



文本感知单图像镜面高光去除

Shiyu Hou^{1,2}, Chaoqun Wang^{1,2}, Weize Quan^{1,2}, Jingen Jiang^{1,2},
and Dong-Ming Yan^{1,2}(✉)

1 中国科学院自动化研究所模式识别国家实验室, 北京
100190{weize.quan,dongming.yan}@nlpr.ia.ac.cn² 中国科学院
大学人工智能学院科学, 北京100049

抽象的。从单个输入图像中去除不需要的镜面高光对于许多计算机视觉和图形任务至关重要。现有的方法通常会去除医学图像和特定对象图像的镜面高光, 但是它们无法处理带有文本的图像。此外, 文本检测和识别社区很少研究镜面高光对文本识别的影响。因此, 在本文中, 我们首先提出并研究文本感知的单图像镜面高光去除问题。核心目标是通过去除文本图像中的高亮部分来提高文本检测和识别的准确性。为了解决这个具有挑战性的问题, 我们首先收集了三个带有细粒度注释的高质量数据集, 这些数据集将适当发布以促进相关研究。然后, 我们设计了一个新颖的两级网络, 其中包含高光检测网络和高光去除网络。高光检测网络的输出提供有关高光区域的附加信息以指导后续的高光去除网络。此外, 我们建议一个测量集, 包括端到端文本检测和识别评估以及辅助视觉质量评估。对我们收集的数据集进行的广泛实验证明了所提出方法的优越性能。

关键词: 镜面高光去除·文本感知·数据集·神经网络

1 简介

由于物体的材质属性和拍摄环境, 现实世界的图像中经常存在镜面高光。人们总是希望减少或消除这些镜面高光以提高视觉质量并促进视觉和图形任务, 例如立体匹配[10,12]、文本识别[16]、图像分割[3,6]和照片一致性[30,31]。示例见图1, 端到端文本检测和识别的性能下降是由于

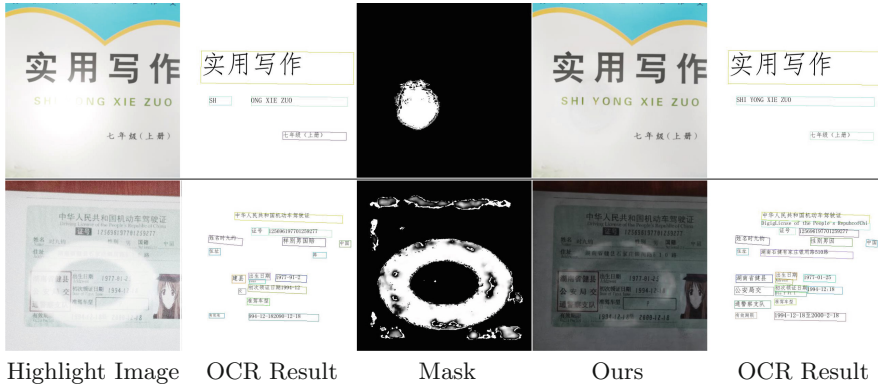


图 1.我们的方法的选定单图像镜面高光去除结果。“面具”和“我们的”分别是我们第一阶段和第二阶段网络的输出。

图像中存在高光，而我们的方法旨在检测并去除高光，以提高后续的 OCR 性能。在过去的几十年中，已经提出了许多方法来解决这个具有挑战性的镜面高光去除问题。这些现有的工作可以大致分为三类：基于二色反射模型的方法[7,23,26,28,32]，基于修复的方法[4,18,21,27]和基于深度学习的方法[8,15,20]。二色反射模型[24]线性地组合了漫反射和镜面反射，随后基于该模型提出了许多方法。这些方法通常需要一些简化假设。另外，在遇到颜色多样、纹理复杂的图像时，往往需要进行分割等预处理操作，效率低、实用性弱。基于修复的方法主要是利用图像修复高光部分后的原始图像内容。来自图像修复社区的技术。这种方法对于大的高光污染来说性能有限。考虑到单图像镜面高光去除的复杂性，最近提出了一些基于深度学习的工作[8,15,20]，例如卷积神经网络（CNN）和生成对抗网络（GAN）。借助深度模型强大的学习能力，这些基于深度学习的方法通常比传统的基于优化的方法具有更好的性能。然而，这些基于深度学习的方法需要大规模的训练数据，特别是配对的真实世界图像和必要的注释，这些数据非常耗时甚至难以收集。现有的镜面高光去除方法主要处理医学图像、自然图像和特定图像。然而，目前还没有针对文本图像的工作。另一方面，对于端到端文本检测和识别，提出了许多方法来处理任意形状和各种方向的文本。据我们所知，很少研究带有镜面高光污染的文本图像的情况。因此，在本文中，

with specular highlight contamination is rarely studied. Therefore, in this paper,

我们对文本感知的单图像镜面高光去除问题进行了广泛的研究，包括数据集收集、网络架构、训练损失和评估指标。我们工作的主要贡献如下：

– 我们首先在文献中提出并研究文本感知的单图像镜面高光去除问题。为了研究这个具有挑战性的问题，我们收集了三个具有细粒度注释的高质量数据集。– 我们提出了一种新颖的两阶段框架，通过两个子网络实现高亮区域检测和删除。高光检测网络提供有用的位置信息以方便后续的高光去除网络。对于训练目标，我们联合利用检测损失、重建损失、GAN 损失和文本相关损失来实现良好的性能。– 为了结果比较，我们建议一个全面的测量集，其中包含端到端文本检测和识别性能和辅助视觉质量评价。

2 相关工作

2.1 基于二色反射模型的方法

二色反射模型[24]假设图像强度可以通过漫反射和镜面反射的线性组合来表示。该模型已广泛用于镜面高光去除。Tan 等人基于最大色度-强度空间中漫反射点和镜面反射点的分布。[28]通过识别漫反射最大色度，然后应用镜面到漫反射机制来分离反射分量。受到局部彩色图像中漫反射最大色度平滑变化的观察的启发，Yang 等人。[32]通过应用双边滤波增强了色度估计的实时性能和鲁棒性。为了利用彩色图像的全局信息进行镜面反射分离，Ren 等人。[23]提出了一种基于二色反射模型的全局色线约束。傅等人。[7]将漫反射和镜面反射图像的估计重新表述为具有稀疏约束的能量最小化，可以近似求解。最近，Son 等人。[26]提出了一种凸优化框架，可以有效地去除自然图像的彩色和非彩色区域中的镜面高光。这些基于二色反射模型的方法在处理具有不同颜色和复杂纹理的图像时通常性能有限。

2.2 Inpainting-Based Methods

修复是通过传播已知区域的信息来完成图像的缺失区域，该技术可用于修复损坏的绘画或删除特定对象[5]。谭等人。[27]首先提出了

aged paintings or remove specific objects [5]. Tan *et al.* [27] first proposed an

基于修复的方法，通过结合基于照明的约束来去除高光。Ortiz 和 Torres [21] 设计了一个连接的矢量滤波器，集成到修复过程中以消除镜面反射。Park and Lee [22] 提出了一种基于彩色线投影的高光修复方法，但是该方法需要使用不同曝光时间拍摄的两张图像。基于修复的高光去除方法也被提出来处理医学图像，例如内窥镜图像 [4] 和阴道镜图像 [18]。然而，这些基于修复的方法仅对具有小区域高光污染的图像有效。

2.3 基于深度学习的方法

与上述两种方法不同，基于深度学习的方法不需要镜面高光模型假设，因此具有处理各种场景的潜力。李等人。[14] 提出感知器人工神经网络来检测牙齿图像的镜面反射，然后应用平滑空间滤波器来递归地校正镜面反射。由于缺乏配对训练数据，Funke 等人。[8] 采用循环 GAN 框架 [33] 并引入自正则化损失目标来减少非镜面区域的图像修改。同样，林等人。[15] 还采用了 GAN 框架并提出了一个多类判别器，其中将生成的漫射图像与真实图像和原始输入图像进行分类。穆罕默德等人。[20] 提出了两种用于去除面部镜面反射的深度模型（Spec-Net 和 Spec-CGAN）。前者以强度通道为输入，后者以 RGB 图像为输入。这些方法主要针对医学图像、特定对象图像或面部图像提出，而我们的工作关注文本图像。

3 个数据集

在文献中，没有公开可用的数据集来研究文本感知的单图像镜面高光去除问题。因此，在这项工作中，我们收集了三个高质量的数据集，包括一个真实数据集和两个合成数据集。数据集收集的流程和配对数据样本的示例如图 2 所示。

3.1 真实数据集

图 2(a) 说明了真实数据集收集的流程。对于真实数据集，我们收集了 2,025 个图像对：具有文本感知镜面高光的图像、相应的无高光图像和指示高光位置的二值掩模图像。图像内容包括身份证、驾驶证等，其中包含大量的文字信息。我们首先将透明塑料膜放在图片上并打开灯。然后，相机拍摄以获得高光图像。相应地，我们通过关闭获得无高光的图像

light image. Correspondingly, we obtain a highlight-free image by turning off

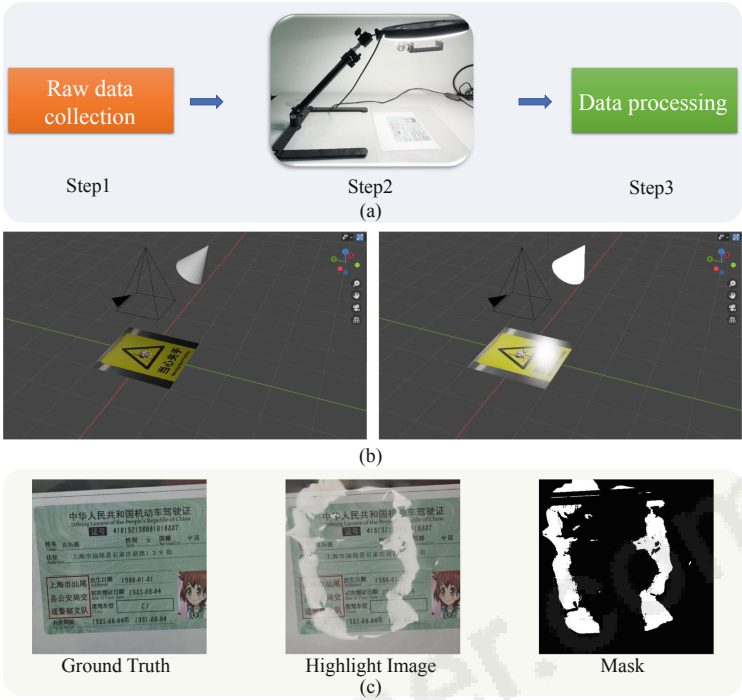


图 2. 真实数据集 (a) 和合成数据集 (b) 的收集管道, 以及配对数据样本 (c) 的示例。

光。通过调整塑料薄膜的位置, 高光的形状和强度会有所不同。将镜面高光图像和无高光图像通过差分 and 多重阈值筛选得到二值掩模图像。我们随机将该数据集 (名为 RD) 分为训练集 (1,800 张图像) 和测试集 (225 张图像)。

3.2 综合数据集

为了进一步丰富我们数据集的多样性, 我们使用 3D 计算机图形软件 Blender 构建了两组合成图像。图 2(b) 显示了合成数据集收集的流程。我们首先从超市和街道收集了 3,679 张带有文本的图像, 以及第 2 节中提到的 2,025 张图像。3.1.然后, 我们使用 Blender with Cycles 引擎自动生成 27,700 组文本感知镜面高光图像、相应的无高光图像和高光蒙版图像。特别是, 高光形状包括圆形、三角形、椭圆形和环形, 以模拟现实场景中的照明条件。材料粗糙度在[0.1,0.3]范围内随机设置, 照明强度在 [40,70]范围内随机选择。为了强制图像的文本区域出现镜面高光, 我们为 Blender 提供了 specular highlight on the text areas of the image, we provide the Blender with

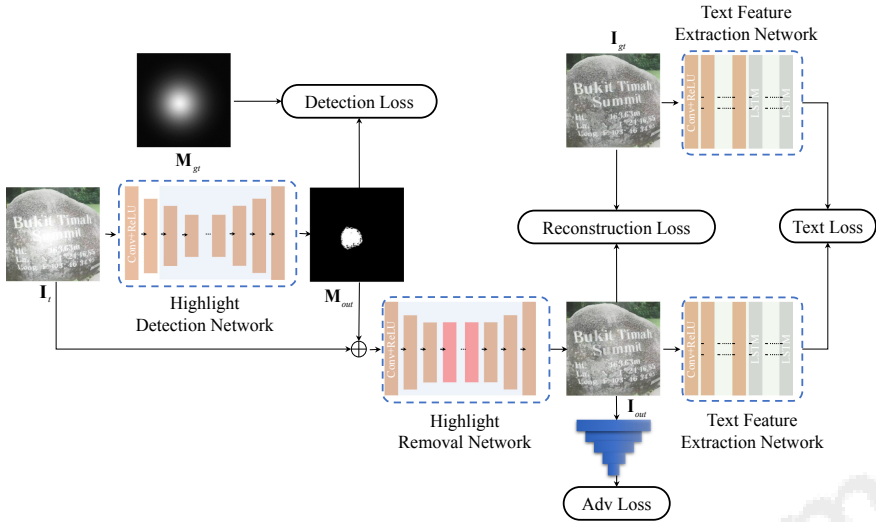


图3.我们提出的镜面高光去除框架的整体结构，它由高光检测网络、高光去除网络和基于补丁的鉴别器组成。符号表示通道级联。

通过文本检测模型CTPN[29]获得的文本区域的位置信息。因为产品或街景类别每张图像包含的文本较少，而身份证和驾驶执照中的文本更密集。在相同的光照条件下，这两种图像中受镜面高光干扰的文本信息的恢复难度是不同的。因此，我们将上述两类图像分为两个数据集，即SD1和SD2。SD1包含12,000个训练集和2,000个测试集。SD2包含12,000个训练集和1,700个测试集。注意，RD和SD2的图像内容是相同的。

4 提出的方法

在这项工作中，我们提出了一个两阶段框架来检测和删除文本图像中的镜面高光。整个架构如图3所示。下面，我们描述网络架构和损失函数的细节。

4.1 网络架构

突出显示检测网络。高光检测网络Net D将具有镜面高光的文本图像 I_t 作为输入并输出指示高光区域的掩模 M 。 M_{out} 的每个元素都在 $[0,1]$ 中，并且一个更大的 indicating the highlight regions. Each element of M_{out} is in $[0,1]$, and a larger

值代表图像I的相应位置被镜面高光覆盖的概率较高。由于 I_t 和 M_{out} 的宽度和高度相同，对于该网络，我们采用由三个下采样和上采样层组成的全卷积架构。每个下采样层后面跟着两个卷积层，每个上采样层后面跟着三个卷积层。

高光去除网络。在获得高光掩模 M_{out} 后，然后应用高光去除网络 Net_R 去除镜面高光并恢复文本信息。作为输入， Net_R 接受输入文本 $imageI_t$ 和检测到的高光蒙版 M_{out} 。我们的高光去除网络 Net_R 的输出是无高光图像 I_{out} 。通过引入 M_{out} ， $networkNet_R$ 可以更加关注高光区域并获得更好的去除性能。对于 Net_R 的网络架构，在这项工作中，我们采用具有跳跃连接的编码器-解码器网络。该网络由两个下采样层、四个残差块和两个上采样层组成。为了进一步增强去除性能，我们还应用了基于补丁的鉴别器[19]。鉴别器 D 包括一个卷积层和五个下采样层，内核大小为 5，步幅为 2。光谱归一化用于稳定鉴别器的训练。

4.2 损失函数

接下来，我们说明用于训练网络的损失函数。

突出显示检测损失。对于高光检测网络的目标函数，我们使用 l_1 损失，即 $L_{NetD} = M_{out} - M_{gt}$ ，其中 M_{gt} 是高光掩模的 groundtruth。

重建损失。重建损失是对像素和特征空间添加约束。像素感知损失由逐像素差异项和总变化项 (TV) 组成： $LP = 5 * I_{out} - I_{gt} + 0.1 * (I_{out}(i, j) - I_{gt}(i-1, j)) + I_{out}(i, j) - I_{gt}(i, j-1)$ 。特征感知损失包括感知损失[11]和风格损失[9]： $LF = 0.05 * \Phi(I_{out}) - \Phi(I_{gt}) + 120 * \Psi(I_{out}) - \Psi(I_{gt})$ ，其中 Φ 是预训练的 VGG-16 [25] 的特征图， $\Psi(\cdot) = \Phi(\cdot)\Phi(\cdot)$ 是 Gram 矩阵 [9]。特征感知损失提高了结果的视觉质量。

甘损失。在高光去除网络中，我们使用基于补丁的判别器 D 来增强结果的视觉真实感。对于 GAN 损失，我们采用 hinge 损失。因此， Net_R 的对抗性损失为 $LG = -E[D(I_{out})]$ 。用于训练判别器 D 的损失公式为 $L_D = E[\max(0, 1 - D(I_{gt}))] + E[\max(0, 1 + D(I_{out}))]$ 。

与文本相关的丢失。在这项工作中，我们的镜面高光去除是文本感知的。这意味着高光去除网络 Net_R 需要更多地关注恢复隐藏在高光后面的文本。为此，我们应用预先训练的文本检测和识别模型来提供监督

pre-trained text detection and recognition models to provide the supervision

关于文本恢复。更具体地说，我们对从上述两个预训练模型中提取的 I_{out} 和 I_{gt} 的特征图添加一致约束，并将文本相关损失表示为 $L_T = c=1 \phi_c(I_{out}) - \phi_c(I_{gt})_1 + \phi_d(I_{out}) - \phi_d(I_{gt})_1$ ，其中 ϕ_c 代表预训练 CTPN 的第 c 层特征图模型[1]， ϕ_d 表示预训练的 DenseNet [1] 的第 d 层特征图。为此，Net D 和 Net R 的总目标函数为 $L = \lambda_{NetD} L_{NetD} + L_P + L_F + \lambda_G L_G + L_T$ 。在所有实验中，我们设置 $\lambda_{NetD} = 10$ 和 $\lambda_G = 0.01$ 。

5 实验

5.1 实施设置

我们的网络是使用 TensorFlow 1.15 实现的。作为 GPU，我们使用 NVIDIA R 的 TITAN RTX。批量大小为 4 的 Adam 优化器 [13] 用于训练我们的网络，其中 $\beta_1 = 0.5$ 且 $\beta_2 = 0.9$ 。学习率初始化为 0.0001。在我们的实验中，所有图像的大小均为 512×512 。请注意，用于结果评估的文本识别模型与文本相关损失中使用的模型不同。

5.2 定性评价

我们在收集的三个数据集上将我们的方法与两种最新的先进镜面高光去除方法进行定性比较：Multi [15] 和 SPEC (Spec-CGAN [20])。结果如图4所示。在这三种方法中，我们的方法可以更好地去除高光并实现优异的端到端文本检测和识别性能。例如，我们的方法成功恢复了第三行中的姓名、地址和 ID 号。由于其盲去除特性，Multi 在第三行和第四行中具有明显的高光残留，而我们的方法可以更好地感知高光区域。与Multi相比，SPEC的结果亮点较少，但SPEC所遵循的cycleGAN框架的文本恢复能力有限。

5.3 定量评价

此外，我们从端到端文本检测和识别性能以及视觉质量方面对上述三种方法进行了定量比较。对于端到端文本检测和识别评估，我们采用通用指标[17]：召回率、精度和f-measure。我们选择当前先进的文本检测和识别算法Paddle OCR [2]来计算这三个指标。对于视觉质量评估，我们利用 PSNR 和 SSIM。表 1 报告了三种方法在我们的三个数据集上的数值结果。由于RD和SD2的图像内容相同，对于真实数据集（RD），我们使用所有三个数据集的RD训练集对在SD2上训练的模型进行微调 we fine-tune the model trained on SD2 using the training set of RD for all three



图 4. 我们的方法与 Multi [15] 和 SPEC [20] 的定性比较。

方法。从表 1 中，我们可以发现我们的方法在端到端文本检测和识别方面实现了最佳性能（参见第 3-5 列）。以召回为例，我们的方法可以将端到端文本检测和识别性能分别提高 6.89% (SD1)、3.07% (SD2) 和 13.65% (RD)。这一改进表明我们的方法可以更好地恢复隐藏在镜面高光后面的原始文本。另外，Multi 和 SPEC 的端到端检测和识别性能有时低于 Light Image。原因是这两种方法同时去除了高光和文本。对于 PSNR 和 SSIM，SPEC 最差，而我们的方法

light and texts as well. For PSNR and SSIM, SPEC is worst, while our method

和 Multi 对于合成数据集具有竞争力，并且我们的方法对于真实数据集优于 Multi。我们的方法的 PSNR 和 SSIM 有时低于 Multi 的，然而，这两个指标与人眼感知的视觉质量并不完全相同。更重要的是，我们关注高光去除后的端到端文本检测和识别性能，视觉质量是辅助方面。

5.4 消融研究

为了验证文本相关损失的有效性，我们进行了消融实验，并在表 2 中报告了相应的结果。我们观察到，我们的方法与文本相关损失的端到端文本检测和识别性能是一致的改进了三个数据集。这表明与文本相关的损失可以强制高光去除网络进行文本感知恢复。此外，我们可以发现我们的方法的端到端文本检测和识别性能已经优于 Multi 和 SPEC（将表 2 中每个数据集的第一行与表 1 中的相应行进行比较），即使有没有与文本相关的损失。

表 1. 我们的方法与两种最新最先进方法的定量比较：Multi [15] 和 SPEC [20]。所有三种方法都分别在我们收集的三个数据集上进行训练和测试。召回率、精度、f-measure 和 SSIM 均以 % 为单位。

Datasets	Methods	Recall \uparrow	Precision \uparrow	F-measure \uparrow	PSNR \uparrow	SSIM \uparrow
SD1	Light Image	85.03	94.70	88.70	17.58	82.37
	Multi (2019)	86.28	94.76	89.30	26.29	89.86
	SPEC (2020)	82.39	93.12	86.31	15.61	68.82
	Ours	91.92	96.32	93.57	22.65	88.33
SD2	Light Image	80.50	95.89	87.10	11.79	66.42
	Multi (2019)	79.21	93.82	84.88	28.99	91.81
	SPEC (2020)	78.87	95.10	85.55	9.66	53.95
	Ours	83.57	95.00	88.42	29.21	90.67
RD	Light Image	64.85	90.60	73.49	17.05	65.04
	Multi (2019)	61.58	87.63	70.72	17.17	64.23
	SPEC (2020)	70.59	91.62	78.38	14.82	52.49
	Ours	78.50	91.34	83.34	21.62	77.19

表 2. 在我们收集的三个数据集上，我们的方法在没有和有文本相关损失的情况下的性能。

Datasets	Methods	Recall \uparrow	Precision \uparrow	F-measure \uparrow	PSNR \uparrow	SSIM \uparrow
SD1	w/o text loss	91.43	94.12	92.75	21.88	87.19
	Ours	91.92	96.32	93.57	22.65	88.33
SD2	w/o text loss	82.69	93.48	87.76	28.12	89.93
	Ours	83.57	95.00	88.42	29.21	90.67
RD	w/o text loss	77.04	89.66	82.87	21.38	76.11
	Ours	78.50	91.34	83.34	21.62	77.19

6 结论和未来工作

在这项工作中，我们研究并解决了单个文本图像的具有挑战性的镜面高光去除问题。为了促进这项研究，我们收集了三个带有细粒度注释的高质量数据集。我们提出了一个包括高光检测网络和高光去除网络的两阶段框架。高光检测网络的输出用作辅助信息，引导高光去除网络更多地关注高光区域。此外，引入了与文本相关的损失以改善文本的恢复。我们的源代码和数据集可以在<https://github.com/weizequan/TASHR>获取。未来，我们希望构建更大、更丰富的数据集来促进相关研究的发展。我们还想设计更有效的网络和损失函数。此外，一个令人兴奋的研究问题是提出更完整和准确的视觉质量测量。

致谢。该工作得到了国家重点研发计划（2019YFB2204104）和国家自然科学基金（6210071649、62172415和61772523）的支持。

References

1.中文ocr: Ctpn加上densenet加上基于ctc的中文ocr. https://github.com/YCG09/chinese_ocr. 访问日期: 20212 年 4 月 30 日。Paddleocr: 基于paddlepaddle 的出色多语言 ocr 工具包。 <https://github.com/PaddlePaddle/PaddleOCR>. 访问日期: 20213 年 4 月 30 日。Arbeláez, P., Maire, M., Fowlkes, C., Malik, J.: 轮廓检测和分层图像分割。IEEE 传输。模式肛门。马赫。英特尔。 33(5), 898–916(2011)4。Arnold, M., Ghosh, A., Ameling, S., Lacey, G.: 内窥镜成像的镜面高光自动分割和修复。图像视频处理。(2010)5。 Bertalmio, M., Sapiro, G., Caselles, V., Ballester, C.: 图像修复。载于: ACMSIGGRAPH, 第 417–424 页 (2000)

6. Fleyeh, H.: 交通标志的阴影和高光不变颜色分割算法。见: IEEE 控制论和智能系统会议 (2006)7. Fu, G., Zhang, Q., Song, C., Lin, Q., Xiao, C.: 真实世界图像的镜面高光去除。计算。图形。论坛 38(7), 253–263 (2019)8. Funke, I., Bodenstedt, S., Riediger, C., Weitz, J., Speidel, S.: 用于内窥镜图像中镜面高光去除的生成对抗网络。见: 医学成像 2018: 图像引导程序、机器人干预和建模, 卷。10576, 页。8–16 (2018) 9. Gatys, L.A., Ecker, A.S., Bethge, M.: 使用卷积神经网络进行图像风格迁移。见: IEEE 计算机视觉和模式识别会议, 第 10 页。2414–2423 (2016)10. 郭X., 陈Z., 李S., 杨Y., 于J.: 深眼: 双目离焦深度焦点堆栈对。见: 中国模式识别与计算机视觉学术会议, 第353–365页 (2019) 11. Johnson, J., Alahi, A., Fei-Fei, L.: 实时风格迁移和超分辨率的感知损失。见: 欧洲计算机视觉会议论文集, 第 11 页。694–711 (2016) 12. Khanian, M., Boroujerdi, A.S., Breuß, M.: 强镜面高光的光度立体。arXiv 预印本 arXiv: 1709.01357 (2017)13. Kingma, D.P., Ba, J.: Adam: 一种随机优化方法。见: 国际学习表征会议 (2015)14. Lee, S.T., Yoon, T.H., Kim, K.S., Kim, K.D., Park, W.: 通过感知器神经网络去除牙齿颜色图像中的镜面反射。在: 国际信号处理系统会议, 卷。1, 第 V1–285–V1–289 (2010)15 页。Lin, J., El Amine Seddik, M., Tamaazousti, M., Tamaazousti, Y., Bartoli, A.: 深度多类对抗性镜面反射消除。见: 图像分析, 第 3–15 页 (2019)16. Long, S., He, X., Yao, C.: 场景文本检测与识别: 深度学习时代。国际。J. 计算机。维斯。129(1), 161–184 (2021)17. Lucas, S., Panaretos, A., Sosa, L., Tang, A., Wong, S., Young, R.: Icdar 2003 稳健阅读竞赛。见: 国际文档分析与识别会议, 第 682–687 页 (2003)18. Meslouhi, O.E., Kardouchi, M., Allali, H., Gadi, T., Benkaddour, Y.A.: 阴道镜图像镜面反射的自动检测和修复。欧洲中央银行计算。科学。1 (2011) 19. Miyato, T., Kataoka, T., Koyama, M., Yoshida, Y.: 生成对抗网络的谱归一化。见: 国际学习代表会议 (2018) 20. Muhammad, S., Dailey, M.N., Farooq, M., Majeed, M.F., Ekpanyapong, M.: Spec-net 和 spec-cgan: 用于去除面部镜面反射的深度学习模型。图像可见。计算。93, 103823 (2020) 21. Ortiz, F., Torres, F.: 一种通过颜色形态消除高光的新修复方法。见: 国际模式识别和图像分析会议, 第 368–376 页 (2005)22. Park, J.W., Lee, K.H.: 使用彩色线投影修复高光。IEICE翻译。信息。系统。90(1), 250–257 (2007)23. Ren, W., Tian, J., Tang, Y.: 具有色线约束的镜面反射分离。IEEE 传输。图像处理。26(5), 2327–2337 (2017)24. Shafer, S.A.: 使用颜色来分离反射分量。颜色。资源。应用10(4), 210–218 (1985)

10(4), 210–218 (1985)

25. Simonyan, K., Zisserman, A.: 用于大规模图像识别的非常深的卷积网络。见：国际学习表征会议 (2015) 26. Son, M., Lee, Y., Chang, H.S.: 基于统计反射模型从自然图像中消除镜面反射。IEEE 传输。图像处理。(2020) 27. Tan, P., Lin, S., Quan, L., Shum, H.Y.: 通过光照约束修复去除高光。载于：IEEE 国际计算机视觉会议, 第 14 页。164–169 (2003) 28. Tan, R.T., Nishino, K., Ikeuchi, K.: 基于色度和噪声分析分离反射分量。IEEE 传输。模式肛门。马赫。英特尔。26 (10) , 1373–1379 (2004) 29. Tian, Z., Huang, W., He, T., He, P., Qiao, Y.: 使用联结文本提议网络检测自然图像中的文本。见：欧洲计算机视觉会议, 第 56-72 页 (2016) 30. Wang, T.C., Efros, A.A., Ramamoorthi, R.: 使用光场相机进行遮挡感知深度估计。载于：IEEE 国际计算机视觉会议, 第 14 页。3487–3495 (2015) 31. 王文、邓瑞、李丽、徐霞：基于感知一致性的图像审美评价。载于：中国模式识别与计算机视觉学术会议, 第 11 页。303–315 (2019) 32. Yang, Q., Wang, S., Ahuja, N.: 使用双边滤波实时镜面高光去除。见：欧洲计算机视觉会议, 第 87-100 页 (2010) 33. Zhu, J.Y., Park, T., Isola, P., Efros, A.A.: 使用循环一致的对抗网络进行不对应的图像到图像的翻译。见：IEEE 国际计算机视觉会议, 第 2242–2251 页 (2017 年)

Computer Vision, pp. 2242–2251 (2017)