



内容列表可在科学直接获得

模式识别

期刊主页: www.elsevier.com/locate/patcog

水下场景之前启发了深水下图像和视频增强

李崇义^{1,*}, 赛义德安瓦尔b,^{c,d}, Porikli

香港城市大学(CityU)计算机科学系,香港b数据61, CSIRO, ACT 2601, 澳大利亚

c澳大利亚国立大学,堪培拉第2600号法令,澳大利亚

d澳大利亚国立大学工程研究学院,堪培拉,ACT 0200, 澳大利亚



article info

章节病史:

收到2019年4月17日

2019年8月23日修订

2019年9月4日接受

2019年9月5日在线发售

关键词:

水下图像和视频增强

和恢复

水下图像合成是模式识别

深度学习

abstract

在水下场景中,波长依赖的光吸收和散射降低了识别年龄和视频的可见性。水下图像和视频的退化影响了水下图像的模式识别、视觉理解和关键特征提取的准确性。本文提出了一种基于水下场景先验的水下图像增强通信神经网络(CNN)模型,称为UWCNN。所提出的UWCNN模型不是估计水下成像模型的参数,而是直接重建清晰的潜水面下图像,这得益于水下场景先验,可用于合成水下图像训练数据。此外,基于光权网络结构和有效的训练数据,我们的UWCNN模型可以很容易地扩展到水下视频中进行逐帧增强。具体来说,将水下成像物理模型与水下场景的光学道具相结合,我们首先合成了包含水类型和退化水平的水下成像数据数据集。然后,设计轻权CNN模型,增强每种水下场景类型,训练by这些响应的训练数据。最后,将该UWCNN模型直接扩展到水下视频增强技术中。现实世界和合成水下图像和视频实验表明,我们的方法可以很好地推广到不同的水下场景。

©2019爱思唯尔有限公司。保留所有权利。

1.介绍

获取清晰的水下图像和视频对于水下场景的感知和理解非常重要,因为自主和远程操作的水下飞行器被广泛用于探索、识别和与海洋环境的交互。然而,水下原始图像和视频很少满足视觉质量的期望,进一步对模式识别、目标检测、关键特征提取等性能提出了挑战。这是因为大多数深度网络都是由高质量图像训练的,或者算法认为输入的图像是清晰的图像。自然地,水下图像由于光吸收和散射的不利影响而退化,包括微浮游植物、溶解有机物和非藻颗粒。此外,当光在水下场景中传播时,它相对于波具有选择性衰减的特性-

光长[1]。[图1](#)给出了光与光波长的衰减图。

这些吸收和散射问题阻碍了水下场景的理解和识别的能力,如水生机器人检查和海洋环境监测。此外,传统的图像增强方法是[2,3]当它们被用来处理水下图像和视频时,你可以炫耀一下。此外,由于缺乏足够和有效的训练数据,基于深度学习的水下图像和视频增强方法的性能不匹配基于深度学习的解决方案的成功,如非分类[4],分析性[5] segmentation[6], 超分辨率[7], recognition[8], 等需要开发水下图像合成和增强方法,以提高水下视觉质量,提高高级视觉任务的性能。

近年来,越来越多的基于深度学习的方法采用了[10,11]已经提出。与传统的基于非学习的方法相比,深度模型具有一定的优势:(i)深度学习提供了较强的扭曲建模能力,促进了区分先验学习;(ii)利用并行方法可以有效地实现深度模型的推理

*通讯作者。

E-mail address: [\(C. 列支敦士登](mailto:lichongyi@tju.edu.cn)

1李崇义和赛义德·安瓦尔对这项工作的贡献相同。

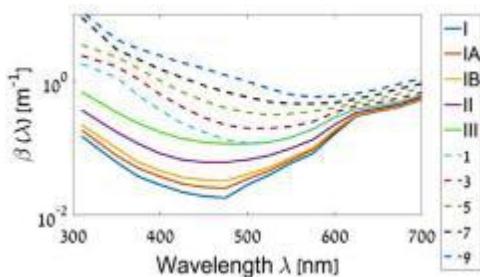


图1。[9]中Jerlov水型的波长依赖光衰减系数 β 。 实线表示开放的海水类型，虚线表示沿海水域类型。Jerlov水类型为开放海洋水域的I、IA、IB、II和III，沿海水域为I至9。i型是最清澈的，iii型是最浑浊的开阔海水。同样地，对于沿海水域，1型是最清澈的，9型是最浑浊的。

处理平台。受最近在模式识别和直接理解[12]方面的成功的启发，我们提出了一种新的水下图像合成算法，并为水下图像和视频增强提供了一个鲁棒的、数据驱动的解决方案。该方法对不同类型的水具有良好的鲁棒性、准确性和灵活性。

贡献：我们利用一种新的水下场景CNN架构训练，设计了一种基于非线性水下图像形成模型的端到端解决方案。我们的模型稳健地保留了去分层的水下图像，并准确地重建了下面的颜色和外观。此外，由于该网络结构，该模型可以很容易地扩展到水下视频。综上所述，

我们提出了一种新的基于水下场景先验的水下图像合成算法，该算法能够模拟处理一组退化的水下图像。据我们所知，它是第一个水下图像合成算法，可以模拟不同的水下类型和退化水平。我们的图像合成可以作为网络训练和全参考图像质量评估的指南。

我们提出了一种新的CNN模型，通过联合优化多项损失，在保持原始结构和纹理的同时，重建水下透明帐篷图像。从光权网络设计和有效的训练数据来看，该模型可以对水下视频进行逐帧增强的扩展。

我们的方法可以很好地推广到合成的和真实世界的水下图像和具有不同颜色和可见性特征的视频。此外，当将有效的先验信息嵌入网络时，轻质网络结构也可以取得良好的效果，从而促进了模式识别、视觉理解等方面的相关设计。

2.相关工作

从不同的角度，现有的水下图像增强和恢复方法可以分为不同的方法。本文将这些方法分为水下图像增强法、水下图像恢复法和补充信息专用法三种广播方法。由于水下视频处理和修复工作很少，本节主要介绍图像处理方法。

2.1. 水下图像增强法

在这一研究中，L等人[13] 将水下图像增强问题作为图像去模糊步骤和颜色校正步骤来处理。Ancuti等人[14] 融合了水下的对比图像和输入的水下图像。在多尺度融合的过程中，使用四个权值来确定哪个像素更适合出现在最终的图像中。在[15]中提出了一种基于颜色校正和水下图像去模糊的混合方法，它利用图像颜色优先值来修正水下图像的颜色转换，并通过改进的图像去模糊算法来提高能见度。Li等人[16] 提出了一种基于弱超监督颜色转移的水下图像颜色校正方法，它可以学习在水下图像和空气图像之间的跨域映射函数。受通用对抗网络（GANs）的启发，Guo等人，[17] 提出了一种多尺度密集的模糊图像增强方法，通过引入多尺度、深度分割和残差学习策略来增强模糊图像的性能。Ancuti等人[18] 修改了他们之前的工作[14] 以减少过度增强和过度暴露的影响。最近，Li等人，[19] 提出了对配对的皮肤年龄和相应的参考图像进行训练的深度基线模型。这些参考年龄是通过不同的方法从增强的结果中主观地选择的。

2.2. 水下图像恢复方法

水下图像恢复方法通常将当前的挑战视为一个反问题，然后构造退化的物理模型，最后估计模型的参数。蒋和陈[20] 将图像去模糊算法与波长相关的补偿算法撕裂存储水下图像相结合，可以去除蓝色的水下图像和人工光的效果。一个红色通道的方法[21] 通过恢复与短波长相关的颜色来恢复水下丢失的图像。Drews等人[22] 提出了一种称为UDCP的水下暗通道航r，它修改了之前的暗通道[23]。有了UDCP，可以估计中传输；然而，当在水中场景中有白色物体或人工光时，UDCP并不总是成立。Li等人[24,25] 将水下图像去模糊算法与对比度增强算法相结合。Peng等人[26] 基于图像模糊度和光吸收的再深水图像，是基于四月的方法。Li等人[27] 提出了一种基于CNN的水下图像，基于合成深水图像的水下图像颜色校正模型。

2.3. 补充-信息专用的方法

补充信息特定的方法通常是利用从极化滤波器、立体图像、场景深度等[28]中获得的附加信息。

3.水下图像配方模型

我们遵循在[20]中提出的水下图像公式模型。该水下图像退化模型已广泛应用于传统的水下图像恢复方法中，可表示为：

$$U\lambda(x) = I\lambda(x) \cdot T\lambda(x) + B\lambda \cdot (1 - T\lambda(x)), \quad (1)$$

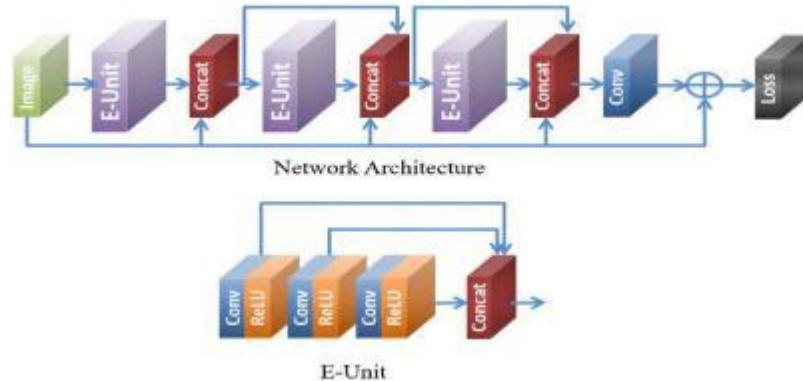


图2。在我们的UWCNN模型中，“Conv”表示卷积层，“Concat”表示堆叠的卷积层，“ReLU”表示修正的线性单元

其中 $U_\lambda(x)$ 是捕获的水下图像； $I_\lambda(x)$ 是 cl 耳潜在图像，也称为场景偏差，我们的目标是撕裂覆盖； B_λ 是均匀的全局背景光； λ 是红色、绿色和蓝膜的光的波长； x 是水下场景中的一个点（为了清晰起见，图像用粗体大写字母表示）。中能量比 $T_\lambda(x)$ 表示从水下场景中的 x 点反射后到达摄像机的场景辐射的百分比，从而导致颜色投射和对比度下降。换句话说， $T_\lambda(x)$ 是 light_λ 波长的函数，以及从场景点 x 到摄像机的距离 $d(x)$ 的 t：

$$T_\lambda(x) = 10^{-\beta_\lambda d(x)} = \frac{E_\lambda(x, d(x))}{E_\lambda(x, 0)} = N_\lambda(d(x)), \quad (2)$$

where β_λ 为波长依赖的介质衰减系数，如图 1 所示。假设由 x 发出的光束在经过传输介质前后的能量分别为 $E_\lambda(x, 0)$ 和 $E_\lambda(x, d(x))$ 。归一化剩余能量比 N_λ 对应于每个距离门控单位的剩余能量与初始能量的比值。它在水中的值取决于光的波长。例如，红光具有较长的波长；因此，其衰减更快，比其他波长吸收更多，这导致大多数水下年龄的蓝色。更多的细节可以在 Ref 中找到。[20]

水下图像退化模型[20]不同于在图像去模糊处理中广泛应用的[29]的图像退化模型，图像去模糊[30]，和图像分辨率[31]。

具体来说，水下图像 d 降解模型与图像去危害模型相似，由于波长-端光吸收和散射特性，它更加复杂。与清晰图像与模糊核卷积运算来模拟模糊年龄的图像去模糊模型和原始图像下采样模拟剩余图像的图像分辨率模型相比，水下图像模型主要关注颜色和视觉性的退化。

4. 提出的UWCNN模型

在这里，我们讨论了提出的UWCNN模型的细节，然后提出一个后处理阶段，以进一步改进进一步的结果。

4.1. 网络架构

受到最近深度网络架构在模式识别[12,32]方面的成功的启发，我们提出了一个轻量级的网络

水下图像和视频增强。图 2 显示了我们的 UWCNN 模型的原始结构，它是一个紧密连接的 FCNN。如下所述，我们将介绍它的基本构建模块和超参数。输入到我们的网络的是一个 RGB 图像 U 。

4.1.1. 残留物

不像[3]等传统的端到端方法通过学习映射函数 $I=f-1(U)$ 直接预测干净的潜像 I ，我们允许我们的网络学习合成的水下图像和它的干净对应图像之间的区别。需要注意的是，这种合成图像生成任务是水下图像增强和恢复领域的一个重要目标，我们将在第 5 节中详细讨论。在随后的层中，在到达非损耗层之前，通过许多卷积滤波器进行处理。虽然我们的网络不是有意的很深，但仍然有消失或爆炸梯度的可能性。为了避免在训练迭代过程中出现这类问题，我们通过在损失功能之前添加网络的输入，即 $\Delta(U, \theta)$ （见下文）：

$$I = U + \Delta(U, \theta), \quad (3)$$

其中，“+”是元素级的加法操作。

4.1.2. 增强单元

UWCNN 有一个模块化的架构，由具有相同结构和组件的增强单元（电子单元）组成。上位 r 和 c 是 ReLU 和卷积的符号，然后卷积和 ReLU 对在第 L 个块中的第一次运算由

$$z_{l,0} = r(c(U; \theta_{l,0}), \quad (4)$$

其中 $z_{l,0}$ 是第 1 个卷积-第 1 个残差增强单元 and $\theta_{l,0}$ 对的输出，0 是与之相关的一组权重和偏差。通过组合卷积-relu 对，我们得到

$z_{l,n} = r(c(\dots r(c(U; \theta_{l,0}) \dots); \theta_{l,n}), \quad (5)$ 第 l 个块的输出是通过沿着每个独立的数据卷积的三维连接，输出 z 和输入图像 U 作为：

$$b_{l,h} = h(z_{l,0}; \dots; z_{l,n}; U). \quad (6)$$

(l+1) 增强单元的输出如下：

$$b_{l+1} = h(z_{l+1,0}; \dots; z_{l+1,n}; U; b_l). \quad (7)$$

最后，我们将所有增强单元链化，将该链的输出与非卷积层与 parameters $\theta_{l+m, n}$ 进行卷积，预测该分量为 $\Delta(U, \theta) = c(b_{l+m, \theta_{l+m, n}})$ 。

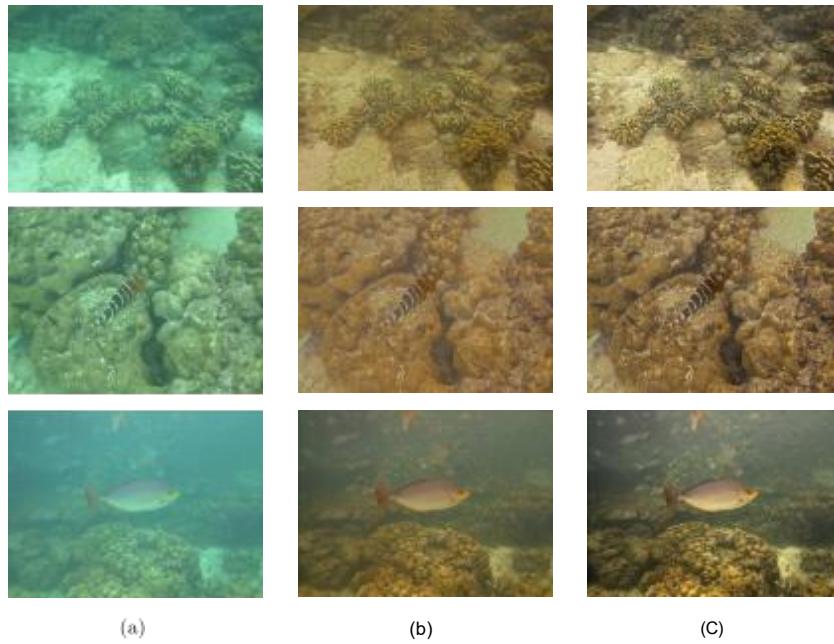


图3。作为定性分析的样本结果。(a)真实世界水下图像。(b)UWCNN的结果。(c)UWCNN+的结果。如图可见，我们的方法(即，UWCNN和UWCNN+)在重建准确和生动的潜在图像时，去除绿色的sh色调。



图4。来自nyu-v2RGB-D数据集[37]的10种合成水下图像

使用一个示例图像及其深度贴图。

4.1.3. 网络层

我们的网络由三个不同的层组成，用不同的颜色表示，如图2所示。第一种类型是由“Conv”表示的卷积层，它由16个大小为 $3 \times 3 \times 3$ 的卷积器组成，为第一层生成16个输出特征图，而随后的卷积层使用 $3 \times 3 \times 16$ 过滤器进行映射。第二种类型是用于引入非利接宽度的激活层“ReLU”。第三种类型是“Con-cat”层，它用于连接每个块之后的所有卷积层。最后一个卷积层表示网络的最终输出。

4.1.5. 网络深度

我们的网络是模块化结构的，由三体积单元组成，其中每个单元同样由三体积层组成。我们在网络的末端有一个卷积1层；因此，使我们的网络的整个深度只有10层。这使得我们的模型在计算上非常昂贵，并且在训练和推理方面非常实用。此外，这种轻权网络结构可以很容易地扩展宽水视频逐帧增强，这是实际应用中所需要的。这种轻盈的网络结构主要得益于嵌入式先验，它促进了网络的训练和推理，鼓励了设计用于模式识别、目标检测和可视化理解的相似网络。

4.1.4. 紧密连接

我们在每个块的末尾堆叠所有的卷积层。这种技术不同于[33]中提供的丹麦网，其中，每个卷积层与同一块中的其他卷积层连接。此外，我们不使用任何完全连接的层或批处理归一化步骤，这使我们的网络工作记忆有效和快速。此外，我们还为每个街区提供输入年龄。卷积层与输入数据的叠加减少了对一个非常深的网络的需要。总之，我们的网络是唯一的，因为(i)输入图像应用于所有的增强单元，并且(ii)它只包含全卷积层，没有任何批处理归一化步骤。

4.1.6. 减少边界伪影

在底层视觉任务中，系统的输出大小需要等于输入。这种要求有时会影响到边界伪影。为了避免这种现象，我们强制执行两种策略：(i)我们在我们的net工作中不使用任何池化层，以及(ii)我们在每个卷积层之前添加零。因此，UWCNN网络的最终输出图像在边界周围几乎是无差异的，并且与输入图像的大小相同。



图5。对来自测试集的样本进行定性比较。(a)原始的水下图像。红色[21]的(b)结果。 (c)UDCP[22]的结果。 (d)ODM [25]的结果。 (e)UIBLA[26]的结果。 (f)我们的研究结果。(g)基本事实。从上到下的第一列中的水下图像类型为I型、3型、5型、7型、9型、I型、II型和III型。我们的方法消除了光的吸收效应，并恢复了原始的颜色，没有任何伪影。

4.2. 网络损失

为了重建图像，我们使用 L_2 损失，它可以很好地保持边缘的锐度和细节，导致边缘模糊会导致很大的误差。我们将估计的残差加到输入的水下图像中，然后计算12个损失：

$$\ell_2 = \frac{1}{M} \sum_{i=1}^M \left| \left[\mathbf{U}(x_i) + \Delta(\mathbf{U}(x_i), \theta(x_i)) \right] - \mathbf{I}^*(x_i) \right|^2, \quad (8)$$

其中 $\mathbf{U}(x_i) + \Delta(\mathbf{U}(x_i), \theta(x_i)) = \mathbf{I}(x_i)$ 是他估计的潜在像 x_i , $i = 1, \dots, M$ 的像素值如式所述。(3) 而我**，我也是那个真实值

此外，我们在目标函数中加入了SSIM损失，将结构和纹理的相似性强加于潜在图像。我们使用灰色图像来计算SSIM分数。对于每个像素 x ，SSIM值在像素体周围的 13×13 图像补丁内计算：

$$SSIM(x) = \frac{2\mu_{l*}(x)\mu_l(x) + c_1}{\mu_{l*}^2(x) + \mu_l^2(x) + c_1} \cdot \frac{2\sigma_{l*l}(x) + c_2}{\sigma_{l*}^2(x) + \sigma_l^2(x) + c_2}, \quad (9)$$

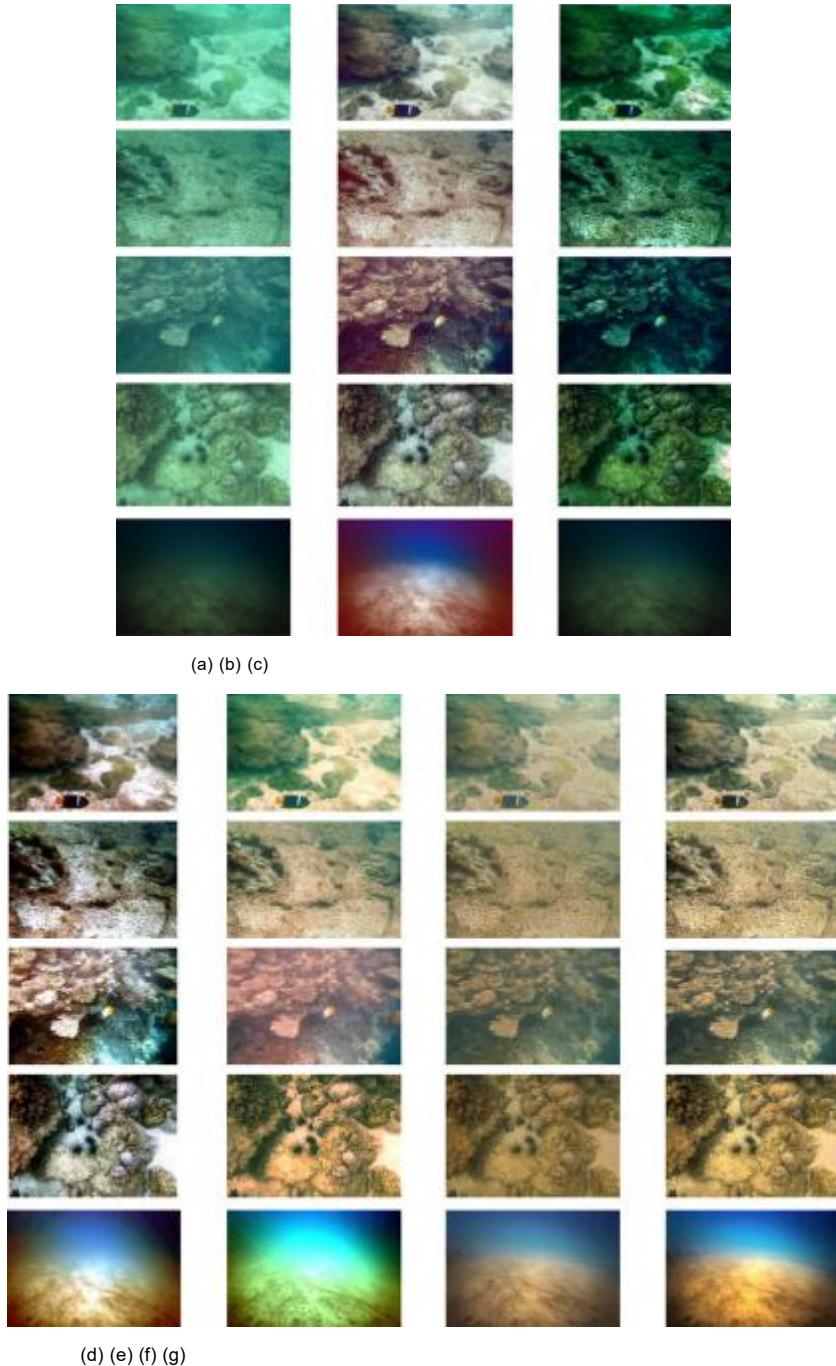
where $\mu_l(x)$ and $\sigma_l(x)$ 响应来自潜在图像 I 的图像补丁的平均和标准化发展，同样， $\mu_{l*}(x)$ and $\sigma_{l*}(x)$ 是来自地面 truth 图像 I^* 的补丁。 covariance $\sigma_{l*l}(x)$ 是在来自 I 和 I^* 的补丁之间计算的像素 x 。我们设置常数 $c_1 = 0.02$ 和 $c_2 = 0.03$ 基于默认的 SSIM 损失。我们的模型对这些默认值不敏感。不过，我们还是为了做一个公平的比较。SSIM 的损失被表示为：

$$L_{SSIM} = 1 - \frac{1}{M} \sum_{i=1}^M SSIM(x_i). \quad (10)$$

最终的损失函数 L 是 MSE 和 SSIM 的聚集： $L = 1.2 + L_{SSIM}$ 。 (11)

4.3. 后处理

UWCNN 生成增强图像没有颜色投射的可见性。然而，由于我们的训练数据对的局限性(将室内图像作为潜在图像，以及使用上述的非水下图像形成模型作为水下对应的室内图像的同步大小的图像



(a) 真实世界的水下图像。红色[21]的(b)结果。

的UWCNN的结果。(g)我们的UWCNN+的结果。我们的方法（即UWCNN和UWCNN+）产生的结果不需要任何视觉伪影、颜色偏差和过饱和度。它还揭示了空间主题和细节。

图像)，增强后的图像的动态范围较低。人们会期望增强的结果会有生动的色彩和更高的对比度。

为了解决这个问题，我们采用了一个简单而有效的后处理调整阶段。我们用后处理表示UWCNN为UWCNN+。图像首先转换为HSI颜色空间。然后，将其在HSI颜色空间中的饱和度和强度分量的范围正常划分为[0,1]，结果为：

$$y_{out} = \frac{y_{in} - y_{min}}{y_{max} - y_{min}}, \quad (12)$$

其中 y_{max} 和 y_{min} 是UWCNN图像中的最大值 m 和最小饱和度或强度值。在这个简化和强度归一化之后，我们将简化的结果转换回RGB颜色空间。

样本结果如图3所示。由此可见，UWCNN有效地消除了现实世界水下图像中主要的绿色畸变，并显著地改善了这些结构，同时保持了自然外观和其年代的真实性。与UWCNN相比，UWCNN+的饱和度和强正态化进一步提高了对比度和亮度，揭示了更多的细节。

表1

合成10种水下图像类型的 $N\lambda$ 值。

类型	I	ia	ib	微光	罗马数字3
蓝色	0.982	0.975	0.968	0.940	0.89
格林	0.961	0.955	0.950	0.925	0.885
红色的	0.805	0.804	0.830	0.800	0.750
类型	1	3	5	7	9
蓝色	0.875	0.800	0.670	0.500	0.290
格林	0.885	0.820	0.730	0.610	0.460
红色的	0.750	0.710	0.670	0.620	0.550

5. 提出的水下图像合成系统算法

不像高级的视觉任务[34–36] 在经常有大型训练数据集的情况下，缺乏具有相应地面真相的水下图像数据集，限制了基于深度学习的水下图像增强素和质量评估的发展。在此基础上，我们提出了一种基于水下成像物理模型和水下场景光学特性的水下图像合成算法。据我们所知，这是第一个基于水下图像的物理模型算法，可以延迟不同的水类型和退化水平，这是实现水下图像和视频增强的重要贡献。

为了合成水下图像退化任务，我们使用了[9]中描述的衰减系数对于海洋和海洋的不同类型（即开放水域的I、IA、IB、II和III，沿海水域的d1、3、5、7和9）。如前所述，i型是最清晰的，iii型是最浑浊的开阔海水。同样地，对于沿海水域，1型是最清澈的，9型是最浑浊的。我们应 用 eq. (1) 和(2) 利用rgb-DNYU-v2室内数据集[37]构建10种水下图像数据集。它由1449张图片组成。我们选择前1000张图像作为训练集，剩下的449张图像作为测试集。

为了合成一个水下图像，我们首先生成一个兰德原子均匀的全球大气光 $0.8 < B\lambda < 1$ 。然后，将深度 $d(x)$ 从0.5 m验证到15 m，然后得到表1所示的红、绿、蓝通道对应的 $N\lambda$ 值。对于每张图像，我们基于随机的 $B\lambda$ 和 $d(x)$ 生成5张水下图像；因此，我们得到了5000个样本的训练集和2495个样本的测试集。为了提高计算效率，我们将这些图像重新设置为 310×230 。我们总共根据不同的水类型合成了10个水下图像数据集。

图4 显示了一个样本的这十种不同类型的地下图像。很明显，i型、ia型和ib型的皮肤和皮肤图像在物理外观和特征细菌学上是相似的。因此，我们从10个模型中选择了8个来显示合成水下图像的结果。

6. 实验评价

在这部分中，我们使用最先进的水下图像增强方法对合成的和真实世界的水下图像进行定性和定量的计算。此外，我们还比较了不同方法在水下视频上的性能。这些进行比较的方法包括UDCP[22]，RED[21]，ODM[25]，和UIBLA[26]。我们使用推荐的参数集运行作者提供的源代码，以产生客观评估的最佳结果。在没有光衰减系数的真实图像中，我们应用我们学到的十大UWCNN模型中的每一个，并呈现出视觉上更有吸引力的结果。这一过程可以通过使用分类阶段来选择来改进

表2

对测试集的定量评估。如上所示，我们的方法在所有水下图像类型的所有指标中都获得了最好的分数。

	类型	RAW	红色	UDCP	视网膜血管血压测定法	乌布拉	我们的
MSE	1	2367.3	3489.7	2062.3	2508.6	2812.6	587.70
	3	2676.5	4953.2	3380.6	3130.1	3490.1	747.50
	5	4851.2	8385.8	6708.9	3488.9	4563.7	1295.1
	7	7381.1	9809.8	8591.6	5337.1	6737.9	2974.1
	9	9060.6	5952.3	9500.1	10634.0	8433.1	4121.5
	I	1449.0	936.9	1020.7	1272.0	1492.2	209.70
	微光	941.9	851.3	1466.0	1401.9	1141.4	251.60
	罗马数字3	1851.0	2240.0	2337.6	1701.1	1697.8	456.40
	1	15.535	15.596	15.757	16.085	15.079	21.790
	3	14.688	12.789	14.474	14.282	13.442	20.251
信号-噪音功率比	5	12.142	11.123	10.862	14.123	12.611	17.517
	7	10.171	9.991	9.467	12.266	10.753	14.219
	9	9.502	11.620	9.317	9.302	10.090	13.232
	I	17.356	19.545	18.816	18.095	17.488	25.927
	微光	20.595	20.791	17.204	17.610	18.064	24.817
	罗马数字3	16.556	16.690	14.924	16.710	17.100	22.633
	1	0.7065	0.7406	0.7629	0.7240	0.6957	0.8558
	3	0.5788	0.6639	0.6614	0.6765	0.5765	0.7951
	5	0.4219	0.5934	0.4269	0.6441	0.4748	0.7266
	斯西姆	0.2797	0.5089	0.2628	0.5632	0.3052	0.6070
斯西姆	9	0.1794	0.3192	0.1624	0.4178	0.2202	0.4920
	I	0.8621	0.8816	0.8264	0.8172	0.7449	0.9376
	微光	0.8716	0.8837	0.8387	0.8251	0.8017	0.9236
	罗马数字3	0.7526	0.7911	0.7587	0.7546	0.7655	0.8795

最好的模式，这就把它作为未来的工作留下了。对于合成数据，我们用前哨处理的结果，因为模型是由合成数据得到的，因此不需要强度y和饱和归一化。最后，我们进行了一项消融研究，以证明我们的网络中每个成分的影响。

6.1. 网络实施与培训

我们使用ADAM训练我们的模型，并将学习率设置为0.0002， β_1 为0.9， β_2 为0.999。我们在轮胎训练过程中确定了学习率。batch的大小被设置为16。优化一个超过20个时代的模型大约需要3个小时。我们使用tensor-flow作为内部(R)i7-6700k CPU、32GB内存和Nvidia GTX 1080 Ti GPU上的深度学习框架。

6.2. 水下合成图像的评价

我们首先展示了从我们的测试集中合成的水下r图像的水下图像增强的结果。在图5中(a)，合成的水下图像与[9]的测量结果一致。红色[21] 对透明型有效，即1型、3型、5型和I型；然而，对于混浊型，即7型、9型、II型和III型，它在这些图像上留下了阴霾，更多，它引入了颜色偏差离子。同样，UDCP [22]产生明显的暗色结果，而ODM[25] 和UIBLA[26] 介绍了官方的颜色或颜色的设计。另一方面，我们的方法不仅提高了图像的可见性，而且恢复了美观的纹理和充满活力而真实的色彩。与其他方法相比，我们的结果的视觉质量反映了地面的真相。

此外，我们量化了在合成测试集上的恢复图像的准确性，包括每种类型的2495个样本。表2中，精度通过三个不同的指标来衡量：均方误差 (MSE)、峰值信噪比 (PSNR) 和结构相似度指数度量 (SSIM) [38]。在MSE和PSNR指标的情况下，较低的MSE (较高的PSNR) 表示结果在图像内容方面更接近地面真相。在SSIM度量的情况下，较高的SSIM分数意味着存在-侮辱在图像结构方面更类似于地面真实值

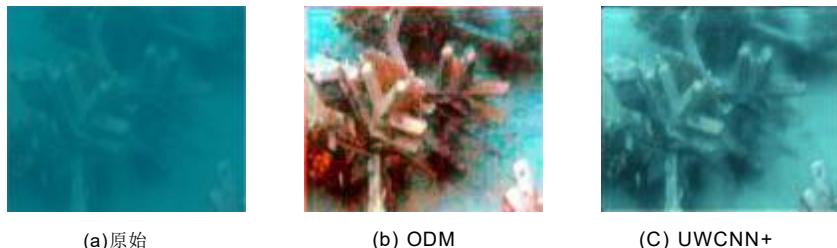


图7。与ODM [25]的比较。 (a)真实世界的水下图像。(b)结果由ODM [25]产生
我们的UWCNN+产生。(如果要解释本图例中对颜色的参考资料,读者可以参考本文的网络版本。)



图8。与UIBLA[26]的比较。 (a)真实世界的水下图像。(b)结果由UIBLA[26]产生。

这是一个失败的案例,因为只有绿色的色调被增强。(c)结果由我们的UWCNN+产生。

和纹理。在这里,给出的结果是平均分数。粗体中的值表示最佳结果。
。

由此可见,在我们测试的所有水下图像增强方法中,我们的方法在所有指标和所有退化类型中都表现最好,证明了它的有效性和鲁棒性。对于SSIM响应,我们的方法至少比最好的方法好10%。同样,我们的PSNR比比较方法更高(MSE分数表明的错误较小)。

6.3. 评估真实世界的水下图像

在这一部分中,我们在真实世界的水下图像上评估了所提出的方法。与竞争方法的视觉比较如图6所示。使用过的
真实世界的水下图像具有不同的色调、光线和对比度。

先看一下图6 可能给人的印象是ODM[25]的结果 和
UIBLA[26] 可能是;然而,谨慎揭示了ODM [25] 由于在
ODM[25]中之前使用的直方图分布,导致了过度增强和过度饱和(除了颜色模型)
并不总是有效的。同样的,由UIBLA [26]产生的图像
该方法不是自然的过度增强,该方法的缺点是背景光的鲁棒性和 p 波估计的中传输分数是次优的。图。7
和8显示ODM [25]的故障案例 和UIBLE [26]。红色[21] 和
UDCP[22] 对输入几乎没有什么影响。相比之下,我们的UWCNN+显示了现实世界缩小的良好结果,没有引入任何人工颜色,颜色模型,过度增强或增强不足的区域。

观察图中的故障情况。7 和8,我们可以发现,ODM的[25]倾向于引入额外的颜色(例如,图7中珊瑚周围的红色) 虽然我们的方法形成了对比,但与ODM[25]的性能相似, 但主要是原始水下图像的真实颜色分布。对于UIBLA [26]的故障案例在图8中, 它加重了绿色的颜色,并产生了视觉上令人不快的结果。与此相反,我们的方法消除了颜色投射,提高了对比和亮度,从而产生了更好的可见性和愉快的感知。



图9。真实世界的水下图像与不同的音调和退化水平。

我们注意到评估和[39,40] 是倾斜的曝光或增强,其中直方图等化方法被认为产生更好的分数。对于更主观的评估,我们进行了用户研究,以提供现实的反馈和量化主观视觉质量。我们从互联网和相关论文中收集了20张真实世界的水下图像。我们在图9中显示了来自这个数据集的样本。图6中显示了一些对应的侮辱。

对于用户研究,我们随机对结果进行排序,然后在屏幕上显示给人类超受试者。有20名参与者具有图像处理的专业知识。每个受试者根据感知到的视觉质量对结果进行排名,从1到5,其中1最差,5i最好。人们期望

表3

用户研究真实世界的水下图像数据集。最好的结果是用粗体显示的。

红色	UDCP	视网膜血管 血压测定法	乌布拉	UWCNN+
分数	2.95	2.55	3.25	3.20 3.35

高对比度、良好的能见度、自然颜色和自然纹理的结果应得到较高的等级，而过度增强/曝光、曝光不足/曝光、颜色模型和伪影的结果应得到较低的等级。平均得分见表3。

我们的UWCNN+获得了最高的排名，这表明我们的方法可以从主观视觉的角度产生更好的真实世界水下图像的性能。

6.4. 对水下视频的评价

为了验证我们的模型的水下视频增强能力，我们对水下视频s进行了实验。由于空间有限，我们只在图中给出了部分实验结果。10 .

表4

不同方法的平均运行时间。最好的结果是用粗体显示的。

时间	红色	UDCP	视网膜血管 血压测定法	乌布拉	UWCNN-C/-G
	3.250	3.319	5.829	47.254	2.250/ 0.225

如图10所示，我们的方法可以修改彩色模型，提高水下视频的对比度。此外，我们在不同帧之间的结果是一致的，并且没有超过闪烁的伪影。相比之下，比较方法在不同帧之间产生一致性增强，降低了其视觉质量。例如，对于第54帧，ODM [25]产生视觉上的愉悦感；然而，这种方法在第1-4帧中引入了红色的模型。其他方法也具有类似的一致性增强性能。此外，我们报告了不同方法的热处理时间（秒），以证明我们的模型可以用于表4中的逐帧视频增强器。大小为640×480的图像的平均运行时间被计算在上述机器上。UWCNN-C/-G表示我们的模型分别只使用CPU或GPU运行。

表4中，我们的UWCNN-G比比较方法快，有很大的边际，这可能受益于GPU加速。

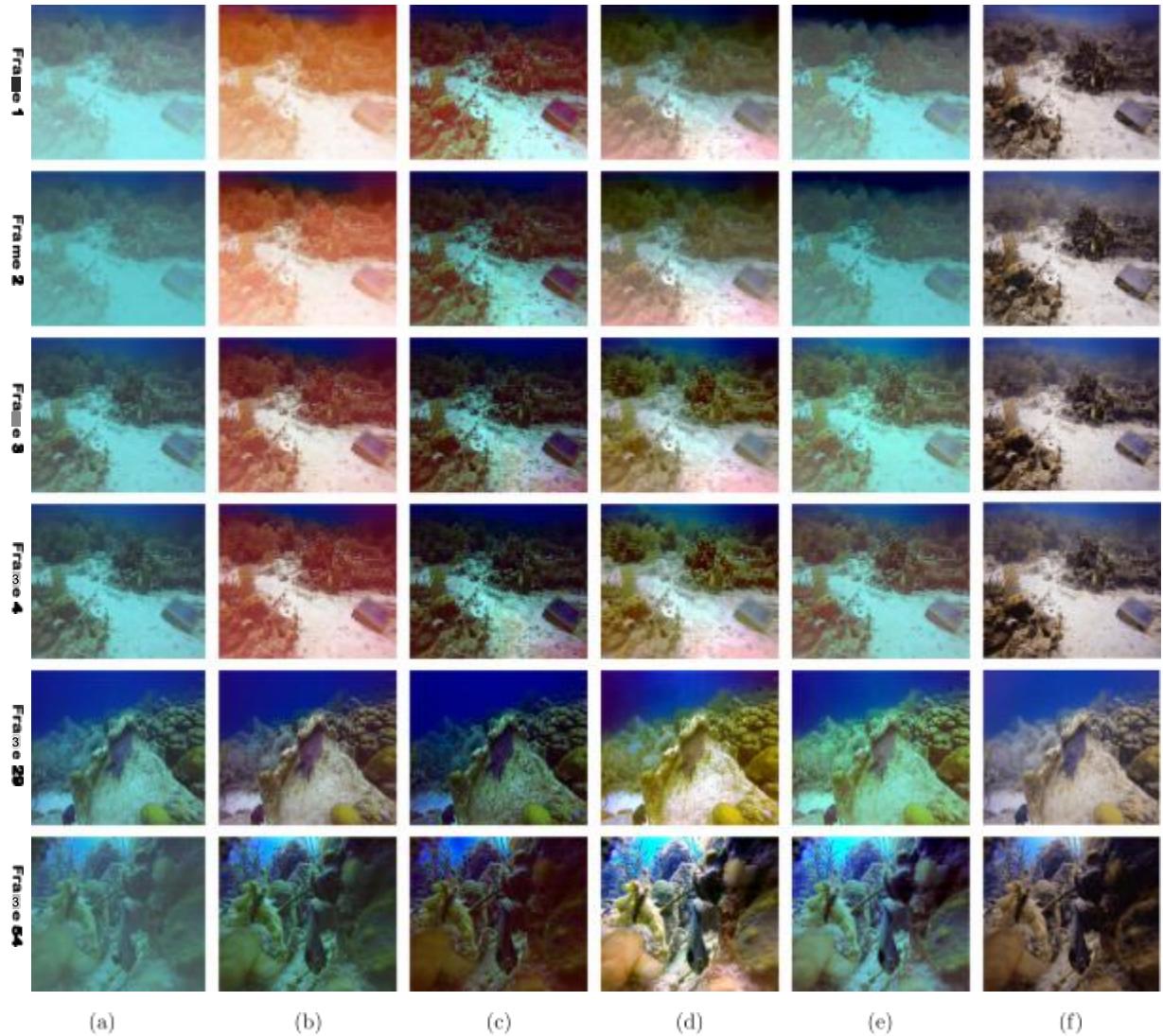


图10. 对水下视频的定性比较。(a)原始水下视频(从上到下是frame 1、帧2、帧3、帧4、帧29和帧54英寸这个视频)。红色[21]的(b)结果。 (c)UDCP[22]的结果。 (d)ODM[25]的结果。 (e)UIBLA[26]的结果。 (f)我们的UWCNN的结果。



(a)型-1型(b)-W/ O SSIM (C) UWCNN (d) GT

图11。SSIM损失重要性的一个例子。(a)一种具有1型退化的水下图像。

因为背景与GT不同。(c)结果由我们的UWCNN产生。(d)地面真理。

然而，我们的UWCNN-Cranks排名第二快，这表明我们的轻量级网络结构也有助于我们的方法的处理速度。

6.5. 消融研究

为了证明我们的网络中每个组件的影响，我们进行了一项基本的研究，包括以下实验：(i)没有重复学习的UWCNN (UWCNN-w/oRL)，(ii)没有密集连接的UWCNN (UWCNN-w/oDC)，以及(iii)没有SSIM损失的UWCNN (UWCNN-w/o SSIM)。由于空间有限，定量评价仅在I型和III型合成测试集上进行。MSE、PSNR、SSIM的平均得分见表5。

表5

I型和III型测试集的定量结果。每个评估的最佳结果用粗体表示，而第二好的结果用下划线表示。

	类型	-w/o RL	-w/o DC	-w/o SSIM	UWCNN
MSE	1	756.96	648.18	398.77	587.70
	罗马数字	542.68	789.76	402.92	456.40
	3				
信号-噪 音功率比	1	20.290	20.805	22.902	21.790
	罗马数字	21.556	20.289	23.026	22.633
	3				
斯西姆	1	<u>0.8450</u>	0.8449	0.8214	0.8558
	罗马数字	<u>0.8579</u>	0.8359	0.8151	0.8795
	3				

从表5，我们可以看到，用残余学习（即UWCNN）取代传统的学习策略（即UWCNN-w/oRL）可以提高性能。将UWCNN与UWCNN-w/oDC进行比较，我们发现密集的连接也可以提高水下图像增强的性能。使用SSIM损失（即UWCNN），以降低MSE和PSNR分数（即UWCNN-w/o SSIM）为代价，改善了结构和纹理的相似性。然而，为了获得更好的主观感知，这些牺牲是不必要的。图11，这说明了SSIM损失的重要性。图11中，在添加SSIM损失后，UWCNN的结果比UWCNN-w/o SSIM的结果背景更平滑。

7. 结论

我们提出了一个受水下场景先验启发的水下图像和视频增强网络。在合成和真实水下图像和视频上的实验证明了该方法的鲁棒性和有效性能。我们的优点是，我们的方法在每个卷积层上只包含10个卷积层和16个特征映射，这提供了对GPU平台形式的高效训练和测试。实验结果还表明，我们网络中使用的剩余学习、学习和SSIM损失定量地提高了性能。

在未来，我们将只使用单一模型来预测UWCNN单盲模型的正确输出

在UWCNN模型增强过程中获得进一步的消除，并在完整的图像退化模型中考虑室内训练数据引起的低对比度。借鉴为模式识别和计算机视觉设计的深度模型的有效网络结构和损失，我们将尝试进一步提高该方法的性能。

确认的内容

这项工作得到了美国国家自然科学协会的资助。
中国的镇静 (61771334) 以及基本科研业务费基金
为中央大学拨款 2019RC039.

参考文献

- [1] D. 阿卡尼雅克, T. Treibitz, 一个修正的水下图像形成, 在: Proc. 电器和电子工程师学会 Int. 会议压缩。Vis. 模式识别。(CVPR), IEEE, 2018, 第 6723–6732 页。
- [2] S. Chikkerur, N. 卡特赖特 V. 戈文达拉朱 使用stft增强指纹功能 分析、模式识别。40 (1) (2007)198–211。
- [3] G. 知识 A. 阿金塔约, S. 萨卡尔, Llnet: 一个深的脚趾编码器方法自然低光图像增强, 模式识别 t.61 (2017) 650–662。
- [4] W. 王, Y. 徐, J. 沈, S. 朱, 关注时尚语法网络 地标检测和服装分类分类, 在: Proc. IEEE 国际。Co nf. 压缩。Vis. 模式识别。(CVPR), IEEE, 2018, pp. 4271–4280。
- [5] F. 阳 K. 燕, S. Lu, H. Jia X. 谢, W. 高, 注意驱动的过程 模式识别, 模式识别。86 (2019)143–155。
- [6] W. 王, J. 沈, F. Porikli, R. 杨, 半监督的视频对象分割器- 超轨迹, IEEE Tr. 模式肛门。马赫数。知识41 (4) (2019)985–998。
- [7] C. 郭, C. 列支敦士登 J. 郭, R. Cong H. Fu, P. Han, 层次结构特性, 驱动器抵抗体- 最终学习深度地图超分辨率, IEEE跨式。图像处理。28 (5) (2019)2545–2557。
- [8] Z. 吴, C. 沈 A.V.D.Hengel, 更广泛或更深入: 重新访问resnet模型 视觉识别, 模式识别。90 (2019)119–133。
- [9] D. 伯曼职业名称 T. Treibitz, S. 阿维丹, 潜水入雾线: 颜色恢复 水下图像, 在: Proc. 小鲱鱼马赫数。Vis. 会议(BMVC), 施普林格, 2017年, pp.1–11。
- [10] W. 王, J. 沈, 深度视觉注意预测, IEEE转换。图像处理。27 (5) (2018)2368–2378。
- [11] H. 宋, W. 王, S. 赵, J.S.K. 林, 金字塔扩张的深处 对于视频显著性目标检测, 在: Proc. 欧洲人会议压缩。Vis. (ECCV), 施普林格, 2018年, 第715–731页。
- [12] J. Gu, Z. 王, J. 库恩, L. 马, A. 沙鲁迪 B. 帅, T. 刘, X. 王, J. 蔡, T. 陈, 卷积神经网络s, 模式识别的最新进展。77 (2018)354–377。
- [13] C. 列支敦士登 J. 郭, 通过模糊和颜色校正增强水下图像, J. 电子。Imag .24 (3) (2015)033023–1 033023–10。
- [14] C. Ancuti, C.O.Ancuti, 通过融合增强水下图像和视频, 在: 过程 IEEE 国际。会议压缩。Vis. 模式识别。(CVPR), IEEE, 2012, pp. 81–88.
- [15] C. 列支敦士登 J. 郭, C. 郭, R. Cong J. 宫, 一种水下图像的混合方法 校正, 模式识别。拉脱维亚的94 (2017)62–67.
- [16] C. 列支敦士登 J. 郭, C. 郭, 从水中出现: 水下图像彩色相关 基于弱监督的颜色传输, IEEE信号过程。拉脱维亚的25 (3) (2018)323–327。
- [17] Y. 郭, H. Li, P. 庄, 使用多尺度的水下图像增强 密集生成对抗网络, IEEE J. Ocean. 工程师(2019) 1–9.
- [18] C. Ancuti, C.O.Ancuti, C. Vleeschouwer, 色彩平衡和融合 图像增强, IEEE图像。图像处理过程。27 (1) (2018)379–393.
- [19] C. 列支敦士登 C. 郭, 仁, R, 丛, 侯, 廖, 水下图像增强基准数据集及以上 (2019) arXiv: 1901.05495。
- [20] J. 蒋, Y. 陈, 水下图像的波长增强 删除和删除侮辱, IEEE跨式。图像处理过程。21 (4) (2012)1756–1769.
- [21] A. Galdran, D. 帕尔多 A. Picn, A. 阿尔瓦雷斯-吉拉 自动红色通道 ter图像恢复。通勤。图像代表26 (2015) 132–145。

- [22] P. 德鲁斯, E. 鼻咽鼓管, S. 波特略, M. 坎波斯, 水下深度估算 以及基于单个图像的图像恢复, IEEE通信输出。图应用程序. 36 (2) (2016) 24–35.
- [23] K. He, J. 太阳 X. 唐 使用暗通道去除单图像雾霾, IEEE 反式模式肛门。马赫数。知识33 (12) (2011) 2341–2343.
- [24] C. 列支敦士登 J. 郭, S. 陈, Y. 唐 Y. Pang J. 王, 水下图像恢复 基于最小信息损失原理和光学特性 德沃特成像, 在: Proc. IEEE国际。会议图像处理过程。(ICIP), IEEE, 2016年, pp. 1993–1997.
- [25] C. 列支敦士登 J. 郭, R. Cong Y. Pang B. 王 水下图像增强系统 采用最小信息损失和直方图分布 反式图像处理过程. 25 (12) (2016) 5664–5677.
- [26] Y. Peng, P. 科斯曼 基于图像模糊度的水下图像恢复 和光吸收, IEEE反式。图像处理过程。26 (4) (2017) 1579–1594.
- [27] J. Li, K. 斯金纳, R. Eustice, M. 罗伯逊, 水门事件: 无监督生成系统 网络可以实现单一水下图像的实时彩色校正, IEEE机器人。Autom. 拉脱维亚的3(1) (2017) 387–394.
- [28] M. Sheinin, Y. 谢克纳, 下一个最好的无水视图, 在: Proc. IEEE国际。会议压缩。Vis. 模式识别它。(CVPR), IEEE, 2017, pp. 1436–1443.
- [29] C. 列支敦士登 C. 郭, J. Guo, P. 韩, H. Fu, R. Cong 受感知启发的单身人士 图像脱模糊网络与细化, IEEE转换。多段时间。(2019) 1.
- [30] Z. 沈, W.L.T. 徐, J. Kautz M. 杨, 深层符号Ic面部去模糊, 见: 过程 IEEE国际。会议压缩。Vis. 模式识别。(CVPR), IEEE, 2018, pp. 8260–8269.
- [31] H. 黄, H. He, X. 迷 J. 张, 人脸图像的超分辨率的使用 典型相关分析, 模式识别。43 (7) (2010) 2532–2543.
- [32] T. Lopes E. 德阿吉亚尔 F.D. 苏扎, T. 奥利维拉-桑托斯 面部表情识别 使用卷积神经网络: 用很少的数据进行裁剪 识别样本顺序, 模式识别。61 (1) (2017) 610–628.
- [33] G. 黄 Z. 刘, L. 范德曼, K. 温伯格, 紧密连接的卷积 调用网络, 在: Proc. IEEE国际。会议压缩。Vis. 模式识别。(CVPR), IEEE, 2017, 第4700–4708页。
- [34] W. 王, J.S.F. 郭, M. 程, A. 博尔吉 回顾视频显著性: 高度规模 基准测试 和一个新的模型, 在: Proc. IEEE国际。会议压缩。Vis. 模式 Recognit. (CVPR), IEEE, 2018, 第4894–4903页。
- [35] C. 周 J. 元, 闭塞行人零件探测器的多标签学习 检测、模式识别。86 (2019) 99–111.
- [36] C. 列支敦士登 R. Cong J. Hou, S. 张, Y. 钱, S. 光, 嵌套网络 用于光学遥感中目标检测的双流金字塔 年龄, IEEE跨。地球科学。重新参议员。(2019) 1.
- [37] N. Silberman, D. Hoiem, P. Kohli R. 费格斯, 室内分割和支持 Argrbd图像中推断, 在: Proc. 欧洲人会议压缩。Vis. (ECCV), 施普林格, 2012, pp. 746–760.
- [38] Z. 王, A. 博维克 H. Sherikh, E. Simoncelli, 图像质量评估: 来自 对结构相似性的错误可见性, IEEE转换。图像处理过程。13 (4) (2004) 600–612.
- [39] M. 杨, A. Sowmya, 一种水下彩色图像质量评价指标, IEEE跨。图像处理过程。24 (12) (2015) 6062–6071.
- [40] K. 帕内塔, C. 一般警报等级 S. Agaian, 受人类视觉系统启发的水下图像质量措施, IEEE J. Ocean. Eng. 41 (3) (2016) 541–551.

李崇义于2018年6月在李崇义获得中国天津大学博士学位。2016年至2017年, 他在澳大利亚国立大学(ANU)工程研究学院学习了一年的访问博士生。现在, 他是香港城市大学(CityU)计算机科学系的博士后研究员。他获得了北京图像与图形学会颁发的优秀博士学位论文奖。他的研究兴趣包括图像处理、计算机视觉和深度学习, 特别是在图像去模糊、水下图像增强、图像分辨率、低光图像增强、显著目标检测等领域。

Saeed安瓦尔是澳大利亚CSIRO(联邦科学和工业研究组织)数据61网络物理系统的研究员。他在澳大利亚国立大学(ANU)和Data61/CSIRO获得博士学位。他一直在巴基斯坦国家计算机和新兴科学多样性(NUCES)担任讲师和助理教授。他的主要研究兴趣是低级视觉、图像增强、图像恢复、计算机视觉化和优化。