

暗光下的对象检测

1 问题背景

弱光环境在生活中非常常见，除了白天在室外我们拥有足够充足的光源，在其他时间里，我们都处于弱光环境下。我们的眼睛在这种弱光环境中可以通过放大瞳孔，采集更多的光线来适应光线的减少，从而能够在较暗的环境中也能够像在强光环境中一样辨别出图像，而哺乳动物猫则相较于人类具有更强的瞳孔调节能力，哪怕在极弱光线的环境中，人类已经难以分辨对象，猫依然能够通过极为有限的光线看清物体。同样具备相似能力的物理设备是照相机，尽管照相机不具备像人或者猫一样的瞳孔自动调节能力，但是在条件允许的条件下照相机同样可以通过直接调节感光强度，调节焦距，增长感光时间等方式在黑夜中捕捉微弱光源得到亮且清晰的图像，然而由于设备噪声和相机稳定性的问题，大部分的方法都有只能在特定的场景使用，例如调节感光强度会相应增加图像噪声，增加感光时间要求相机稳定且拍摄静态物体[2]，另外这种调节需要人为调控，而且调控往往需要专业的摄影有技巧，所以在大部分情况下我们仅能通过相机获得昏暗的图像。

针对这些昏暗图像已经有一些图像增亮的方法被提出，其中传统的图像增亮方法主要有基于直方图均衡的方法[19, 28]和基于Retinex的方法[7, 20]，另外基于深度学习的图像增亮方法也有许多被提出[4, 16, ?, 11, 24]，这些方法往往以较暗的RGB图像为深度网络的输入，并以指定的较亮的增强了的RGB图像为输出，在这些方法中，数据集以成对出现，相同场景的较暗的图像和较亮的图像配成一对。另外近年来也有直接基于相机传感器的图像增亮的方法被提出[2]，在这种方法中，作者直接提取照相机感光传感器中的光照数据(Raw Data)，然后直接将光照数据乘以相应的亮度增强系数，并将这些数据输入到深度全卷机中得到亮度增强的RGB图像。由于使用了相机传感器数据，这种方法可以从极弱的人眼完全无法分辨的光照信息中，将光照强度增强并进行修复得到良好光照的图像。另外基于相机传感器数据的使用深度卷积神经网络替代传统ISP架构[25]的方法也被提出，由传感器数据转换成RGB数据有更好的表现。

另外在对象检测任务的研究中[15]，尽管随着全球各地研究者的不断研究，COCO[13]，Pascal VOC[3]等主流数据集的mAP指数在不断进步，然而其中的弱光照图像的检测效果依然显著弱于强光照图像的检测效果，甚至弱光照数据完全无效。进一步分析数据集，其中弱光条件的检测图像仅占图像总量的2%，可以说目前的端到端图像检测数据集和方法对弱光环境并没有做专门的考虑，针对弱光环境下的图像的对象检测任务亟待解决。针对这个问题专门针对于弱光环境对象检测的数据集被提出[15]，以供研究者研究。

在调研的过程中我们发现目前已经有很多基于RGB图像的数据增亮方法被提出，这些方法直接作用于RGB数据，并输出相应正常亮度的图像，这些RGB数据只是相对较暗，相对较难检测，肉眼通过细致观察或通过简单的提亮处理就可以准确识别出其中的对象，而针对极暗甚至全黑的图像，由于RGB数据的精度不足，极暗条件下，难以有信息保存，这类图像难以通过RGB图像增亮方法恢复。而光照传感器数据(Raw Data)是光照的原始数据，保存了光照的完备信息，其浮点精度较高，在极暗的光照采集中，保留了相较于RGB更多的信息，对于昏暗以及肉眼完全无法看到的极暗场景，[2]工作表明，Raw Data数据都具备良好的图像恢复效果。

[2]已经表明Raw Data类型数据在亮度增强上相较于RGB数据有得天独厚的优势，其光照

信息完全无损，且精度更高，在极暗的肉眼无法分辨的场景下依然可以有效将图像进行复原，而RGB数据仅能在及其有限的范围内对图像做亮度增强，基于以上观察我们希望对使用Raw Data数据来对极暗环境下的对象检测展开研究，希望利用Raw Data数据的特有优势，来解决RGB数据在弱光条件下表现较差的问题。

另外调查发现，早在2014年已有Raw Data类型的对象检测数据集Pascal Raw发布[18]，针对RGB图像的弱光条件的检测数据集ExDark[15]，但是目前仍然没有弱光条件下的Raw Data对象检测数据，为了克服这个问题，我们直接以清华校园为取材场地，在夜晚对不同光照条件下和不同相机参数条件下的行人，自行车和汽车进行拍照，并提取Raw Data数据，并在后期人工对这些数据集进行边界框标注，构建极弱光条件下的Raw Data对象检测数据集，目前有约700张对象检测Raw Data图像。另外出于亮度增强的研究需要，我们还对清华大学的约100个静态场景提取了不同曝光时间下的夜晚图像，这些图像有完全相同的内容和不同的采集到的光照强度，在必要条件下将用于辅助弱光条件下的Raw Data对象检测。

针对以上的问题和条件，我们最终希望基于我们采集标注的数据构建一个基于RAW Data数据的弱光条件下的对象检测模型。不同于将RGB数据直接输入到目前已有的对象检测网络中去，我们首先将RAW Data数据输入到一个亮度增强网络中，并输出相应的正常亮度的RGB数据，然后将正常亮度的RGB数据输入到对象检测框架中去，以提升对象检测的检测能力。进一步地，我们希望将这两个模型进行改良，并构建融合成一个单独的网络，实现针对基于Raw Data的弱光条件下的端到端对象检测模型。

我们初步计划在本论文中做出以下三点贡献：

1. 我们采集标注了首个弱光条件下的Raw Data对象检测数据集，来弥补极弱光条件下Raw Data类型数据缺失的空白。
2. 我们通过实验验证了Raw Data数据不仅在极暗图像亮度增强方面具有突出的效果，同样在极暗环境下的对象检测同样有效。
3. 我们提出了一个使用弱光Raw Data数据的端到端对象检测框架，实验表明该框架相较于其他对象检测网络对于弱光图像数据具有更好的表现。

2 调研

2.1 RAW数据

相机传感器获取的数据是RAW格式数据，分为6*6的X-Trans矩阵和2*2的Bayer阵列。一般是Bayer阵列格式，由2*2的G/R/B/G组成，将各个通道的像素重新组合为RGB三个通道并通过插值等方法补全像素的过程称为demosaicing。之后在经过相机硬件的ISP架构进行白平衡、去噪、锐化、颜色空间转换、伽马矫正等等转化为RGB格式数据，这是对Raw Data数据处理的传统的pipeline，但这些过程依赖于特定的相机，不同相机的处理不尽相同。Jiang等人[26]提出了局部、线性可学习的L3滤波器近似现在成像系统中复杂的非线性管道。但这些方法由于信噪比SNR较低而并不能很好的处理暗光条件下的成像。Hasinoff等人[8]描述了处理手机相机的burst image成像pipeline，它通过对齐和融合多张RAW图像来获得效果更好的图片。由于Raw Data的精度范围相较于RGB格式更大，没有因为图像处理硬件ISP造成的复杂的噪声影响，相对来说噪声较合成的RGB图像更加纯粹，因此人们思考可以将它用于图像去噪和暗图增强上。比如Chen C[2]等人提出将Raw Data重新排列为四个通道的数据后使用FCN进行图像增强。Eli Schwartz等人[25]提出一个完整的端到端的摄像机图像信号处理管道的深度神经网络模型，学习了从原始低光照图像到最终视觉上效果引人注目的图像的映射，并且实现了低等级任务，如去马赛克和去噪，以及高等级任务，例如颜色校正和图像调整。

2.2 图像去噪

当在低光照环境下拍摄照片时，由于光线的少量进入，拍摄的图片通常非常暗，伴随有大量的噪声，且颜色也无法反映真实世界的颜色。传统的图像去噪方法有最小化全变差、使用小波域处理、核范数极小化等方法[17]，这些方法通常基于特定的图像先验，例如平滑度，稀疏度，低等级或自相似性。研究人员还研究了深度网络在去噪中的应用，包括堆叠稀疏去噪自动编码器(SSDA)，可训练非线性反应扩散(TNRD)，多层感知器，深度自动编码器和卷积网络[17]。当在某些噪声水平上进行训练时，这些数据驱动的方法可以与诸如BM3D和稀疏编码[17]之类经典技术竞争。最近，Tim Brooks等人[1]提出传统的图片去噪通常使用的是合成图像进行训练。由于合成图像与真实图像的差异，在图像处理pipeline中，除去去噪外，其他方面如增益，色彩校正，色调映射经常被忽略了。为了解决这个问题，[1]提出了一种模拟图像处理pipeline每个步骤和反转的方法，对从RAW数据到RGB数据的处理过程进行合理建模，将RGB数据转换为RAW数据后利用U-Net[23]结构进行图像去噪，再通过process过程将RAW数据转换为RGB数据。但该方法在对处理过程建模时对测试数据集的特性依赖比较大，因此也有一定的局限性。Di Zhao等人[27]考虑可以拍摄相同场景的多个帧来增强信噪比，提出一种端到端的循环完全卷积网络(RFCN)来处理极端低光环境下拍摄的图片，将原始照片直接映射到sRGB输出，以产生最佳图像或形成多帧去噪图像序列。

2.3 图像增强

暗光的图像增强一直是图像处理领域待解决的问题。图像增强的基本方法分为三种，分别是基于直方图均衡化、基于Retinex和基于学习的方法。

基于直方图均衡化的方法有自适应的直方图均衡(AHE)[19]和限制对比度的自适应直方图均衡(CLAHE)[28]。AHE是计算那些指定区域的直方图，并用这些值去重新分配图像亮度来改变图像的对比度。这种方法更适合改进图像的局部对比度和增强图像边缘来获得更多细节。但它的问题在于对比度增强的同时，也放大了图像的噪音。CLAHE通过在计算累积分布函数(CDF)前用预先定义的阈值来裁剪直方图以达到限制放大幅度的目的。这是处理图像增强问题的传统方法。

Retinex理论合理地假设了图像的构成。如果将观察者看到的图像看成是一幅带有乘性噪声的图像，那么入射光的分量就是一种乘性的，相对均匀且变换缓慢的噪音。Retinex算法所做的就是合理地估计图像中各个位置的噪音，并除去它。如基于中心Retinex的单尺度Retinex(SSR)[7]类似于自然视觉科学中广泛使用的高斯差分(DOG)函数，它将反射率视为最终增强结果。多尺度Retinex(MSR)[20]可以被认为是几种不同SSR输出的加权和。

近年来卷积神经网络的兴起也为图像增强提供了新的思路。Gharbi M 等人[4]提出的 HDR-Net结合传统的图像处理如双边滤波的方法和神经网络结合，快速高效地进行图像增强。它将输入图片降采样为原来的一半，对小图进行局部特征和全局特征提取并进行融合，进一步得到每一张特征图对应的局部线性变换，并利用双边网格存下这些线性变换的系数。对原图进行像素级处理后得到引导图，根据引导图对双边网格做切分，得到大图中逐像素的变换系数，原图中乘以系数后即可得到增强后的图像。Lore K G等人[16]提出了一种基于深度自动编码器的方法，用于识别来自低光图像的信号特征并自适应地增亮图像，而不会过度放大具有高动态范围的图像中的较亮部分。它训练不同低照度图像信号的特征来实现自适应变亮和去噪，主要是通过非线性暗化和添加高斯噪声的方法来模拟低照度环境，进行图像对比度增强。Chen C等人[2]认为对于非常暗光和非常短曝光时间的图像，之前的图像增强方法并不适用。因此作者提出可以利用RAW数据进行极暗图像的图像增强。对于RAW数据进行重新排列，如将Bayer数据分到四个通道中并进行两倍的下采样，再减去黑色像素后乘以放大系数，通过FCN的去噪网络后，最后网络输出是12通道只有一半分辨率的图像数据，通过一个sub-pixel层可以恢复到原来的分辨率。

它可以实现极暗环境下的图像增强，但也有一些不足。比如放大系数需要手动调节；训练后的只针对特定相机，泛化性能不是很好；对图像处理的速度不足够快、内存消耗大等。Li C等人[11]提出了一种用于弱照明图像增强的卷积神经网络LightenNet，它将弱照明图像作为输入并输出其照明图，随后用于获得基于Retinex模型的增强图像。Ruixing Wang等人[24]提出不是直接学习图像到图像的映射，而是在网络中引入中间照明，通过估计图像到光照的映射来增强曝光不足的图片，并根据各种照明约束和先验知识设计Smooth Loss函数，从而能够在增强结果中恢复清晰的细节，鲜明的对比度和自然色彩。

2.4 对象检测

对象检测也是计算机视觉中的一项基本任务。首先对象检测可以看做一个分类问题，将图像中所有可能的位置获取一定大小的窗口送入分类器进行分类。传统的对象检测方法是使用基于梯度直方图的特征提取方法提取图像特征后送入支持向量机分类器中进行分类。Girshick等人[6]提出第一个将深度学习应用到对象检测领域的区域卷积神经网络(R-CNN)，其中选择性搜索的对象候选算法使用局部提醒生成对象所在的所有可能的位置的搜索框，并将这些框送入卷积神经网络中预测每个小块的类别。更进一步地提出可以先抽特征后再抽取ROI(region of interest)区域，提出ROI Pooling[5]，引入多任务学习，除了ROI区域的选择外是一个端到端的可训练模型，针对bounding box回归设计了更合适的对离群点更敏感的loss函数。He Kaiming等人[10]引入空间金字塔池化，提出SPP-Net，将不同大小的特征图concat后输入FC层。Ren Shaoqing等人[22]提出区域提议网络(Region Proposal Network,RPN)，每个位置有n个不同大小的bounding box，使用CNN产生ROI，从而提出Faster R-CNN。Redmon[21]等人提出YOLO算法将检测看做回归问题，将整张图作为网络的输入，直接在图像的多个位置上回归出这个位置的目标边框。Liu等人提出SSD[14]将 YOLO的回归思想和Faster R-CNN的anchor box机制结合，通过在不同卷积层的特征图上预测物体区域，输出离散化的多尺度、多比例的default boxes坐标，同时利用小卷积核预测一系列候选框的边框坐标补偿和每个类别的置信度。Lin Tsung-Yi等人[12]提出特征金字塔网络(Feature Pyramid Network,FPN)，同时利用低层特征高分辨率和高层特征的高语义信息，通过融合这些不同层的特征达到预测的效果。He Kaiming等人[9]提出的Mask R-CNN中RoIAlign将任意大小的建议区域转化为固定的维度表示，结合语义分割和对象检测任务，通过多任务学习来提升对象检测的效果。

3 数据集

我们使用Canon EOS 70d在清华校园中采集的图像数据，包括RAW数据(.CR2)和RGB 数据(.JPG)。数据集由两部分组成，分别是相同的场景多张不同曝光值的不同光照数据和暗光条件下的对象检测数据，目前分别采集了98个场景共652张不同光照图像和590张暗光条件下的“people”、“car”、“bicycle”数据集，日后还有待继续完善。同时也考虑加入RGB格式的暗光对象检测数据集ExDark(Exclusively Dark dataset)[15]，它包含7363张暗光下的图像和12个对象类别，都是RGB类型的数据(.jpg)。考虑使用将RGB格式数据转换为RAW格式来增加数据。

我们采集的数据集如下：

(1)不同光照数据集：相同的场景多张不同曝光值的图像数据，内容多是静止的雕塑楼房建筑物等。同样的场景固定ISO和焦距，只调整曝光时间，采集多张欠曝光和过曝光的数据。

(2)暗光下的对象检测数据集：夜晚暗光条件下在校园中拍摄的人、车、自行车的图像数据，使用PASCAL VOC[3]相同的标注格式进行标注处理。

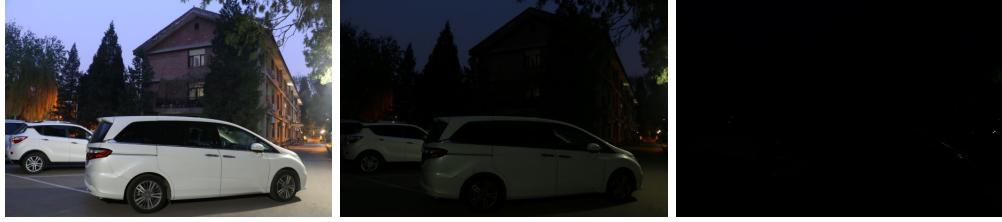


图 1: 我们采集的相同场景不同曝光时间的数据



图 2: 我们采集的暗光下的对象检测数据

4 评估指标

4.1 图像质量评估指标

峰值信噪比(Peak Signal to Noise Ratio,PSNR) 是一种全参考的图像质量评价指标。

$$MSE = \frac{1}{H * W} \sum_{i=1}^H \sum_{j=1}^W (X(i, j) - Y(i, j))^2 \quad (1)$$

$$PSNR = 10 \log_{10} \frac{(2^n - 1)^2}{MSE} \quad (2)$$

其中，MSE表示当前图像X和参考图像Y的均方误差，H、W分别为图像的高度和宽度；n为每像素的比特。PSNR的单位是dB，数值越大表示失真越小。PSNR是最普遍和使用最为广泛的一种图像客观评价指标，然而它是基于对应像素点间的误差，即基于误差敏感的图像质量评价。

结构相似性(structural similarity,SSIM)，也是一种全参考的图像质量评价指标，它分别从亮度、对比度、结构三方面度量图像相似性。SSIM取值范围[0,1]，值越大，表示图像失真越小。

4.2 对象检测评估指标

交并比IoU(Intersection over union):衡量两个区域的重叠程度，是两个区域重叠部分面积占二者总面积(重叠部分只计算一次)的比例，即交叉面积与合并面积之比。

平均精度AP(Average Precision):每个预测结果有一个预测框，预测类别以及置信概率。当预测类别正确且IoU、置信概率都大于一定阈值时，该预测结果被认为预测正确。针对某个分类，固定IoU阈值，将预测结果按置信概率从高到低排列，将每个结果往上当作预测属于该分类，往下当作预测不属于该分类，分别计算precision和recall，画出precision-recall(PR)曲线。将某个查全率r对应的查准率p取为查全率 $r \geq p$ 时最大的查准率，可得到平滑的PR曲线，曲线下的面积即为AP。

多类平均精度mAP(mean Average Precision):是对多个类别准确度的综合度量。VOC[3]数据集计算的是IoU阈值为0.5时各个类别AP的均值，COCO数据集[13]在VOC标准的基础上，取IoU阈值为0.5, 0.55, 0.6, ..., 0.95时各个mAP的均值。

5 初步技术方案

我们在技术实现上分为两个部分，第一个部分是将弱光条件下采集到的Raw Data数据转换为正常光照的RGB数据，在这个技术实现中，我们初步采用[2]中的方法，首先将Raw Data从单个通道转换成4个通道并乘以一个光照强度系数，然后将Raw Data输入到一个U-net[23]网络中，U-net网络负责将Raw Data数据转换为相应的RGB图像，另外我们还考虑将其他的基于RGB数据的亮度增强网络直接移植到Raw Data数据上来，通常的如HDRNet[4]，直接以Raw Data数据为输入，并产生相应的指定亮度的Raw Data或是RGB数据。

第二个部分是将正常亮度的RGB数据输入到目前已有的对象检测网络中去，并评估对象检测网络效果，初步地我们首先选取SSD，Faster RCNN，Yolo[14, 22, 21]等经典网络进行初步实验，在后期我们会扩展实验更多的目前较为先进的网络。

另外在初步目标达成后，我们还会将图像增亮网络和已有的对象检测网络进行合并，并构建新的基于Raw Data的无关光照强度的对象检测网络。

6 实现计划

在实现计划上，网络部分我们将这个过程分成三步走。

首先我们考考虑将图像增亮网络和检测网络分成单独的模块，先通过图像增亮网络将弱光条件下采集到Raw Data数据转换成正常光照强度的RGB数据，然后将RGB数据作为检测网络的输入，进行对象检测。

然后我们考虑将增亮网络和检测网络简单嫁接到一起，成为一个整体网络，然后做网络的联合训练。

最后我们考虑设计一个全新的端到端网络，实现完全的针对若光照条件的Raw Data的端到端对象检测。

在数据采集和标注方面，我们用指定型号单反相机，在清华大学的夜晚环境中，在不同的相机参数下，对行人，自行车，汽车三个对象进行拍摄，来构成弱光检测数据集，另外为了训练图像增亮网络，我们使用相同型号的单反相机，对清华大学的大约100个场景的从0.01s曝光时间到15s曝光时间范围内进行了取景拍摄。

然后我们使用已经训练好的图像增亮网络对数据集进行增亮处理，并对已有的增亮的数据集使用已有的标准对象检测标注软件进行标注。

由于针对Raw Data的图像处理方法较少，我们专门研究了Raw Data的数据形式和处理方法，并专门写Raw Data数据增强库，以应用于未来在Raw Data数据训练过程中的数据增强。

参考文献

- [1] Tim Brooks, Ben Mildenhall, Tianfan Xue, Jiawen Chen, Dillon Sharlet, and Jonathan T. Barron. Unprocessing images for learned raw denoising. *CoRR*, abs/1811.11127, 2018.
- [2] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to see in the dark. *CoRR*, abs/1805.01934, 2018.
- [3] M. Everingham, S. M. A. Eslami, L. Van Gool, C. K. I. Williams, J. Winn, and A. Zisserman. The pascal visual object classes challenge: A retrospective. *International Journal of Computer Vision*, 111(1):98–136, January 2015.

- [4] Michaël Gharbi, Jiawen Chen, Jonathan T Barron, Samuel W Hasinoff, and Frédo Durand. Deep bilateral learning for real-time image enhancement. *ACM Transactions on Graphics (TOG)*, 36(4):118, 2017.
- [5] Ross B. Girshick. Fast R-CNN. *CoRR*, abs/1504.08083, 2015.
- [6] Ross B. Girshick, Jeff Donahue, Trevor Darrell, and Jitendra Malik. Rich feature hierarchies for accurate object detection and semantic segmentation. *CoRR*, abs/1311.2524, 2013.
- [7] Qin Guang. The research on a new image enhancement algorithm based on retinex theory. In Song Lin and Xiong Huang, editors, *Advanced Research on Computer Education, Simulation and Modeling*, pages 336–342, Berlin, Heidelberg, 2011. Springer Berlin Heidelberg.
- [8] Samuel W. Hasinoff, Dillon Sharlet, Ryan Geiss, Andrew Adams, and Marc Levoy. Burst photography for high dynamic range and low-light imaging on mobile cameras. *Acm Transactions on Graphics*, 35(6):1–12, 2016.
- [9] Kaiming He, Georgia Gkioxari, Piotr Dollár, and Ross B. Girshick. Mask R-CNN. *CoRR*, abs/1703.06870, 2017.
- [10] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Spatial pyramid pooling in deep convolutional networks for visual recognition. *CoRR*, abs/1406.4729, 2014.
- [11] Chongyi Li, Jichang Guo, Fatih Porikli, and Yanwei Pang. Lightennet: A convolutional neural network for weakly illuminated image enhancement. *Pattern Recognition Letters*, 104:15 – 22, 2018.
- [12] Tsung-Yi Lin, Piotr Dollár, Ross B. Girshick, Kaiming He, Bharath Hariharan, and Serge J. Belongie. Feature pyramid networks for object detection. *CoRR*, abs/1612.03144, 2016.
- [13] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C Lawrence Zitnick. Microsoft coco: Common objects in context. In *European conference on computer vision*, pages 740–755. Springer, 2014.
- [14] Wei Liu, Dragomir Anguelov, Dumitru Erhan, Christian Szegedy, Scott E. Reed, Cheng-Yang Fu, and Alexander C. Berg. SSD: single shot multibox detector. *CoRR*, abs/1512.02325, 2015.
- [15] Yuen Peng Loh and Chee Seng Chan. Getting to know low-light images with the exclusively dark dataset. *CoRR*, abs/1805.11227, 2018.
- [16] Kin Gwn Lore, Adedotun Akintayo, and Soumik Sarkar. Llnet: A deep autoencoder approach to natural low-light image enhancement. *Pattern Recognition*, 61:650–662, 2017.
- [17] Mukesh C Motwani, Mukesh C Gadiya, Rakhi C Motwani, and Frederick C Harris. Survey of image denoising techniques. In *Proceedings of GSPX*, pages 27–30, 2004.
- [18] Ta David Omid-Zohoor, Alex and Boris Murmann. Pascalraw: Raw image database for object detection. *Stanford Digital Repository*. Available at: <http://purl.stanford.edu/hq050zr7488>, 2015.

- [19] Stephen M Pizer, E Philip Amburn, John D Austin, Robert Cromartie, Ari Geselowitz, Trey Greer, Bart ter Haar Romeny, John B Zimmerman, and Karel Zuiderveld. Adaptive histogram equalization and its variations. *Computer vision, graphics, and image processing*, 39(3):355–368, 1987.
- [20] Zia-ur Rahman, Daniel J Jobson, and Glenn A Woodell. Multi-scale retinex for color image enhancement. In *Proceedings of 3rd IEEE International Conference on Image Processing*, volume 3, pages 1003–1006. IEEE, 1996.
- [21] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*, pages 779–788, 2016.
- [22] Shaoqing Ren, Kaiming He, Ross B. Girshick, and Jian Sun. Faster R-CNN: towards real-time object detection with region proposal networks. *CoRR*, abs/1506.01497, 2015.
- [23] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *CoRR*, abs/1505.04597, 2015.
- [24] Chi-Wing Fu Xiaoyong Shen Wei-Shi Zheng Jiaya Jia Ruixing Wang, Qing Zhang. Under-exposed photo enhancement using deep illumination estimation. *CVPR*, 2019.
- [25] Eli Schwartz, Raja Giryes, and Alex M Bronstein. Deepisp: learning end-to-end image processing pipeline. *arXiv preprint arXiv:1801.06724*, 2018.
- [26] Brian A. Wandell. Learning the image processing pipeline. *IEEE Transactions on Image Processing*, PP(99):1–1, 2017.
- [27] Di Zhao, Lan Ma, Songnan Li, and Dahai Yu. End-to-end denoising of dark burst images using recurrent fully convolutional networks. *CoRR*, abs/1904.07483, 2019.
- [28] John B Zimmerman, Stephen M Pizer, Edward V Staab, J Randolph Perry, William McCartney, and Bradley C Brenton. An evaluation of the effectiveness of adaptive histogram equalization for contrast enhancement. 1987.