

DOI: 10.19650/j.cnki.cjsi.J2209518

# 高速摄影仪成像噪声去除方法研究\*

陈怀安<sup>1</sup>, 卢小银<sup>2</sup>, 单奕萌<sup>3</sup>, 阚艳<sup>3</sup>, 金一<sup>1</sup>

(1. 中国科学技术大学工程科学学院 合肥 230026; 2. 合肥中科君达视界技术股份有限公司 合肥 230051;  
3. 国营芜湖机械厂 芜湖 241000)

**摘要:**高速摄影仪在超高帧率下(>10 000 FPS)易产生噪声,该噪声分布复杂,难以获取与有噪图像完全对应的清晰图像。针对该问题,提出一种基于非理想配对图像的卷积去噪网络训练方法。首先利用高速和低速摄影仪拍摄相同场景图像,获得有噪图像及与其对应的非理想配对清晰图像;然后,建立基于卷积神经网络的深度去噪模型,结合亮度一致化和图像对齐方法,实现非理想配对图像的监督学习,从而去除成像噪声;最后,引入模型量化技术将模型参数和激活值由32位浮点数量化为8位定点数,降低模型大小、内存需求和运行时间。实验结果表明,提出的去噪方法可有效去除高速摄影仪成像噪声,相比于其他方法,去噪图像峰值信噪比提高1.96 dB,结构相似性提高1.95%;通过模型量化,模型大小降低4倍,内存需求降低45.62%,运行时间降低37.5%。

**关键词:**高速摄影;图像去噪;非理想配对;卷积神经网络;模型量化

中图分类号: TH744 文献标识码: A 国家标准学科分类代码: 460.4035

## Research on the method of imaging noise removal for high-speed camera

Chen Huaian<sup>1</sup>, Lu Xiaoyin<sup>2</sup>, Shan Yimeng<sup>3</sup>, Kan Yan<sup>3</sup>, Jin Yi<sup>1</sup>

(1. School of Engineering Science, University of Science and Technology of China, Hefei 230026, China;  
2. Hefei Agile Device Co., Ltd., Hefei 230051, China; 3. Wuhu State-owned Factory of Machining, Wuhu 241000, China)

**Abstract:** The high-speed camera tends to produce noise when it works under an ultra-high frame rate (>10 000 FPS). The distribution of the produced noise is complex. Thus, it is very difficult to obtain the noisy-clean image pairs. To address this problem, a training scheme for the convolutional denoising networks based on the non-ideal noisy-clean pairs is proposed. Firstly, the cameras with high and low frame rates are leveraged to capture the images of the same scenes, through which the noisy images and the corresponding non-ideal paired clean images are achieved. Then, a deep denoising model based on a convolutional neural network is formulated to achieve supervised learning with the non-ideal noisy-clean image pairs by utilizing the brightness consistency and image alignment methods, contributing to removing imaging noise. Finally, a model quantization technique is introduced to quantize the values of parameters and activations, which helps them to transform the 32 bit floating-point number to an 8 bit fixed-point number, and thus greatly reduces the model size, memory consumption, and running time. Experimental results show that the proposed denoising method can effectively remove the imaging noise of a high-speed camera. Compared with other methods, the peak signal-to-noise ratio and structural similarity of the denoised image are improved by at least 1.96 dB and 1.95%, respectively. In addition, with the help of the model quantization technique, the model size is reduced by 4 times, and the memory consumption and running time are decreased by 45.62% and 37.5%, respectively.

**Keywords:** high-speed photography; image denoising; non-ideal paired; convolutional neural network; model quantization

## 0 引言

高速摄影技术具有高时间分辨率、非接触、快响应等优点,可在极短时间内捕获高速运动目标状态,在缺陷检测<sup>[1]</sup>、参数测量<sup>[2]</sup>、诊断测试<sup>[3]</sup>等领域具有广泛应用。近年来,随着电子信息与仪器科学的发展,基于互补金属氧化物半导体(complementary metal oxide semiconductor, CMOS)的数字式高速摄影仪成为实现高速摄影技术的重要仪器。然而,相比于低速摄影仪,高速摄影仪曝光时间极短,导致其图像信号弱,噪声信号相对较强;同时,由于成像数据量急剧增大,需要多路模数转换器才能实现数据采集,高频采样下易导致电压变化,致使图像出现条纹等噪声,严重影响后续诊断测试、参数测量等任务的精度,亟需开展针对高速摄影仪的成像噪声去除方法研究。

传统图像去噪算法可在一定程度上缓解高速摄影仪的成像噪声问题。例如,在空间域通过局部或非局部滤波(如高斯滤波)可实现图像去噪,该类算法实现简单,计算复杂度低,但在去除噪声的同时也会滤除大量纹理细节,去噪效果较差。随着图像去噪技术的发展,基于先验的方法逐渐成为主流。该类方法假设图像退化模型已知,通过精心设计图像先验作为正则项,实现噪声去除。典型算法如三维块匹配协同滤波算法(block-matching and 3D filtering, BM3D)<sup>[4]</sup>、加权核范数最小化算法<sup>[5]</sup>、K-奇异值分解算法<sup>[6]</sup>等。其中, BM3D 是最成功的方法之一,它通过结合稀疏性先验和非局部自相似先验实现图像去噪。在此基础上, Mairal 等<sup>[7]</sup>分析了稀疏性先验和非局部自相似先验的内在联系,利用稀疏编码方法结合两种先验,进一步提升了去噪性能。Dong 等<sup>[8]</sup>引入稀疏编码噪声的概念,进而提出非局部中心化稀疏表征模型,成功应用在图像去噪等任务。尽管上述方法可以实现图像噪声去除,但受限于人工设计的图像先验,该类方法去噪性能有限,且针对不同图像往往需要手动调整模型参数,给实际应用带来困难。

近年来,深度学习模型在各类视觉任务<sup>[9-11]</sup>中取得了突破性进展,促使大量工作研究基于卷积神经网络的图像去噪模型<sup>[12-16]</sup>。例如, Zhang 等<sup>[13]</sup>利用残差学习和批归一化构建了深度去噪卷积网络(denoising convolutional neural networks, DnCNNs),取得了比传统先验方法更好的性能;在此基础上,该团队通过将输入图像下采样为子图像并引入噪声强度图,提出了快速灵活去噪网络(fast and flexible denoising convolutional neural network, FFDNet)<sup>[14]</sup>,取得了去噪性能和计算量的良好权衡。Park 等<sup>[15]</sup>提出了基于编解码结构的分层图像去噪网络,利用具有大量参数的深度模型实现了去噪性能的显著提升。Chen 等<sup>[16]</sup>提出了需求导向去噪网络

(demand-oriented network, DONet),实现了图像去噪模型在计算量、参数量和去噪性能三方面的平衡。上述方法虽然能取得很好的图像去噪性能,但都需要大量成对的有噪和清晰图像才能实现模型训练。面向实际场景,往往难以获取该类数据。为此 Lehtinen 等<sup>[17]</sup>提出有噪图像到有噪图像的训练方法,该方法无需清晰图像即可实现模型的监督训练。在此基础上, Krull 等<sup>[18]</sup>证明了通过自监督学习可训练去噪网络。这些方法虽然无需成对的有噪和清晰图像,但其去噪性能无法与利用清晰图像监督训练的模型相比,且往往要求噪声均值为 0。

面向高速摄影仪,上述方法难以适用。一方面,高频采样下的条纹噪声均值不为 0,导致自监督或无监督学习方法难以满足要求;另一方面,噪声均值不为 0 也导致无法采用多帧平均的方法获得与有噪图像完全配对的清晰图像,使得基于成对数据训练的方法难以用来训练去噪模型。为此,本文提出基于非理想配对图像的卷积神经网络去噪模型训练方法。首先,利用基于相同 CMOS 图像传感器研制的高速和低速摄影仪拍摄相同场景图像,将低速摄影仪拍摄的高质量图像看作清晰图像,使高速摄影仪成像质量能够逼近低速摄影仪成像质量。由于摄影仪参数、成像视野不完全一致,难以采集理想配对的有噪图像与清晰图像。为此,利用非理想配对图像训练基于卷积神经网络的深度去噪模型,结合亮度一致化和图像对齐方法,实现非理想配对图像的监督学习,有效去除成像噪声。最后,引入模型量化技术将模型参数和激活值由 32 位浮点数量化为 8 位定点数,减少模型大小、内存需求和运行时间。提出的方法在自主研制的数字式高速摄影仪上得到应用,实验结果表明,提出的方法可有效去除高速摄影仪成像噪声。

## 1 高速摄影仪成像噪声去除方法

### 1.1 非理想配对数据集构建

高速摄影仪在超高帧率下易产生成像噪声,其中因采样电压不一致导致的条纹噪声均值不为 0,因而无法采用多帧平均的方法获得与有噪图像完全对应的清晰图像。为此,本文采用高速和低速摄影仪拍摄相同场景图像,其中,低速摄影仪曝光时间较长且无需高频采样,成像质量更高,因此,其拍摄图像可视为清晰图像;高速摄影仪曝光时间极短,噪声信号相对较强,且其成像数据量极大,需通过多路模数转换器高频采样,易导致图像产生条纹等噪声,成像质量差,其拍摄图像为有噪图像。在实际拍摄过程中,尽管可以通过人工调节曝光时间、增益、成像视野等参数使两个相机拍摄图像尽可能一致,但该精细操作极其浪费时间且仍难以保证两组图像在亮度和成像视野上实现像素级匹配。因此,本文采集非理想配

对的有噪图像和清晰图像,即允许采集的有噪和清晰图像存在一定的位置偏移和亮度偏差,然后利用该图像训练深度去噪模型。

本文采用一台自主研发的帧率可达13 500 FPS的高速摄影仪和一台参数设置在100 FPS的低速摄影仪组成数据采集系统。系统成像视野取两台摄影仪共有视野,图像分辨率为1 280 pixels×960 pixels。本文共采集12个场景图像,通过调节补光灯及相机参数,使每个场景采集25个不同亮度水平的图像。通过该系统,共采集300组非理想配对的有噪和清晰图像。考虑到实际工业检测和测量任务中主要采用灰度图像,本文主要研究灰度图像去噪。图1展示了采集的两组不同亮度水平的有噪和清晰图像,从图中可以看出,高速摄影仪拍摄的有噪图像具有明显的噪声,而低速摄影仪拍摄的图像噪声较少,成像质量较高。

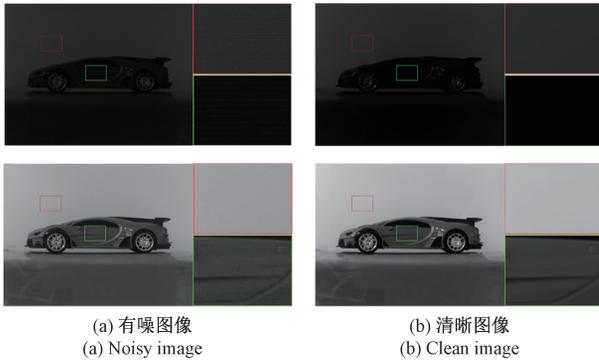


图1 采集的有噪图像和清晰图像

Fig. 1 Illustration of the captured noisy and clean images

## 1.2 非理想配对图像训练方法

### 1) 非理想配对下的图像去噪建模

噪声导致的图像退化模型可以表达为:

$$\mathbf{y} = \mathbf{x} + \boldsymbol{\eta} \quad (1)$$

式中: $\mathbf{y}$ 表示有噪图像, $\boldsymbol{\eta}$ 表示噪声, $\mathbf{x}$ 表示与有噪图像完全配对的无噪清晰图像。图像去噪的目标是从有噪图像 $\mathbf{y}$ 中恢复出清晰图像 $\mathbf{x}$ 。利用上述退化模型,基于深度残差学习建立的去噪模型可以表达为:

$$\hat{\mathbf{x}} = \mathbf{y} - D(\mathbf{y}, \boldsymbol{\Theta}) \quad (2)$$

式中: $D(\cdot)$ 表示噪声预测模型, $\boldsymbol{\Theta}$ 表示模型可学习参数, $\hat{\mathbf{x}}$ 表示预测的去噪图像,参数 $\boldsymbol{\Theta}$ 的学习往往需要与有噪图像完全配对的清晰图像。然而,由于高速和低速摄影仪成像参数不同,有噪和清晰图像在亮度和成像视野方面并不完全一致,直接进行端到端训练会破坏原图像的位置和亮度信息。因此,本文将非理想配对的清晰图像与有噪图像间的退化关系建模为:

$$\hat{\mathbf{y}} = \varphi(\mathbf{x}(p + f(p))) + \boldsymbol{\eta} \quad (3)$$

式中: $\hat{\mathbf{y}}$ 表示与清晰图像 $\mathbf{x}$ 非理想配对的有噪图像,也即

高速摄影仪的观测图像。 $\varphi(\cdot)$ 表示亮度一致化函数, $p$ 表示像素位置, $f(p)$ 表示在像素 $p$ 处有噪图像与清晰图像的位置偏移。基于上述退化模型,本文建立非理想配对下的去噪模型:

$$\hat{\mathbf{x}} = W(\varphi(\hat{\mathbf{y}} - D(\hat{\mathbf{y}}, \boldsymbol{\Theta})), f(p)) \quad (4)$$

式中: $W(\cdot)$ 是位置对齐函数<sup>[19]</sup>。训练后,实际去噪图像为 $\hat{\mathbf{x}} = \mathbf{x} - D(\mathbf{x}, \boldsymbol{\Theta})$ , $\hat{\mathbf{x}}$ 是与摄影仪观测的有噪图像 $\hat{\mathbf{y}}$ 配对的去噪图像。亮度一致化和位置对齐在训练过程中自动完成,在推断过程无需执行,因而推断过程不引入额外的计算量和参数。

提出的去噪模型训练框架如图2所示,该框架包括去噪网络模型、亮度因子估计器和光流估计模型。亮度因子估计器估计有噪图像与清晰图像的亮度差异,从而构建亮度一致函数 $\varphi(\cdot)$ ;光流估计模型计算位置偏移 $f(p)$ ,从而实现图像对齐。具体而言,有噪图像首先通过去噪模型 $D(\cdot)$ 得到去噪图像,同时,有噪图像和清晰图像输入亮度因子估计器计算亮度差异;然后,利用光流估计网络计算有噪图像与清晰图像间的光流,为保证光流计算过程中有噪图像与清晰图像亮度尽可能一致,提出的框架也对有噪图像进行亮度一致操作;最后,基于估计的光流和亮度因子,可实现预测图像与清晰图像的亮度一致和图像对齐,进而实现监督训练。值得注意的是,提出的方法训练过程虽然会有复杂的数据处理过程,但这些过程并不需要在测试阶段进行。去噪模型训练完成后,可直接应用于高速摄影仪的图像处理,将高速摄影仪的有噪图像输入去噪模型,即可获得清晰图像。

### 2) 去噪模型结构

去噪模型结构如图2所示,该模型采用编解码网络,其下采样通过步长为2的3×3卷积实现,上采样通过双线性插值完成,每个尺度都包含一个残差模块,该残差模块由步长为1的1×1和3×3卷积层、归一化层<sup>[20]</sup>以及激活层<sup>[21]</sup>构成。此外,为保留细节特征,编码过程残差模块提取的特征将通过跳跃连接与同尺度的解码过程残差模块融合。去噪模型通过残差学习实现噪声去除,即输入图像与预测的残差噪声相加得到最终的去噪图像。

### 3) 亮度一致化

伽玛变换是摄影仪调整图像亮度的重要方法,本文通过伽玛变换函数来模拟亮度一致化函数,即 $\varphi(\hat{\mathbf{y}}) = \hat{\mathbf{y}}^\gamma$ , $\gamma$ 即为亮度因子。变换后图像的总亮度应与清晰图像的总亮度尽可能一致,即:

$$\sum_{h=1}^H \sum_{w=1}^W \hat{\mathbf{y}}(h, w)^\gamma = \sum_{h=1}^H \sum_{w=1}^W \mathbf{x}(h, w) \quad (5)$$

式中: $H$ 和 $W$ 分别表示采集图像的高和宽, $h$ 和 $w$ 是像素位置索引,通过数值分析方法可求得 $\gamma$ 的近似值。

在实际训练过程中, $\gamma$ 可预先求解,在模型训练时为常数。上述亮度一致化操作可使采集的有噪图像和清晰

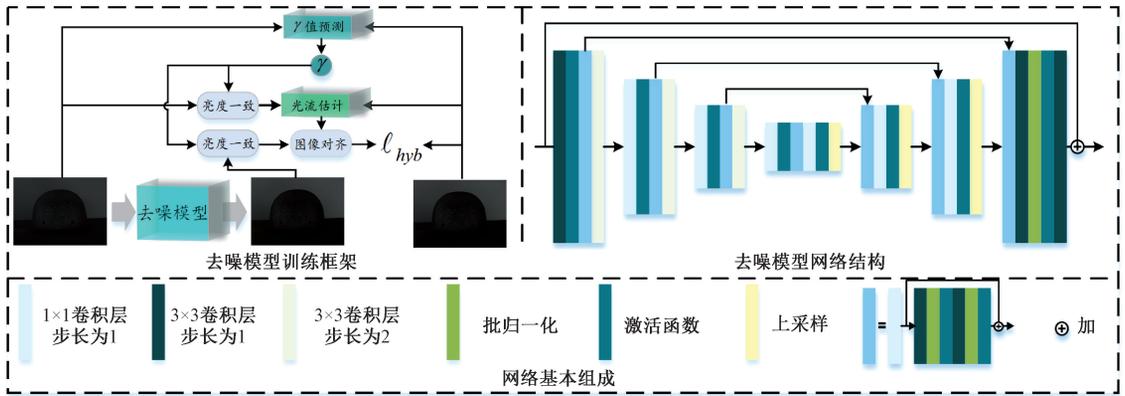


图2 非理想配对图像训练框架

Fig. 2 The training framework for non-ideal paired images

图像整体亮度尽可能一致,有效促进后续的图像对齐和模型训练过程。

#### 4) 图像对齐

本文基于光流估计方法实现图像对齐,给定有噪图像 \$\hat{\mathbf{y}}\$ 和非理想配对的清晰图像 \$\mathbf{x}\$,首先利用光流估计模型预测二者之间的位置偏移 \$f\$:

$$f = \psi(\varphi(\hat{\mathbf{y}}), \mathbf{x}) \quad (6)$$

式中: \$\psi(\cdot)\$ 表示光流估计模型,本文利用预训练的空间金字塔网络 (spatial pyramid network, SPyNet)<sup>[22]</sup> 作为 \$\psi(\cdot)\$。获得位置偏移 \$f\$ 后,即可将去噪图像 \$\hat{\mathbf{x}}\$ 向理想清晰图像 \$\mathbf{x}\$ 对齐,得到 \$\tilde{\mathbf{x}}\$:

$$\tilde{\mathbf{x}}(p) = W(\hat{\mathbf{x}}(p), f(p)) \quad (7)$$

对齐操作可减小成像视野不一致导致的位置偏移。

#### 5) 模型训练

本文利用混合损失 \$\ell\_{hyb}\$ 训练去噪网络,该损失包含均方差损失 \$\ell\_{mse}\$ 和全变分损失 \$\ell\_{tv}\$,定义为:

$$\ell_{hyb} = \ell_{mse} + \lambda \ell_{tv} \quad (8)$$

式中: \$\lambda\$ 是平衡权重。均方差损失可提升模型的保真度,全变分损失可有效抑制带来明显梯度变化的噪声。具体而言,均方差损失可以表达为:

$$\ell_{mse} = \frac{1}{2BHW} \sum_{b=1}^B \sum_{h=1}^H \sum_{w=1}^W \|\tilde{\mathbf{x}}_b(h, w) \times \mathbf{O}_b(h, w) - \mathbf{x}_b(h, w) \times \mathbf{O}_b(h, w)\|^2 \quad (9)$$

全变分损失可以表达为:

$$\ell_{tv} = \frac{1}{BHW} \sum_{b,h,w} ((\tilde{\mathbf{x}}_b(h+1, w) - \tilde{\mathbf{x}}_b(h, w))^2 + (\tilde{\mathbf{x}}_b(h, w+1) - \tilde{\mathbf{x}}_b(h, w))^2) \quad (10)$$

式中: \$B\$ 表示每次训练输入的图像数量,即批大小, \$b\$ 是图像在每个批次中的索引, \$\mathbf{O}\$ 表示一致区域掩膜<sup>[19]</sup>,可通过光流获得。通过上述损失函数,可实现模型训练,从而获得与观测有噪图像 \$\hat{\mathbf{y}}\$ 配对的去噪图像 \$\hat{\mathbf{x}}\$。本质上,本文的对齐是高速摄影仪和低速摄影仪在成像视场上的对

齐,由于两台相机难以调整至视场完全一致,因此,本文通过图像对齐的方式取得两个相机公共的视场。

#### 1.3 去噪模型量化

基于卷积神经网络的去噪模型参数量较大,且对内存需求较大,不易于硬件部署。受文献[23]启发,本文将模型量化技术引入图像去噪模型。由于训练后直接量化会引入较大的量化误差,导致去噪模型精度大幅下降,本文通过量化感知训练微调模型从而减少量化过程的精度损失,在减少权重和激活值位宽的同时使模型保持良好的去噪性能。

针对模型的激活值,本文采用非对称量化方法:

$$\tilde{\mathbf{v}} = R\left(F\left(\frac{\mathbf{v} - \beta}{s}, m, n\right)\right) \quad (11)$$

式中: \$\mathbf{v}\$ 表示原去噪模型的激活值, \$\tilde{\mathbf{v}}\$ 表示量化后去噪模型的激活值, \$R(\cdot)\$ 表示四舍五入函数, \$\beta\$ 表示量化的偏置参数, \$s\$ 表示量化的尺度因子, \$F(\cdot)\$ 表示截断函数,可将输入值截断在最小值 \$n\$ 和最大值 \$n\$ 之间。本文中, \$m = -2^{b-1}\$, \$n = 2^{b-1} - 1\$, \$b\$ 是量化的位宽。量化感知训练需将量化后的模型反量化为浮点数减少量化误差,即:

$$\mathbf{v}' = \tilde{\mathbf{v}} \times s + \beta \quad (12)$$

式中: \$\mathbf{v}'\$ 表示反量化后的激活值。尺度因子 \$s\$ 和偏置参数 \$\beta\$ 均是可学习参数,尺度因子 \$s\$ 的梯度更新可以表达为:

$$\frac{\partial \mathbf{v}'}{\partial s} = \frac{\partial \tilde{\mathbf{v}}}{\partial s} + \tilde{\mathbf{v}} \approx \begin{cases} -\frac{\mathbf{v} - \beta}{s} + R\left(\frac{\mathbf{v} - \beta}{s}\right), & m < \frac{\mathbf{v} - \beta}{s} < n \\ m \text{ or } n, & \text{其他} \end{cases} \quad (13)$$

偏置参数的梯度更新可以表达为:

$$\frac{\partial \mathbf{v}'}{\partial \beta} = \frac{\partial \tilde{\mathbf{v}}}{\partial \beta} + 1 \approx \begin{cases} 0, & m < \frac{\mathbf{v} - \beta}{s} < n \\ 1, & \text{其他} \end{cases} \quad (14)$$

在式(13)和(14)中,  $\frac{\partial \tilde{\nu}}{\partial s}$  和  $\frac{\partial \tilde{\nu}}{\partial \beta}$  通过直通估计器 (straight through estimator, STE)<sup>[23]</sup> 求解近似值。

针对模型中的权重参数,由于权重分布近似关于零点对称分布<sup>[24]</sup>,本文采用对称量化方法:

$$\tilde{\mathbf{w}} = R\left(F\left(\frac{\mathbf{w}}{s_2}, m, n\right)\right) \quad (15)$$

$$\mathbf{w}' = \tilde{\mathbf{w}} \times s_2 \quad (16)$$

式中:  $s_2$  是用于权重量化的尺度因子,  $\mathbf{w}$  表示原去噪模型的权重,  $\tilde{\mathbf{w}}$  是量化后去噪模型的权重,  $\mathbf{w}'$  是反量化后的权重。尺度因子的梯度更新可以表达为:

$$\frac{\partial \mathbf{w}'}{\partial s_2} = \frac{\partial \tilde{\mathbf{w}}}{\partial s_2} s_2 + \tilde{\mathbf{w}} \approx \begin{cases} -\frac{\mathbf{w}}{s_2} + R\left(\frac{\mathbf{w}}{s_2}\right), & m < \frac{\mathbf{w}}{s_2} < n \\ m \text{ or } n, & \text{其他} \end{cases} \quad (17)$$

通过上述方法,可实现去噪模型的量化压缩,大幅降低模型大小和资源需求。

## 2 实验结果与分析

### 2.1 平台设置及实验细节

#### 1) 平台设置

成像系统由自主研发的高速和低速摄影仪构成,高速摄影仪帧率为 13 000 FPS,低速摄影仪帧率为 100 FPS。算法运行平台环境为 Ubuntu 20.04, CUDA 11.4, cuDNN 8.0.4, 处理器为 Intel(R) Xeon(R) Gold 6258R。去噪模型基于 PyTorch 框架实现,在单张 Titan RTX3090 显卡上实现模型训练和测试。

#### 2) 实验设置

从采集的图像数据中随机选择 50 对图片作为测试集,其余 250 对图片作为训练集。去噪模型共训练 500 周期,图像随机剪裁为 96×96 的图像块(patch size),批大小(batch size)设置为 10,初始学习率设置为 0.0004,每 100 个周期学习率下降 2 倍。模型量化中,量化位宽设置为 8。

#### 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(structural similarity, SSIM)作为质量评价指标。峰值信噪比是信号最大可能功率和信号中噪声功率的比值,图像信号中该指标表达为:

$$P(\tilde{\mathbf{x}}, \mathbf{x}) = 10 \times \lg\left(\frac{\text{Max}_x^2}{\text{MSE}(\tilde{\mathbf{x}}, \mathbf{x})}\right) \quad (18)$$

$$\text{MSE}(\tilde{\mathbf{x}}, \mathbf{x}) = \frac{1}{H \times W} \sum_{h=1}^H \sum_{w=1}^W (\tilde{\mathbf{x}}(h, w) - \mathbf{x}(h, w))^2 \quad (19)$$

式中:  $P(\tilde{\mathbf{x}}, \mathbf{x})$  计算信号  $\tilde{\mathbf{x}}$  和  $\mathbf{x}$  的 PSNR,  $\text{Max}_x$  为信号中可能的最大值,  $\text{MSE}$  表示均方误差。结构相似性可衡量两幅图像的结构相似度,计算公式为:

$$Q(\tilde{\mathbf{x}}, \mathbf{x}) = \frac{2\mu_{\tilde{\mathbf{x}}}\mu_{\mathbf{x}} + c_1}{\mu_{\tilde{\mathbf{x}}}^2 + \mu_{\mathbf{x}}^2 + c_1} \times \frac{2\sigma_{\tilde{\mathbf{x}}\mathbf{x}} + c_2}{\sigma_{\tilde{\mathbf{x}}}^2 + \sigma_{\mathbf{x}}^2 + c_2} \quad (20)$$

式中:  $Q(\tilde{\mathbf{x}}, \mathbf{x})$  计算信号  $\tilde{\mathbf{x}}$  和  $\mathbf{x}$  的 SSIM,  $\mu_{\tilde{\mathbf{x}}}$ ,  $\mu_{\mathbf{x}}$ ,  $\sigma_{\tilde{\mathbf{x}}}$ ,  $\sigma_{\mathbf{x}}$  分别表示图像  $\tilde{\mathbf{x}}$  和  $\mathbf{x}$  的均值和方差,  $\sigma_{\tilde{\mathbf{x}}\mathbf{x}}$  表示图像  $\tilde{\mathbf{x}}$  和  $\mathbf{x}$  的协方差。提出方法得到的去噪图像计算 PSNR 和 SSIM 需经过亮度一致化和图像对齐。

### 2.2 对比结果

本节对提出方法进行定量和定性分析,“提出方法(-)”表示提出的去噪模型直接训练而不使用亮度一致和图像对齐操作。本文将提出的方法与基于空间滤波的方法(高斯滤波)、基于先验的方法(BM3D<sup>[4]</sup>)和基于深度学习的去噪方法(DnCNN<sup>[13]</sup>、FFDNet<sup>[14]</sup>和DONet<sup>[16]</sup>)进行对比。高斯滤波和BM3D方法通过手动调节方差选择最佳性能参数,高斯滤波窗口大小设置为 10×10;此外,基于深度学习的方法均在建立的数据集上重新训练。定量对比结果如表1所示,可以看出,传统方法去噪性能较差,无法与基于深度学习的方法相比;DnCNN、FFDNet和DONet直接利用非配对图像进行监督学习,由于图像不是逐像素匹配,去噪模型在学习去噪的同时,也会学习位置偏移和亮度偏差,去噪性能有限;相比之下,提出的方法能够显著提升图像去噪结果,将PSNR提升至少 1.96 dB,SSIM提升至少 1.95%,该结果证明了提出方法的优越性。此外,与不使用亮度一致化和图像对齐操作得到的去噪模型相比,PSNR提升了至少 1.5 dB,SSIM提升了 1.46%,该结果证明了亮度一致化和图像对齐操作在非理想配对图像训练方法中的有效性。图3和4展示了暗光环境和正常光照环境拍摄图像去噪结果的视觉对比,高斯滤波得到的去噪图像非常模糊,去除噪声的同时,损失大量细节信息;BM3D方法虽然能有效去除图像噪声,但也会使图像变模糊且有残余噪声;DnCNN和DONet得到的去噪图像也会有残余噪声;相比之下,提出的方法得到的去噪图像更够获得更好的视觉效果。

### 2.3 非理想配对图像训练方法适用性分析

为进一步验证提出的非理想配对图像训练方法的适用性,本文将提出的方法应用到其他基于深度学习的图像去噪模型中。实验结果如表2所示,“FFDNet+”、“DnCNN+”和“DONet+”分别表示将提出的非理想配对图像训练方法应用到FFDNet、DnCNN和DONet上训练得到的去噪模型。结果表明,利用提出的非理想配对图像训练方法比直接用深度学习模型学习非理想配对图像得到的去噪效果更好。图5展示了直接使用非理想配对图像训练的深度去噪模型和基于提出的训练方法训练的

表1 不同去噪方法结果对比

Table 1 Comparison results of different denoising methods

评价指标	高斯滤波	BM3D	FFDNet	DnCNN	DONet	提出方法(-)	提出方法
PSNR	28.62	30.01	33.38	33.44	33.72	34.18	35.68
SSIM	0.845 2	0.895 1	0.927 5	0.948 6	0.949 8	0.954 3	0.968 3
运行时间/ms	5.9	7 720	18.8	85.3	32.7	57.6	57.6

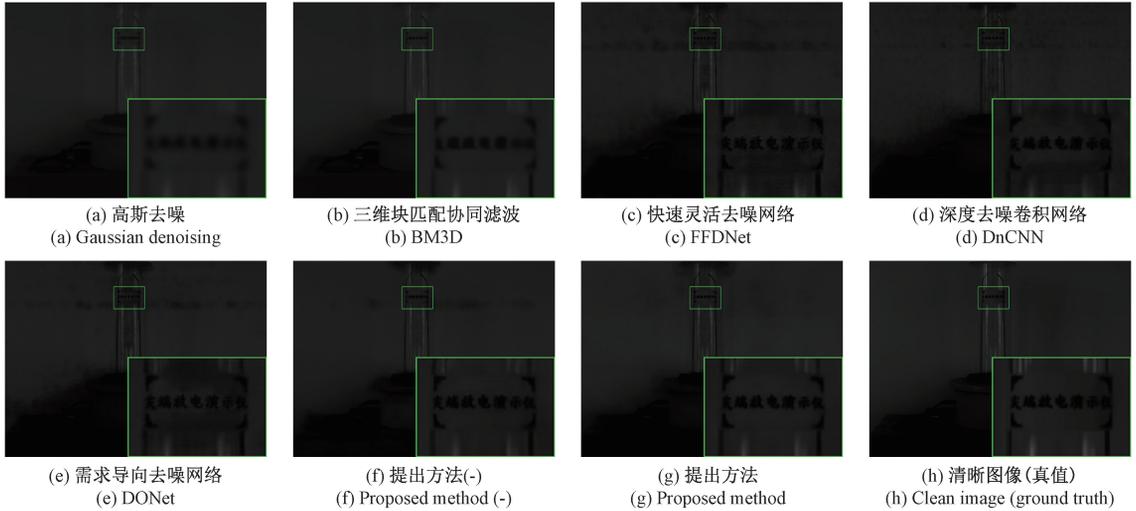


图3 不同方法在暗光环境下的去噪结果对比

Fig. 3 Comparison results achieved by different methods under the low-light environment

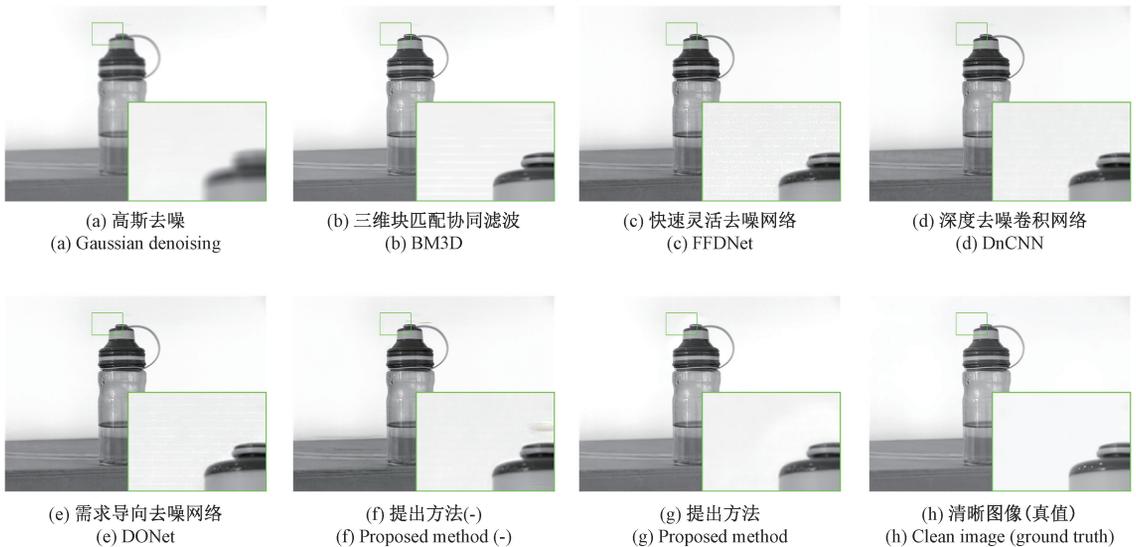


图4 不同去噪方法在正常光照环境下的去噪效果对比

Fig. 4 Comparison results achieved by different methods under the normal-light environment

表2 提出的框架应用在其他去噪模型上得到的对比结果

Table 2 Comparison results achieved by applying the proposed framework to other denoising models

评价指标	FFDNet	FFDNet+	DnCNN	DnCNN+	DONet	DONet+
PSNR	33.38	34.23	33.44	34.78	33.72	35.03
SSIM	0.927 5	0.933 2	0.948 6	0.951 5	0.949 8	0.958 1

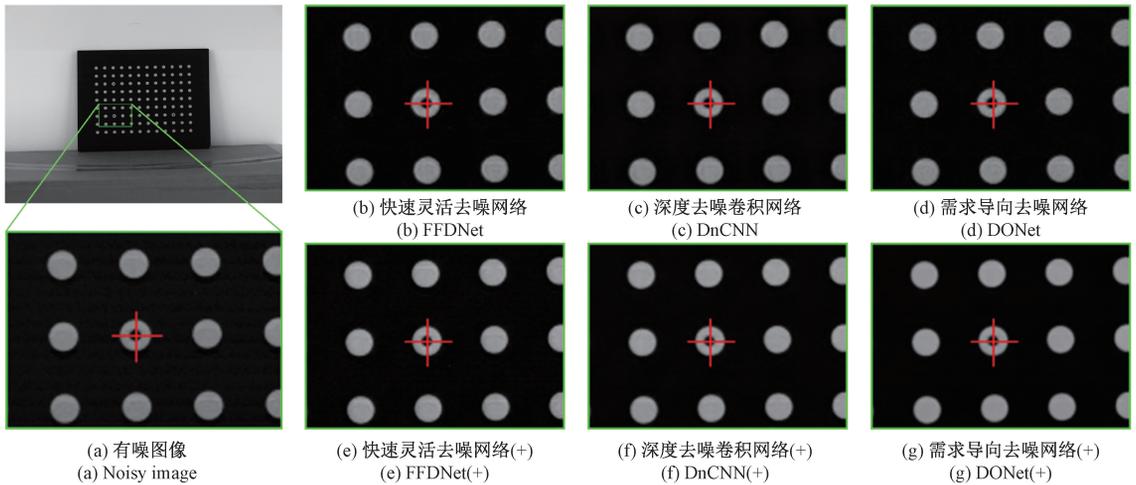


图 5 提出的框架应用在其他去噪模型上得到的可视化对比结果

Fig. 5 Visual comparison results achieved by applying the proposed framework to other denoising models

深度去噪模型的定性对比结果,“十字”是在所有去噪图像的同位置标注的参考点,可以看出直接使用非理想配对图像训练去噪模型会使模型学习位置偏移,导致去噪模型破坏原图像的位置信息。相比之下,使用提出的非理想配对图像训练方法可帮助去噪图像更好地保留原始位置信息。

### 2.4 模型量化性能评估

去噪模型量化前后定量对比结果如表 3 所示,引入的模型量化方法将模型参数和激活值由 32 位浮点数量化为 8 位定点数。量化后的模型去噪性能与原去噪模型性能相近,其峰值信噪比降幅为 0.15 dB,结构相似性降幅更是只有 0.000 7。与此同时,量化操作能将模型大小降低 4 倍,内存需求降低 45.62%,运行时间降低 37.5%,该结果表明量化后的去噪模型可以用更少的资源获得与原模型相近的去噪性能。

表 3 量化前后去噪模型的性能对比

Table 3 Comparison results of the unquantized and quantized denoising models

是否量化	PSNR	SSIM	位宽	模型尺寸/M	内存需求/M	运行时间/ms
不量化	35.68	0.968 3	32	3.3	4 193	57.6
量化	35.53	0.967 6	8	0.8	2 280	36.0

### 2.5 量化感知训练与训练后直接量化的模型性能对比

去噪模型训练后直接量化会引入较大的量化误差,导致模型去噪性能大幅下降。量化感知训练可在训练过程中对模型进行微调,减少量化过程的性能下降。训练后直接量化与量化感知训练得到的去噪模型对比结果如

表 4 所示,相比于训练后直接量化,量化感知训练可以更好的保证去噪模型性能,能将 PSNR 提升 5.34 dB,SSIM 提升 18.32%。该结果证明了在去噪任务中,量化感知训练对去噪模型量化的重要性。

表 4 直接量化与量化感知训练得到的模型去噪结果对比  
Table 4 Comparison results of the denoising models obtained by direct quantization and quantization perception training

量化方法	PSNR	SSIM
训练后直接量化	30.19	0.817 8
量化感知训练	35.53	0.967 6

### 2.6 权重参数 $\lambda$ 敏感性分析

权重参数对模型的去噪结果有很大影响,为此,本节对权重参数  $\lambda$  进行敏感性分析。分析结果如图 6 所示,当  $\lambda$  过小时,模型主要由均方差损失主导,由于缺少梯度约束,噪声不能被完全去除;随  $\lambda$  值增大,全变分损失可有效促进模型优化,去噪性能得到提升;然而,当  $\lambda$  过大时,均方差损失主导的保真度受到影响,致使去噪模型性能下降。实验分析发现, $\lambda$  在 0.02 时可以取得较好的去噪性能,因此,本工作中, $\lambda$  值设置为 0.02。

### 2.7 实际场景应用分析

本节进一步将提出方法在电弧发生场景中进行应用测试。电弧发生过程非常短暂,为观察电弧形状和变化过程,需要用高速摄影仪记录其发生过程。实验平台如图 7 所示,将自研高速摄影仪放置在三脚架上,拍摄放电器的电弧发生过程,拍摄视场约 15 cm,发生时间约 500  $\mu$ s,拍摄帧率为 71 000 FPS。实验拍摄图像如图 8 所示,第 1 行是未使用提出方法进行去噪的实验结果,第 2

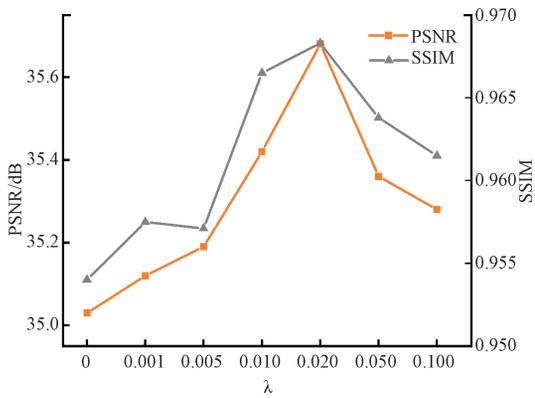


图6 权重参数  $\lambda$  敏感性分析

Fig. 6 Sensitivity analysis of the weight parameter  $\lambda$

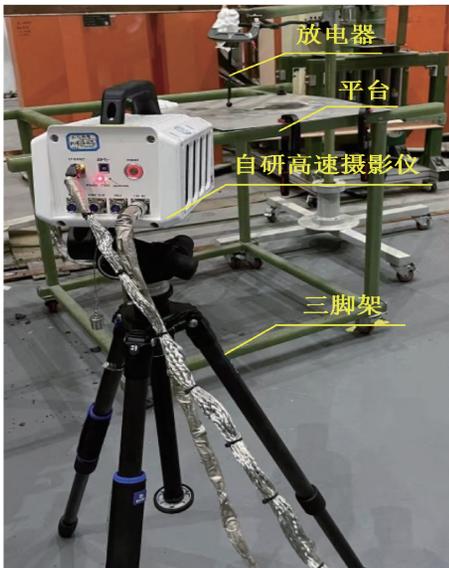


图7 实验平台

Fig. 7 Experimental platform

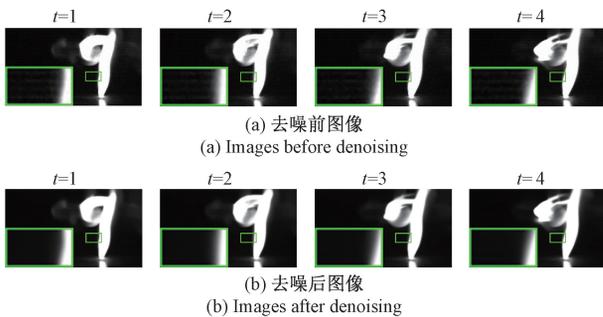


图8 提出方法在电弧放电实验中的应用分析

Fig. 8 Analysis of the proposed denoising method applied to electro-discharge experiments

行是使用提出方法去噪后的结果。可以看出去噪后电弧周围的条纹噪声及随机噪声得到有效抑制,可以更清晰的记录电弧发生过程的形状,该结果证明了提出方法在实际应用场景的有效性。

### 3 结 论

本文针对高速摄影仪在超高帧率下易产生噪声,且难以获取与有噪图像完全配对的清晰图像的问题,建立了非理想配对的图像去噪数据集,提出了基于非理想配对图像的卷积神经网络去噪模型训练方法,通过亮度一致化和图像对齐方法,实现了非理想配对图像的监督学习,有效去除了成像噪声。此外,通过引入模型量化方法将去噪模型参数和激活值由32位浮点数量化为8位定点数,大幅降低了模型大小、内存需求和运行时间。提出的方法在自主研制的数字式高速摄影仪上得到成功应用,后续将进一步降低模型运行的资源需求,实现模型在FPGA中的部署。

### 参考文献

[ 1 ] 张堃,王震,张培建,等. 面向高速视觉检测的精确抓拍安全策略研究[J]. 仪器仪表学报, 2018, 39(2): 232-240.  
ZHANG K, WANG ZH, ZHANG P J, et al. Research on accurate capture security strategy for high-speed visual inspection[J]. Chinese Journal of Scientific Instrument, 2018, 39(2): 232-240.

[ 2 ] LIU Z, HE Q B, CHEN S Q, et al. Time-varying motion filtering for vision-based nonstationary vibration measurement[J]. IEEE Transactions on Instrumentation and Measurement, 2020, 69(6): 3907-3916.

[ 3 ] 王耀东,余祖俊,石井抱. 高速立体视觉与机械运动异常动作三维检测[J]. 仪器仪表学报, 2013, 34(10): 2161-2169.  
WANG Y D, YU ZH J, ISHII I. High-speed stereo vision based three dimension abnormal operation detection of mechanical motion[J]. Chinese Journal of Scientific Instrument, 2013, 34(10): 2161-2169.

[ 4 ] DABOV K, FOI A, KATKOVNIK V, et al. Image denoising by sparse 3-D transform-domain collaborative filtering[J]. IEEE Transactions on Image Processing, 2007, 16(8): 2080-2095.

[ 5 ] GU S, ZHANG L, ZUO W, et al. Weighted nuclear norm minimization with application to image denoising [C]. Proceedings of the IEEE Conference on Computer Vision

- and Pattern Recognition. Piscataway, NJ: IEEE, 2014, 2862-2869.
- [ 6 ] ELAD M, AHARON M. Image denoising via sparse and redundant representations over learned dictionaries [ J ]. IEEE Transactions on Image Processing, 2006, 15( 12 ): 3736-3745.
- [ 7 ] MAIRAL J, BACH F, PONCE J, et al. Non-local sparse models for image restoration [ C ]. Proceedings of the IEEE Conference on International Conference on Computer Vision, IEEE, 2009: 2272-2279.
- [ 8 ] DONG W, ZHANG L, SHI G, et al. Nonlocally centralized sparse representation for image restoration [ J ]. IEEE Transactions on Image Processing, 2013, 22 ( 4 ): 1620-1630.
- [ 9 ] 李佳星, 赵勇先, 王京华. 基于深度学习的单幅图像超分辨率重建算法综述 [ J ]. 自动化学报, 2021, 47( 10 ): 2341-2363.
- LI J X, ZHAO Y X, WANG J H. A review of single image super-resolution reconstruction algorithms based on deep learning [ J ]. Acta Automatica Sinica, 2021, 47( 10 ): 2341-2363.
- [ 10 ] 杨珂, 方诚, 段黎明. 基于深度学习模型融合的铸件缺陷自动检测 [ J ]. 仪器仪表学报, 2021, 42( 11 ): 150-159.
- YANG K, FANG CH, DUAN L M. Automatic detection of casting defects based on deep learning model fusion [ J ]. Chinese Journal of Scientific Instrument, 2021, 42( 11 ): 150-159.
- [ 11 ] 张加劲. 基于注意力机制和 CNN-BiLSTM 模型的航空发动机剩余寿命预测 [ J ]. 电子测量与仪器学报, 2022, 36( 8 ): 231-237.
- ZHANG J J. Remaining useful life estimation of aeroengine based on CNN-BiLSTM and attention mechanism [ J ]. Journal of Electronic Measurement and Instrumentation, 2022, 36( 8 ): 231-237.
- [ 12 ] 盖杉, 鲍中运. 基于深度学习的高噪声图像去噪算法 [ J ]. 自动化学报, 2020, 46( 12 ): 2672-2680.
- GAI SH, BAO ZH Y. High noise image denoising algorithm based on deep learning [ J ]. Acta Automatica Sinica, 2020, 46( 12 ): 2672-2680.
- [ 13 ] ZHANG K, ZUO W, CHEN Y, et al. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising [ J ]. IEEE Transactions on Image Processing, 2017, 26( 7 ): 3142-3155.
- [ 14 ] ZHANG K, ZUO W, ZHANG L. FFDNet: Toward a fast and flexible solution for CNN-based image denoising [ J ]. IEEE Transactions on Image Processing, 2018, 27( 9 ): 4608-4622.
- [ 15 ] PARK B, YU S, JEONG J. Densely connected hierarchical network for image denoising [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019.
- [ 16 ] CHEN H, JIN Y, DUAN M, et al. DOF: A demand-oriented framework for image denoising [ J ]. IEEE Transactions on Industrial Informatics, 2021, 17( 8 ): 5369-5379.
- [ 17 ] LEHTINEN J, MUNKBERG J, HASSELGREN J, et al. Noise2noise: Learning image restoration without clean data [ C ]. International Conference on Machine Learning, Stockholmsmässan, Cambridge, MA: PMLR, 2018.
- [ 18 ] KRULL A, BUCHHOLZ T O, JUG F. Noise2void-learning denoising from single noisy images [ C ]. Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition. Piscataway, NJ: IEEE, 2019, 2129-2137.
- [ 19 ] GUPTA A, JOHNSON J, ALAHI A, et al. Characterizing and improving stability in neural style transfer [ C ]. Proceedings of the IEEE International Conference on Computer Vision. Piscataway, NJ: IEEE, 2017, 4067-4076.
- [ 20 ] IOFFE S, SZEGEDY C. Batch normalization: Accelerating deep network training by reducing internal covariate shift [ C ]. International Conference on Machine Learning. Cambridge, MA: PMLR, 2015.
- [ 21 ] MAAS A L, HANNUN A Y, NG A Y. Rectifier nonlinearities improve neural network acoustic models [ C ]. International Conference on Machine Learning, Cambridge, MA: PMLR, 2013.
- [ 22 ] RANJAN A, BLACK M J. Optical flow estimation using a spatial pyramid network [ C ]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, NJ: IEEE, 2017, 4161-4170.
- [ 23 ] BHALGAT Y, LEE J, NAGEL M, et al. Lsq+: Improving low-bit quantization through learnable offsets and better initialization [ C ]. Proceedings of the IEEE/

CVF Conference on Computer Vision and Pattern Recognition, Piscataway, NJ: IEEE 2020, 696-697.

- [24] FARAONE J, FRASER N, BLOTT M, et al. Sq: Learning symmetric quantization for efficient deep neural networks[C]. Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Piscataway, NJ: IEEE, 2018, 4300-4309.

## 作者简介



**陈怀安**, 2017 年于安徽大学获得学士学位, 2022 年于中国科学技术大学获得博士学位, 现为中国科学技术大学博士后, 主要研究方向为高速成像与图像处理。

E-mail: anchen@mail.ustc.edu.cn.

**Chen Huaian** received his B. Sc. degree from Anhui University in 2017 and received his Ph. D. degree from University of Science and Technology of China in 2022. He is currently a postdoctoral research fellow at University of Science and Technology of China. His research interests include high-speed imaging and image processing.



**卢小银**, 2010 和 2013 年于中国科学技术大学分别获得学士和硕士学位, 现任合肥中科君达视界技术股份有限公司研发总监, 主要研究方向为高速成像与图像处理。

E-mail: lxy@gaosuxiangji.com

**Lu Xiaoyin** received his B. Sc. degree and M. Sc. degree both from University of Science and Technology of China in 2010 and 2013, respectively. He is currently the research & development director at Hefei Agile Device Co., Ltd. His research interests include high-speed imaging and image processing.



**阚艳**, 2001 年于北京航空航天大学获得学士学位, 2018 年于南京航空航天大学获得硕士学位, 现为上海交通大学博士研究生, 同时担任国营芜湖机械厂厂长, 主要研究方向为航天科学与工程、材料科学、自动化技术。

E-mail: kanyan5720@gmail.com

**Kan Yan** received his B. Sc. degree from Beihang University in 2001 and M. Sc. degree from Nanjing University of Aeronautics and Astronautics in 2018. He is currently pursuing his Ph. D. degree at Shanghai Jiao Tong University. He is also the director of the Wuhu State-owned Factory of Machining. His current research interests include aerospace science and engineering, materials science, and automation technology.



**金一**(通信作者), 2007 年于江南大学获得学士学位, 2013 年于中国科学技术大学获得博士学位, 现为中国科学技术大学工程科学学院副教授, 主要研究方向包括智能检测、图像处理、人工智能。

E-mail: jinyi08@ustc.edu.cn

**Jin Yi** (Corresponding author) received his B. Sc. degree from Jiangnan University in 2007 and received his Ph. D. degree from University of Science and Technology of China in 2013. He is currently an associate professor with the School of Engineering Science at University of Science and Technology of China. His current research interests include intellectual detection, image processing, and artificial intelligence.