文字母数据,对模型的泛化性进行检验。通过遮挡采集到的光斑信号,验证模型在受到外界影响下的恢复情况。实验结果表明,URNet 网络具有较好的泛化性,并在外界影响下仍能得到较好的恢复结果。相比于散斑自相关成像技术,URNet 网络恢复结果质量更高。相比于传统方法,基于深度学习的成像方法具有较好的成像效果,但是该方法实际上是一种统计学习方法,并不具有普适性,对与学习库没有相似分类特征的对象,该方法成像效果差,甚至无法成像。因此,以后的研究工作重点将为不同种类复杂目标的非视域成像。

## 参 考 文 献

- [1] Repasi E, Lutzmann P, Steinvall O, *et al.* Advanced short-wavelength infrared range-gated imaging for ground applications in monostatic and bistatic configurations[J]. *Applied Optics*, 2009, 48(31): 5956-5969.
- [2] Laurenzis M, Christnacher F, Velten A. Study of a dual mode SWIR active imaging system for direct imaging and non-line-of-sight vision[J]. *Proceedings of SPIE*, 2015, 9465: 946509.
- [3] Pandharkar R, Velten A, Bardagjy A, *et al.* Estimating motion and size of moving non-line-of-sight objects in cluttered environments [C]//2011 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 20-25, 2011, Colorado Springs, CO, USA. New York: IEEE, 2011: 265-272.
- [4] Naik N, Zhao S, Velten A, *et al.* Single view reflectance capture using multiplexed scattering and time-of-flight imaging [J]. *ACM Transactions on Graphics*, 2011, 30(6): 171.
- [5] Heide F, Xiao L, Heidrich W, *et al.* Diffuse mirrors: 3D reconstruction from diffuse indirect illumination using inexpensive time-of-flight sensors [C]//2014 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), June 23-28, 2014, Columbus, OH, USA. New York: IEEE, 2014: 3222-3229.
- [6] Wu D, Velten A, O'Toole M, *et al.* Decomposing global light transport using time of flight imaging[J]. *International Journal of Computer Vision*, 2014, 107(2): 123-138.
- [7] Kirmani A, Hutchison T, Davis J, *et al.* Looking around the corner using ultrafast transient imaging [J]. *International Journal of Computer Vision*, 2011, 95(1): 13-28.
- [8] Gupta O, Willwacher T, Velten A, *et al.* Reconstruction of hidden 3D shapes using diffuse reflections [J]. *Optics Express*, 2012, 20(17): 19096-19108.
- [9] Velten A, Willwacher T, Gupta O, *et al.* Recovering three-dimensional shape around a corner using ultrafast time-of-flight imaging [J]. *Nature Communications*, 2012, 3: 745.
- [10] Garipey G, Tonolini F, Henderson R, *et al.* Detection and tracking of moving objects hidden from view[J]. *Nature Photonics*, 2016, 10(1): 23-26.
- [11] Singh A K, Naik D N, Pedrini G, *et al.* Looking through a diffuser and around an opaque surface: a holographic approach[J]. *Optics Express*, 2014, 22(7): 7694-7701.
- [12] Vinu R V, Kim K, Somkuwar A S, *et al.* Single-shot optical imaging through scattering medium using digital in-line holography [J/OL]. (2016-03-24) [2019-01-01]. <https://arxiv.org/abs/1603.07430>.
- [13] Yaqoob Z, Psaltis D, Feld M S, *et al.* Optical phase conjugation for turbidity suppression in biological samples[J]. *Nature Photonics*, 2008, 2(2): 110-115.

- [14] Vellekoop I M, Lagendijk A, Mosk A P. Exploiting disorder for perfect focusing[J]. *Nature Photonics*, 2010, 4(5): 320-322.
- [15] Liu Y, Ma C, Shen Y C, *et al.* Focusing light inside dynamic scattering media with millisecond digital optical phase conjugation[J]. *Optica*, 2017, 4(2): 280-288.
- [16] Bertolotti J, van Putten E G, Blum C, *et al.* Non-invasive imaging through opaque scattering layers[J]. *Nature*, 2012, 491(7423): 232-234.
- [17] Katz O, Heidmann P, Fink M, *et al.* Non-invasive single-shot imaging through scattering layers and around corners via speckle correlations[J]. *Nature Photonics*, 2014, 8(10): 784-790.
- [18] Judkewitz B, Horstmeyer R, Vellekoop I M, *et al.* Translation correlations in anisotropically scattering media[J]. *Nature Physics*, 2015, 11(8): 684-689.
- [19] Li L, Li Q, Sun S, *et al.* Imaging through scattering layers exceeding memory effect range with spatial-correlation-achieved point-spread-function[J]. *Optics Letters*, 2018, 43(8): 1670-1673.
- [20] Li S, Deng M, Lee J, *et al.* Imaging through glass diffusers using densely connected convolutional networks[J]. *Optica*, 2018, 5(7): 803-812.
- [21] Lyu M, Wang H, Li G, *et al.* Exploit imaging through opaque wall via deep learning[J/OL]. (2017-08-09) [2019-01-01]. <https://arxiv.org/abs/1708.07881>.
- [22] Li Y Z, Xue Y J, Tian L. Deep speckle correlation: a deep learning approach toward scalable imaging through scattering media[J]. *Optica*, 2018, 5(10): 1181-1190.
- [23] Ronneberger O, Fischer P, Brox T. U-Net: convolutional networks for biomedical image segmentation[M]//Navab N, Hornegger J, Wells W, *et al.* *Medical image computing and computer-assisted intervention-MICCAI 2015*. Lecture notes in computer science. Cham: Springer, 2015, 9351: 234-241.
- [24] 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), June 27-30, 2016, Las Vegas, NV, USA. New York: IEEE, 2016: 770-778.
- [25] Peng C, Zhang X Y, Yu G, *et al.* Large kernel matters—improve semantic segmentation by global convolutional network[C]//2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), July 21-26, 2017, Honolulu, HI, USA. New York: IEEE, 2017: 1743-1751.