Extreme Low-Light Environment-Driven Image Denoising over Permanently Shadowed Lunar Regions with a Physical Noise Model

Ben Moseley
University of Oxford
Oxford, UK
bmoseley@robots.ox.ac.uk

Valentin Bickel
ETH Zurich/ MPS Goettingen
Zurich/ Goettingen, CH/ GER

Ignacio G. López-Francos
NASA Ames Research Center
Moffett Field, CA, USA

Loveneesh Rana
University of Luxembourg
Luxembourg, LUX

Abstract

Recently, learning-based approaches have achieved impressive results in the field of low-light image denoising. Some state of the art approaches employ a rich physical model to generate realistic training data. However, the performance of these approaches ultimately depends on the realism of the physical model, and many works only concentrate on everyday photography. In this work we present a denoising approach for extremely low-light images of permanently shadowed regions (PSRs) on the lunar surface, taken by the Narrow Angle Camera on board the Lunar Reconnaissance Orbiter satellite. Our approach extends existing learning-based approaches by combining a physical noise model of the camera with real noise samples and training image scene selection based on 3D ray tracing to generate realistic training data. We also condition our denoising model on the camera's environmental metadata at the time of image capture (such as the camera's temperature and age), showing that this improves performance. Our quantitative and qualitative results show that our method strongly outperforms the existing calibration routine for the camera and other baselines. Our results could significantly impact lunar science and exploration, for example by aiding the identification of surface water-ice and reducing uncertainty in rover and human traverse planning into PSRs.

1. Introduction

Low-light image denoising is an important field of computer vision. Low-light environments occur in many different applications, such as night-time photography, astronomy and microscopy [4, 7, 50], and the images captured usually have very poor quality. Fundamentally, this is due to the very limited number of photons arriving at the camera, meaning that noise sources such as photon noise and sensor-related noise strongly dominate [7, 22].

Direct approaches for removing these noise sources include increasing the exposure time or by using burst photography and aligning and combining the images during post-processing. However, these approaches have significant downsides; for example longer exposure times can cause blurring and burst photography can still be sensitive to dynamic scenes [18, 20]. Instead, image denoising can be applied, for which there exists a rich literature [3, 13, 11].

An emerging approach is to use learning-based denoising methods on short exposure images [4, 46, 48]. Whilst promising, a disadvantage is that large amounts of labelled training data are often required to prevent overfitting, which can be expensive to collect, or unavailable [4, 50, 1]. Some recent methods alleviate this problem by using physical models to generate synthetic training data [46], but ultimately the generalisation performance depends on the realism of the physical model.

Furthermore, many recent works focus on denoising low-light images of everyday scenes [4, 20, 46, 35]. In this work we present a denoising approach for a specialised set of extremely low-light images; namely those of the Permanently Shadowed Regions (PSRs) on the lunar surface, taken by the Narrow Angle Camera (NAC) on board the Lunar Reconnaissance Orbiter (LRO) satellite [6, 38]. Improving the quality of these images could have significant scientific impact, for example by enabling the detection of science targets and obstacles in PSRs, reducing the uncertainty in rover and human traverse planning and aiding the identification of surface-exposed water-ice [26].

Whilst the NAC is a scientific camera operating in a remote environment, many of its underlying noise sources are the same as those found in other applications. In this work we extend existing learning-based methods in several respects. Firstly, in order to avoid overfitting and improve generalisation performance, we combine a realistic physical noise model of the camera with real noise samples from dark frames to generate realistic training data. We also use 3D ray tracing to select training image scenes which best match the illumination conditions expected in PSRs. Secondly, instead of using the noisy image as the only input to the learned model as is typically done [4, 46, 48], we condition our model on the camera's environmental metadata available at the time of image capture, allowing us to account for the effect of external factors such as camera temperature on the noise. Finally, many recent works focus on images from CMOS sensors, whilst we formulate a physical noise model for a CCD sensor [4, 46, 48, 28, 44].

Our main contributions are summarised as follows:

- We present a novel application of learning-based low-light image denoising for images of PSRs on the lunar surface.

- We extend existing approaches by combining a physical noise model with real noise samples and scene selection based on 3D ray tracing to generate training data; and by conditioning our model on the camera's environmental metadata at the time of image capture.

- Our quantitative and qualitative results show that our method strongly outperforms the existing NAC calibration routine and other baselines.

We name our approach Hyper-Effective Noise Removal U-Net Software, or HORUS[1].

2. Related work

Image denoising is a well-studied area of computer vision [3, 13, 11, 14]. Many different approaches have been proposed, for example those based on spatial-domain filtering [13], total variation [40, 31], sparse representations [10, 29, 9] and transformed-domain filtering [34, 8]. Most algorithms are based on the assumption that signal and noise have different statistical properties which can be exploited to separate them, and therefore crafting a suitable image prior (such as smoothness, sparsity, and self-similarity) is critical [40, 10, 8, 15]. More recently, deep learning based methods have become hugely popular [49, 30, 45, 16, 2, 24, 23, 25]. In these approaches,

deep neural networks are used to learn an implicit representation of signal and noise. Approaches such as DnCNN [49] have shown significant improvements over popular traditional methods such as BM3D [8].

For denoising low-light images, simple traditional methods include histogram equalisation and gamma correction [5], whilst more sophisticated methods include using wavelet transforms [27], retinex-based methods [32, 47, 17] and principle component analysis [41]. Deep learning methods are becoming popular too [48, 46, 28, 50, 4, 36, 35]. For example, [4] proposed an end-to-end network with a U-Net [39] architecture to convert short-exposure images to their corresponding long-exposure images, whilst [48] proposed a frequency-based enhancement scheme.

Whilst promising, a major disadvantage of deep learning methods is that they often require large amounts of training data [1, 4, 50]. An alternative approach is to synthetically generate noisy images, although many works use simple additive Gaussian noise models or similar, leading to poor performance on real images [33]. A potential improvement is to incorporate both synthetic and real images into the training set [16]. For low-light images, [46] focused on improving the realism of their synthetic data by incorporating a comprehensive set of noise sources based on a physical noise model of CMOS sensors and showed that this performed as well as a deep neural network trained on real clean-noisy image pairs.

Specifically for low-light PSR LRO NAC images, [42] attempted to remove noise using a Canny edge detector and a Hough transform with local interpolation. Their method worked sufficiently well for images with relatively high photon counts but failed for low photon counts, i.e., the vast majority of high-latitude PSRs. The current calibration routine of the camera also attempts to remove noise in the images, but is not specifically designed for images of PSRs [38, 19]. We improve on these approaches by using state-of-the-art learning-based low-light denoising techniques.

3. Instrument overview

The LRO has been orbiting the Moon since 2009 at an altitude varying between 20 and 200 km and over its lifetime its NAC has captured over 2 million images of the lunar surface, enough to cover the surface multiple times over [6, 38]. The NAC consists of two nominally identical cameras ("left" and "right") which capture publicly-available[2] 12-bit panchromatic optical images [38, 19]. Each camera consists of a 700 mm focal length telescope, which images onto a Kodak KLI-5001G 5064-pixel CCD line array, providing a 10 μrad instantaneous field of view. The array is oriented perpendicular to the direction of flight of the satellite and 2D images are captured by taking multiple line

---

[1]The name HORUS represents our interest in complementing the existing LRO NAC image processing pipeline, named Integrated Software for Imagers and Spectrometers (ISIS) [12]. Isis also happens to be the mother of Horus in Egyptian mythology.
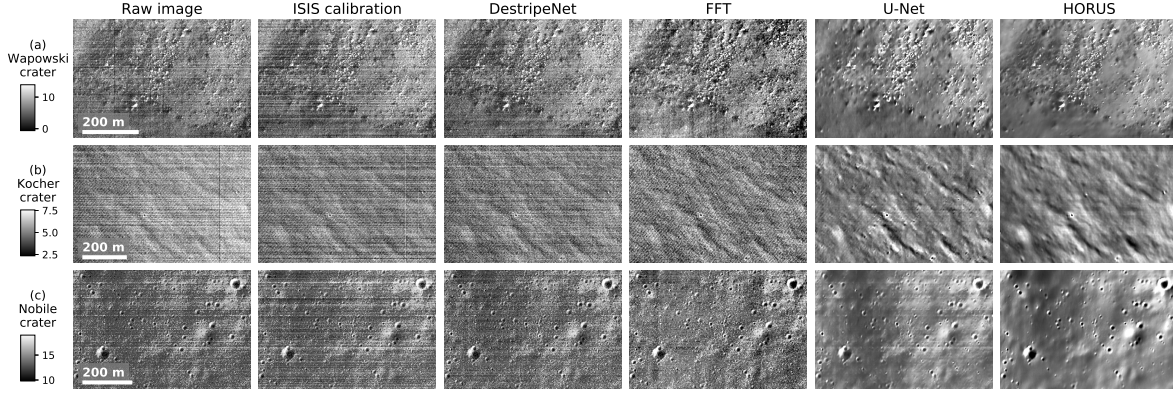
Figure 1. Examples of HORUS applied to 3 real PSR images, compared to the baseline approaches. Grayscale colorbars show the estimated mean photon count S in DN for each HORUS plot. Raw/ISIS image credits to LROC/GSFC/ASU.
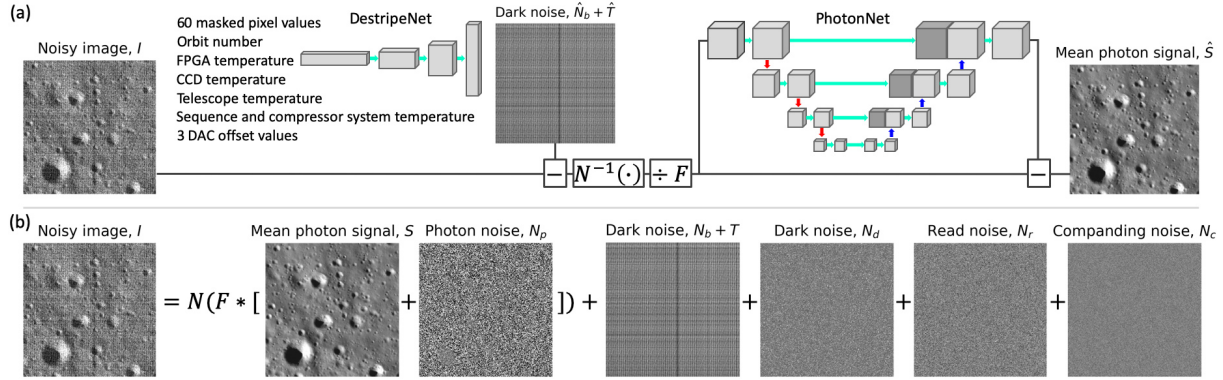

Figure 2. HORUS denoising approach and physical noise model. (a) HORUS denoising workflow. The input to the workflow is the raw, noisy low-lit image I. First a model of the dark bias and dark current is subtracted from the image, which is estimated using a convolutional decoder from the environmental and masked pixel data available at the time of image capture. Next, the inverse nonlinearity and flatfield corrections are applied. Lastly, residual noise sources are estimated and subtracted using a network with a U-Net architecture. (b) Physical noise model. We assume that the raw image I contains companding noise, read noise, dark current and bias noise, photon noise and a nonlinearity and flatfield response. The quantity we wish to recover is the mean photon signal, S. S image credit to LROC/GSFC/ASU.

scans as the spacecraft moves. Because of this motion, the in-track spatial resolution of the images depends on the exposure time as well as the spacecraft altitude and typically both the in-track and cross-track resolution are in the range $0.5 - 2.0$ m. Each camera has two operating modes, "regular" and "summed", where the summed mode typically uses twice the exposure time and sums adjacent pixels during image capture. This is frequently used to maximise the signal received over low-lit regions of the Moon (such as PSRs), at the expense of halving the spatial resolution. For every image the NAC also records over 50 fields of environmental metadata at the time of capture, which for example includes the camera's temperature and orbit number, and the values of 60 masked pixels located on the edges of the CCD.

Of central interest to this study are the NAC images of PSRs. These are located around the lunar poles and whilst they receive no direct sunlight, PSRs can be illuminated by very low levels of indirect sunlight scattered from their surroundings. Figure 1 (column 1) shows example raw NAC images captured over 3 PSRs at the lunar south pole.

4. Physical noise model

In this extremely low-light setting these images are dominated by noise. As part of HORUS, we propose a physical model of this noise. The model is informed by the model developed by [19], who carried out comprehensive characterisation of the instrument before and during its deployment, as well as standard CCD theory [7], and is given by

$$I = N(F * (S + N_p)) + N_b + (T + N_d) + N_r + N_c , \quad (1)$$

where I is the raw image detector count in Data Numbers (DN) recorded by the camera, N is the nonlinearity response, F is the flatfield response, S is the mean photon signal (which is the desirable quantity of interest), N_p is

photon noise, N_b is the dark bias, $T + N_d$ is the dark current noise, N_r is read out noise and N_c is companding noise. A depiction of our noise model is shown in Figure 2 (b). In the following we describe each noise source in greater detail.

The *photon noise* N_p is due to the inherent randomness of photons arriving at the CCD, and obeys

$$(S + N_p) \sim \mathcal{P}(S) \,, \tag{2}$$

where $\mathcal{P}$ denotes a Poisson distribution. The strength of this noise depends on the mean rate of photons hitting the CCD, and it represents a fundamental limitation for any sensor.

The *dark bias* N_b is a deterministic noise source which is due to an artificial voltage offset applied to the CCD to ensure that the Analogue-to-Digital Converter (ADC) always receives a positive signal, and varies pixel-to-pixel, manifesting itself as vertical stripes in the image. Different offsets are commanded depending on the temperature of the NAC, and we also consider the possibility that the bias changes over the camera's lifetime as it degrades.

The *dark current noise* is generated by thermally-activated charge carriers in the CCD which accumulate over the exposure duration and obeys

$$(T + N_d) \sim \mathcal{P}(T) \,, \tag{3}$$

where T is the mean number of thermally-activated charge carriers which depends on the CCD temperature. This noise source varies pixel-to-pixel and with each image line and introduces horizontal and vertical stripes in the image.

The *read out noise* N_r is stochastic system noise introduced during the conversion of charge carriers into an analogue voltage signal and is estimated by [19] to have a standard deviation around 1.15 DN for both cameras.

A perfect camera is expected to have a linear response between the recorded DN value and the mean photon signal S. However during laboratory characterisation [19] showed that the response of the NAC becomes nonlinear for DN counts below 600 DN. They proposed an empirical *nonlinearity correction* to correct for this effect, given by

$$N^{-1}(x) = (x + d) - \frac{1}{ab^{(x+d)} + c} \,, \tag{4}$$

where x is the recorded pixel DN value and $N^{-1}(x)$ is the estimated true DN value. Here a, b, c and d are free parameters which [19] estimated through experimental calibration. In this work we use the average parameter values reported by [19] for each camera, using the same values across all pixels. For computing the forward function $N(x)$ we are not aware of an analytical inverse of Equation 4 and instead use numerical interpolation. The nonlinearity correction curves used are shown in Figure 3 (a).

The *flatfield correction* is used to correct for pixel-to-pixel variations in the sensitivity of the camera, which may be due to the optics or the detector, such as vignetting or
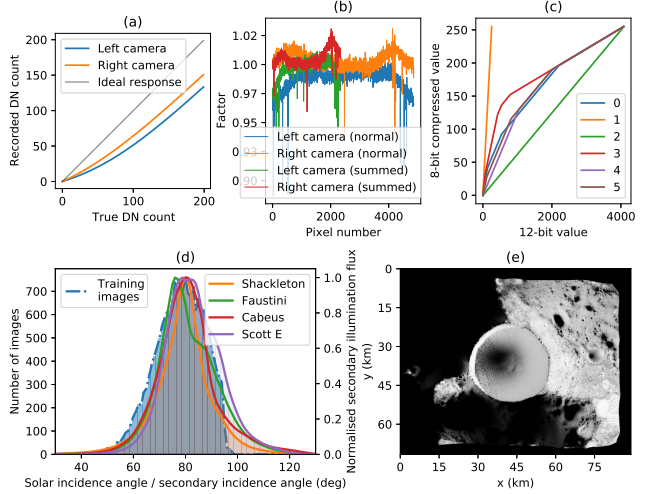

Figure 3. NAC characteristics and training image selection. (a) Nonlinearity response. Separate responses are used for the left and right cameras. (b) Flatfield response. Separate flatfields are used for each camera and operating mode. (c) The six compression schemes commandable for compressing NAC images from 12-bit values to 8-bit values. (d) Distribution of solar incidence angles in the training set for HORUS compared to the distribution of secondary incidence angle of scattered rays using 3D ray tracing over 4 PSRs. (e) Render of the secondary illumination radiance from ray tracing of the Shackleton crater PSR at the lunar south pole (white indicates increased secondary illumination).

particulates on the detector. The forward response is modelled by multiplying the flatfield F, which is a vector of gain values, point-wise with each pixel. [19] experimentally measured separate flatfields for each camera and each operating mode and we use the same, shown in Figure 3 (b).

Finally, the NAC images are compressed (companded) from 12-bits to 8-bits using a lossy compression scheme before downlink to Earth [38]. This introduces *companding noise* and represents a fundamental limit on DN resolution. Six different compression schemes can be commanded; for PSR images scheme 3 in Figure 3 (c) is frequently used, which more accurately reconstructs lower DN values, leading to a maximum absolute reconstruction error after decompression of 2 DN for counts below 424 DN.

Compared to the noise model proposed by [19], our noise model explicitly considers photon noise, stochastic dark current noise, read noise and companding noise.

5. Training data generation

5.1. Hybrid noise generation

To train HORUS we generate a large dataset of clean-noisy image pairs. Noisy images are generated by combining our physical noise model (Equation 1) with noise sampled from real dark calibration frames. An example of a
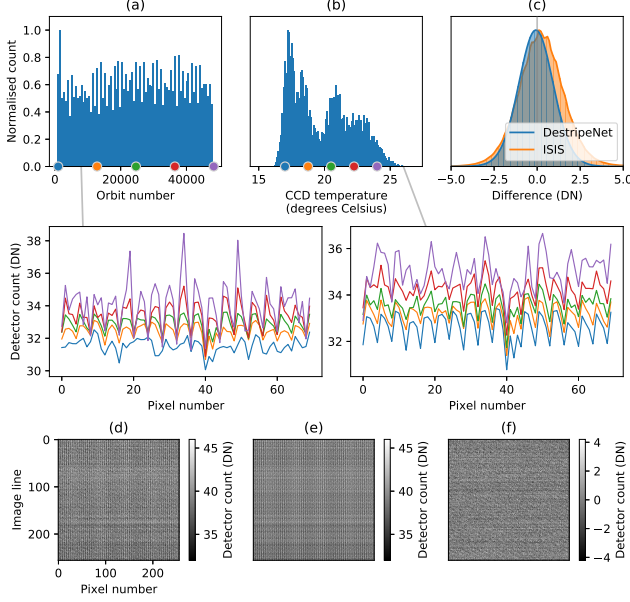
Figure 4. Example dark noise predictions generated by DestripeNet. (a) Histogram of orbit numbers over the dark calibration frames. The linked line plot shows the predicted dark level using DestripeNet for the first 70 pixels of the right camera in summed mode when varying the orbit number and fixing the other metadata inputs to those of a randomly selected dark frame. Color-coded dots in the histogram show the value of the orbit number for each line. (b) Similar set of plots for the CCD temperature. (c) Histograms of the difference between the DestripeNet and ISIS dark frame predictions and the ground truth dark calibration frame over the test set of summed mode images. (d) Example portion of a real dark calibration frame. (e) Corresponding DestripeNet dark level prediction. (f) Difference between (d) and (e).

real dark calibration frame is shown in Figure 4 (d). These are captured by the NAC whenever the LRO passes over the non-illuminated side of the Moon and nominally receive no sunlight. We therefore assume that they sample the following terms in our physical noise model[3],

$$D \simeq N_b + (T + N_d) + N_r . \tag{5}$$

Over 70,000 dark calibration frames have been captured, which cover the entire lifetime of the instrument over a wide range of camera states and environmental conditions; the histogram of orbit numbers and CCD temperatures at the time of image capture over the entire set are shown in Figure 4. We assume that these frames are a representative and sufficient sample of the dark noise distribution of the camera.

Given an input clean image representing the mean photon signal $\tilde{S}$, we generate its corresponding noisy image $\tilde{I}$

---

using the following equation,

$$\tilde{I} = N(F * (\tilde{S} + N_p)) + D + N_c , \tag{6}$$

where D is a randomly selected dark noise frame, N_p is synthetic, randomly generated Poisson noise and N_c is the noise generated by compressing and decompressing the noisy image through the NAC companding scheme 3. By using a hybrid noise model, Equation 6 does not entirely rely on the assumptions of our physical noise model and we ensure that the distribution of noise in our training data is as close as possible to the real noise distribution.

5.2. Image preprocessing and selection based on 3D ray tracing

Clean images $\tilde{S}$ are generated by extracting millions of randomly selected 256×256 image patches cropped from NAC images of the lunar surface in normal sunlit conditions. We define a sunlit image as any image with a median DN > 200 and our assumption is that in this regime the noise sources in Equation 1 become negligible, such that we can use sunlit images as a proxy for S. An example sunlit image is shown in Figure 2 (b).

An important difference between sunlit images and the test-time images of PSRs are the illumination conditions; whilst sunlit images are illuminated by direct sunlight, PSRs are only illuminated by secondary (backscattered) light. In order to match the illumination conditions of the training data to the test-time data as best we can, we select sunlit images with similar solar incidence angles to the secondary illumination incidence angles expected over PSRs. 3D ray tracing is performed over 4 example PSRs using 30 m/pixel spatial resolution LRO Lunar Orbiter Laser Altimeter elevation data [37] and a Lambertian bidirectional reflection function. The solar incidence angle is available as metadata for each sunlit image, and we match the distribution of solar incidence angles in our training images to the distribution of secondary illumination angles from ray tracing, shown in Figure 3 (d).

Before using the sunlit images as training data we rescale their DN values to those expected for PSR images. Each image patch is divided by its median DN value and multiplied by a random number drawn from a uniform distribution over the range 0-60 DN, which is what we expect for typical PSRs. Images are sampled from the entire lunar surface with no preference on their location and in total over 1.5 million image patches $\tilde{S}$ are extracted for each camera in each operating mode.

6. Experiments

6.1. Denoising workflow

The workflow HORUS uses to denoise images is shown in Figure 2 (a). The input is the noisy low-lit image, I, af-

		ISIS	DestripeNet	FFT	U-Net	HORUS
Mode	**Camera**			L1 / PSNR / SSIM		
Normal	Left	5.68 / 31.12 / 0.61	5.58 / 31.33 / 0.62	4.93 / 33.38 / 0.80	1.64 / 42.50 / 0.96	**1.30 / 44.63 / 0.97**
	Right	5.09 / 32.23 / 0.67	4.99 / 32.44 / 0.68	4.61 / 34.20 / 0.83	1.43 / 44.08 / 0.96	**1.23 / 45.35 / 0.97**
Summed	Left	5.45 / 31.51 / 0.65	5.44 / 31.59 / 0.65	4.50 / 34.04 / 0.83	1.69 / 42.34 / 0.95	**1.52 / 43.33 / 0.96**
	Right	4.96 / 32.48 / 0.70	4.92 / 32.59 / 0.70	4.21 / 34.82 / 0.85	1.60 / 42.93 / **0.96**	**1.43 / 44.02 / 0.96**

Table 1. Synthetic test set performance of HORUS, compared to the baseline approaches. All metrics compare the ground truth image $\tilde{S}$ to the estimated denoised image and the average performance over all images is reported. Higher is better, apart from L1 error.
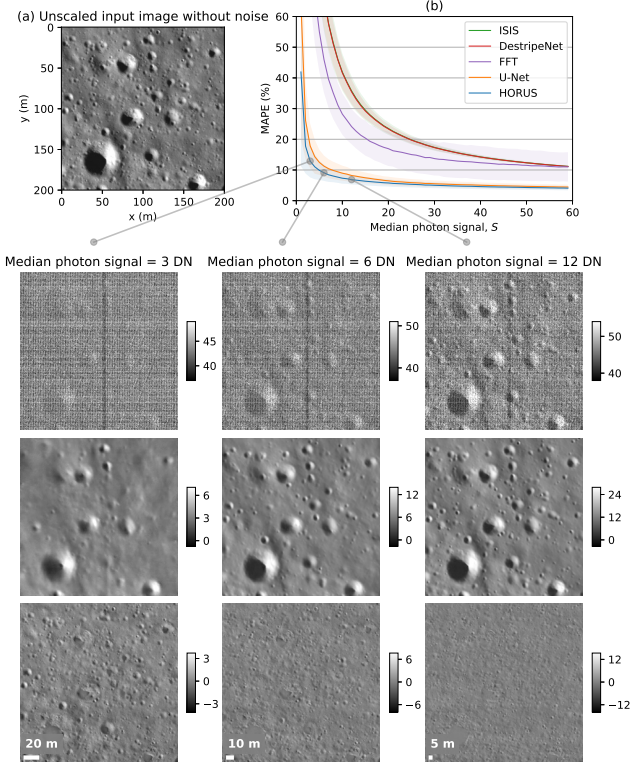

Figure 5. Synthetic test set performance of HORUS with varying signal strengths. (a) Unscaled ground truth image S used below. (b) Mean absolute percentage error between the ground truth image and the denoised image, binned by median DN value of the ground truth image, for HORUS and the baselines over the synthetic test set of summed mode images. Filled regions show ±1 standard deviation. The image grid shows the ground truth image from (a) scaled to different median photon counts with noise added according to our physical model (top row), HORUS denoising of this image (middle row) and the difference between HORUS and the ground truth image (bottom row). Colorbar shows DN counts. Raw image credits to LROC/GSFC/ASU.

ter decompression. First a convolutional decoder called DestripeNet is used to predict the dark bias and mean dark current, given a set of environmental metadata and the masked pixel values at the time of image capture, which is sub-tracted from the input image. Next, the inverse nonlinearity and flatfield corrections are applied. Lastly, a network with a U-Net [39] architecture called PhotonNet is used to estimate the residual noise sources in the image, which are subtracted from the image.

The input to DestripeNet is a vector of 8 selected environment metadata fields available at the time of image capture (listed in Figure 2 (a)), concatenated with the masked pixel values. Our assumption is that this vector is sufficient to predict the dark bias and mean dark current for an image. DestripeNet is trained using the real dark calibration images as labels and a L2 loss function. Under a L2 loss the network learns to predict the mean of its output variable conditioned on its inputs [43], such that, under the assumptions of Equation 5, it estimates the quantity $(N_b + T)$.

Given the DestripeNet prediction $(\hat{N}_b + \hat{T})$, the input to PhotonNet is

$$J = N^{-1}(I - (\hat{N}_b + \hat{T}))/F . \qquad (7)$$

We train PhotonNet using the synthetic image pairs $(\tilde{I}, \tilde{S})$ described in Sections 5.1 and 5.2. Before inputting the noisy training images $\tilde{I}$ into the network, Equation 7 is applied using their DestripeNet prediction, and we use $\tilde{J} - \tilde{S}$ as training labels. Thus, PhotonNet learns to estimate the (transformed) residual noise sources in the image, namely the photon noise, stochastic dark current noise, read noise, companding noise, and any residual dark bias and mean dark current noise which DestripeNet failed to predict.

DestripeNet is trained to predict each image line separately and is comprised of 13 1D transposed convolutional layers with a filter length and stride of 2 and ReLU activation functions. The number of hidden channels starts constant at 512 and then halves every layer from 512 to 16 in the last layer. Each metadata field is independently normalised before input. PhotonNet operates on 2D image patches and uses a standard U-Net architecture [39] with 4 downsampling steps and 1 convolutional layer per scale with LeakyReLU activation functions with a negative slope of 0.1. The number of hidden channels doubles after every downsampling step from 32 to 512. DestripeNet is trained before PhotonNet and its predictions used for training PhotonNet are pre-computed. Both networks are trained using

the Adam stochastic gradient descent algorithm [21] with a L2 loss function, using a batch size of 400 image lines for DestripeNet and 10 image patches for PhotonNet. Learning rates of 1×10^{-5} and 1×10^{-4} are used for DestripeNet and PhotonNet respectively. We divide the synthetic image dataset $(\tilde{I}, \tilde{S})$ into a training, validation and test set (72:18:10). Finally, to account for possible differences in the two cameras and their two operating modes, separate networks are trained for each possible configuration.

6.2. Baselines

We compare our approach to a number of baselines;

Current NAC calibration routine (ISIS). We use the Integrated Software for Imagers and Spectrometers (ISIS) [12], which is the current calibration routine for NAC images. The calibration removes an estimate of the dark bias and mean dark current noise by fitting a cyclic trend to the masked pixel values, and also applies the nonlinearity and flatfield corrections [19].

DestripeNet only. For this case, we only subtract the DestripeNet prediction from the image and then apply the nonlinearity and flatfield correction, i.e. PhotonNet is not used.

Fast Fourier Transform (FFT) filtering. We develop a hand-crafted FFT denoising algorithm. The 2D image is transformed into the frequency domain and its zero-frequency components are replaced with the mean of their nearest neighbours to remove horizontal and vertical dark noise stripes in the image. A mild low-pass filter using a 2D Gaussian kernel with a standard deviation of 0.25 m^{-1} and the nonlinearity and flatfield corrections are also applied.

End-to-end U-Net. Instead of two separate networks, we train a single U-Net to directly estimate $\tilde{S}$ given $\tilde{I}$ as input, without using any metadata. This end-to-end strategy is typical in many existing works [4, 28, 46]. The U-Net has the same architecture and training scheme as PhotonNet.

All relevant baselines are trained on and/or have their hyperparameters selected using the same training data as HORUS, and all are tested on the same data as HORUS.

7. Results

7.1. Results on synthetic images

Table 1 shows the quantitative performance of HORUS compared to the baselines across the synthetic test set of images. We find that HORUS gives the strongest performance across all metrics for this dataset, and significantly outperforms the ISIS, DestripeNet-only and FFT baselines.

Example HORUS denoising of a synthetic test image with varying mean photon counts are shown in Figure 5. We also plot the mean absolute percentage error between the ground truth image $\tilde{S}$ and the denoised image, binned by median DN value of the ground truth image, for HORUS and the baselines over the test set of summed mode

images. We find that the performance of all the methods degrades with lowering photon counts, however HORUS gives the lowest error across all DN values. The example denoised images suggest that HORUS can denoise synthetic images with median photon counts as low as 3-6 DN. The difference plots suggest that the minimum feature size HORUS can reliably resolve correlates strongly with the photon count, and for median photon counts of $\sim$12 DN this is $\sim$5 m.

7.2. DestripeNet results

The histogram of the error between the DestripeNet prediction and the real dark calibration frame over the calibration frame test set is shown in Figure 4 (c). We find that DestripeNet is typically able to reconstruct the dark frames to within 1 DN. We plot scans of its prediction over the orbit number and CCD temperature whilst keeping the other metadata inputs fixed. We find that these predictions are physically interpretable: increasing the CCD temperature increases the dark DN level, which is expected, and the variance of the prediction increases with orbit number, possibly indicating a degradation of the instrument over time.

7.3. Results on real images

Qualitative examples of HORUS and the baselines applied to 3 real images of different PSRs are shown in Figure 1. We find that HORUS gives the most convincing results on these images; compared to the ISIS, DestripeNet-only and FFT baselines it removes much more of the high frequency stochastic noise in the images, and compared to the U-Net it removes more of the residual dark noise stripes in the image and introduces less low-frequency noise.

Ground truth qualitative verification. A limitation in this setting is that we do not have real image pairs of noisy and clean images to quantitatively test performance. Instead, we verify our work by using Temporary Shadowed Regions (TSRs). These are regions which transition between being sunlit and shadowed over the course of a lunar day. We compare HORUS denoised shadowed images of TSRs to their raw sunlit images in Figure 6, using the sunlit images as ground truth. For images with sufficient photon counts (e.g. Wapowski crater in Figure 6) HORUS is able to resolve the vast majority of topographic features, e.g., impact craters as small as $\sim$4 m across. For images with lower photon counts (e.g. Kocher crater in Figure 6) HORUS is able to resolve most large-scale topographic features, although higher frequency details are sometimes lost. In sharp contrast, the raw input image is affected by intense noise, making any meaningful observations difficult. Furthermore, we do not observe any hallucinations in the HORUS images.

Overlapping image qualitative verification. We perform a second analysis to verify performance, by comparing three
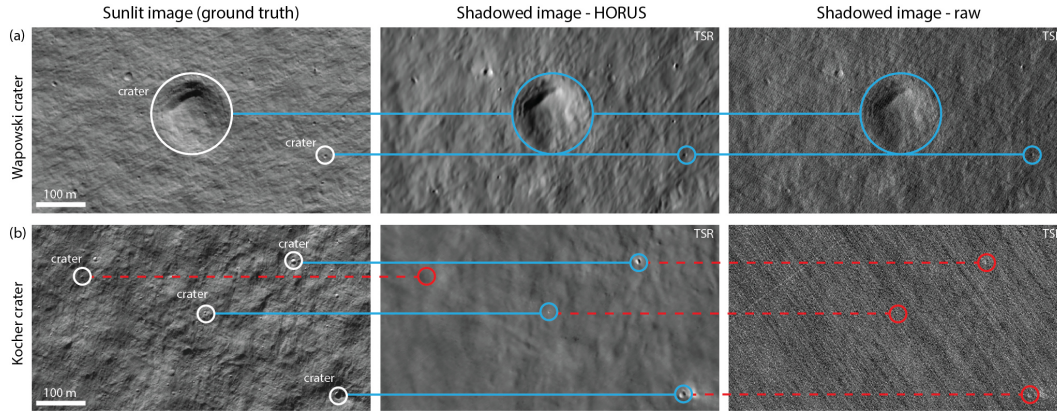
Figure 6. Qualitative verification of HORUS using map-projected sunlit (left) and shadowed image pairs in (a) Wapowski crater and (b) Kocher crater; HORUS-denoised frames (center) are compared with their raw input frames (right). Some features are resolved in both the HORUS and raw frames (blue marks), but HORUS resolves a significantly larger number of features. With decreasing photon counts (example (b)), HORUS struggles to resolve smaller features (here less than ∼10 m across, red marks), resulting in a smoother image. Differences in shadowing are caused by different sunlit/secondary illumination incidence angles. Raw image credits to LROC/GSFC/ASU.
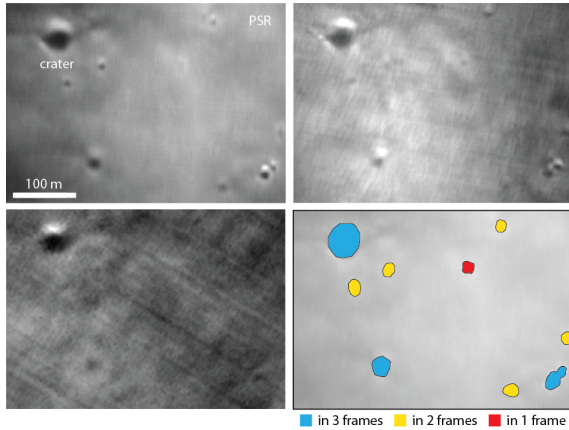

Figure 7. In-PSR qualitative verification of 3 overlapping map-projected HORUS frames with varying levels of signal (decreasing clockwise from top left to bottom left). Some features are present in all frames (blue), some in 2 (yellow), and some only in 1 (red).

overlapping HORUS images taken over a PSR at the lunar south pole (Figure 7). Here we do not have definitive ground truth as all images are shadowed, but we can analyse whether topographic features appear consistent throughout all three HORUS frames. As illustrated in Figure 7, we can trace topographic features through all three frames. However, the size of resolvable features increases with decreasing photon counts: The smallest feature we can identify in the highest photon count frame (top left) is ∼7 m across, while the smallest feature in the lowest photon count frame (bottom left) is ∼14 m across. This is consistent with the synthetic observations in Figure 5.

8. Conclusion

We have presented a learning-based extreme low-light denoising method for improving the image quality of permanently shadowed regions at the lunar poles. Our work is novel in several respects; we are the first to apply learning-based denoising in this setting and we built on state-of-the-art approaches by combining a realistic physical noise model with real noise samples and scene selection based on 3D ray tracing to generate realistic training data. We also conditioned our learned model on the camera's environmental metadata at the time of image capture. Our quantitative and qualitative results showed that our method strongly outperformed the existing calibration routine of the camera and other baselines. Future work will look at quantitatively assessing HORUS using downstream tasks on real images such as crater counting and elevation modelling. Further refinement of our training image scene selection to specific tasks could also be studied; for example by training our networks to image transition zones between PSRs and sunlit regions, which are particularly relevant for future exploration missions.

References

[1] Abdelrahman Abdelhamed, Stephen Lin, and Michael S Brown. A High-Quality Denoising Dataset for Smartphone Cameras. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 1692–1700, 2018. 1, 2

[2] Abdelrahman Abdelhamed, Radu Timofte, Michael S Brown, Songhyun Yu, Bumjun Park, Jechang Jeong, Seung-Won Jung, Dong-Wook Kim, Jae-Ryun Chung, Jiaming Liu, Yuzhi Wang, Chi-Hao Wu, Qin Xu, Yuqian Zhou, Chuan Wang, Shaofan Cai, Yifan Ding, Haoqiang Fan, Jue Wang, Kai Zhang, Wangmeng Zuo, Magauiya Zhussip, Dong Won, Park Shakarim, Soltanayev Se, Young Chun, Zhiwei Xiong, Chang Chen Muhammad, Haris Kazutoshi, Akita Tomoki, Yoshida Greg, Shakhnarovich Norimichi, Ukita Syed, Waqas Zamir, Aditya Arora, Salman Khan Fahad, Shahbaz Khan, Ling Shao, Sung-Jea Ko, Dong-Pan Lim, Seung-Wook Kim, Seo-Won Ji, Sang-Won Lee, Wenyi Tang, Yuchen Fan, Ding Liu, Thomas S Huang, Deyu Meng, Lei Zhang, Hongwei Yong, Yiyun Zhao, Pengliang Tang, Yue Lu, Raimondo Schettini, Simone Bianco, Simone Zini, Chi Li, Yang Wang, and Zhiguo Cao. NTIRE 2019 Challenge on Real Image Denoising: Methods and Results. In *CVPR 2019*, 2019. 2

[3] A. Buades, B. Coll, and J. M. Morel. A review of image denoising algorithms, with a new one. *Multiscale Modeling and Simulation*, 4(2):490–530, jul 2005. 1, 2

[4] Chen Chen, Qifeng Chen, Jia Xu, and Vladlen Koltun. Learning to See in the Dark. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 3291–3300, 2018. 1, 2, 7

[5] H. D. Cheng and X. J. Shi. A simple and effective histogram equalization approach to image enhancement. *Digital Signal Processing: A Review Journal*, 14(2):158–170, 2004. 2

[6] Gordon Chin, Scott Brylow, Marc Foote, James Garvin, Justin Kasper, John Keller, Maxim Litvak, Igor Mitrofanov, David Paige, Keith Raney, Mark Robinson, Anton Sanin, David Smith, Harlan Spence, Paul Spudis, S. Alan Stern, and Maria Zuber. Lunar reconnaissance orbiter overview: The instrument suite and mission. *Space Science Reviews*, 129(4):391–419, apr 2007. 1, 2

[7] Frederick R. Chromey. *To Measure the Sky*. Cambridge University Press, oct 2016. 1, 3

[8] Kostadin Dabov, Alessandro Foi, Vladimir Katkovnik, and Karen Egiazarian. Image denoising by sparse 3-D transform-domain collaborative filtering. *IEEE Transactions on Image Processing*, 16(8):2080–2095, aug 2007. 2

[9] Weisheng Dong, Xin Li, Lei Zhang, and Guangming Shi. Sparsity-based image denoising via dictionary learning and structural clustering. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 457–464. IEEE Computer Society, 2011. 2

[10] Michael Elad and Michal Aharon. Image denoising via sparse and redundant representations over learned dictionaries. *IEEE Transactions on Image Processing*, 15(12):3736–3745, dec 2006. 2

[11] Linwei Fan, Fan Zhang, Hui Fan, and Caiming Zhang. Brief review of image denoising techniques. *Visual Computing for Industry, Biomedicine, and Art*, 2(1):7, dec 2019. 1, 2

[12] L. Gaddis, J. Anderson, K. Becker, T. Becker, D. Cook, K. Edwards, E. Eliason, T. Hare, H. Kieffer, E. M. Lee, J. Mathews, L. Soderblom, T. Sucharski, J. Torson, A. McEwen, and M. Robinson. An Overview of the Integrated Software for Imaging Spectrometers (ISIS). In *Lunar and Planetary Science Conference*, Lunar and Planetary Science Conference, page 387, Mar. 1997. 2, 7

[13] Bhawna Goyal, Ayush Dogra, Sunil Agrawal, B. S. Sohi, and Apoorav Sharma. Image denoising review: From classical to state-of-the-art approaches. *Information Fusion*, 55:220–244, mar 2020. 1, 2

[14] Shuhang Gu and Radu Timofte. A Brief Review of Image Denoising Algorithms and Beyond. In *Part of the The Springer Series on Challenges in Machine Learning book series (SSCML)*, pages 1–21. Springer, Cham, 2019. 2

[15] Shuhang Gu, Lei Zhang, Wangmeng Zuo, and Xiangchu Feng. Weighted nuclear norm minimization with application to image denoising. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 2862–2869. IEEE Computer Society, sep 2014. 2

[16] Shi Guo, Zifei Yan, Kai Zhang, Wangmeng Zuo, and Lei Zhang. Toward convolutional blind denoising of real photographs. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2019-June, pages 1712–1722. IEEE Computer Society, jul 2019. 2

[17] Xiaojie Guo, Yu Li, and Haibin Ling. LIME: Low-light image enhancement via illumination map estimation. *IEEE Transactions on Image Processing*, 26(2):982–993, feb 2017. 2

[18] Samuel W. Hasinoff, Jonathan T. Barron, Dillon Sharlet, Ryan Geiss, Florian Kainz, Jiawen Chen, 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, nov 2016. 1

[19] D. C. Humm, M. Tschimmel, S. M. Brylow, P. Mahanti, T. N. Tran, S. E. Braden, S. Wiseman, J. Danton, E. M. Eliason, and M. S. Robinson. Flight Calibration of the LROC Narrow Angle Camera. *Space Science Reviews*, 200(1-4):431–473, 2016. 2, 3, 4, 7

[20] Ahmet Serdar Karadeniz, Erkut Erdem, and Aykut Erdem. Burst Photography for Learning to Enhance Extremely Dark Images. Technical report, jun 2020. 1

[21] Diederik P. Kingma and Jimmy Ba. Adam: A Method for Stochastic Optimization. *ArXiv e-prints*, dec 2014. 7

[22] Mikhail Konnik and James Welsh. High-level numerical simulations of noise in CCD and CMOS photosensors: review and tutorial. *ArXiv e-prints*, dec 2014. 1

[23] Alexander Krull, Tim Oliver Buchholz, and Florian Jug. Noise2void-Learning denoising from single noisy images. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, volume 2019-June, pages 2124–2132, nov 2019. 2

[24] Jaakko Lehtinen, Jacob Munkberg, Jon Hasselgren, Samuli Laine, Tero Karras, Miika Aittala, and Timo Aila. Noise2Noise: Learning image restoration without clean data. In *35th International Conference on Machine Learning, ICML 2018*, volume 7, pages 4620–4631. International Machine Learning Society (IMLS), mar 2018. 2

[25] Victor Lempitsky, Andrea Vedaldi, and Dmitry Ulyanov. Deep Image Prior. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, pages 9446–9454, 2018. 2

[26] Shuai Li, Paul G. Lucey, Ralph E. Milliken, Paul O. Hayne, Elizabeth Fisher, Jean-Pierre Williams, Dana M. Hurley, and Richard C. Elphic. Direct evidence of surface exposed water ice in the lunar polar regions. *Proceedings of the National Academy of Sciences*, 115(36):8907–8912, 2018. 1

[27] Artur Łoza, David R. Bull, Paul R. Hill, and Alin M. Achim. Automatic contrast enhancement of low-light images based on local statistics of wavelet coefficients. *Digital Signal Processing: A Review Journal*, 23(6):1856–1866, dec 2013. 2

[28] Paras Maharjan, Li Li, Zhu Li, Ning Xu, Chongyang Ma, and Yue Li. Improving extreme low-light image denoising via residual learning. In *Proceedings - IEEE International Conference on Multimedia and Expo*, volume 2019-July, pages 916–921. IEEE Computer Society, jul 2019. 2, 7

[29] Julien Mairal, Francis Bach, Jean Ponce, Guillermo Sapiro, and Andrew Zisserman. Non-local sparse models for image restoration. In *Proceedings of the IEEE International Conference on Computer Vision*, pages 2272–2279, 2009. 2

[30] Xiao Jiao Mao, Chunhua Shen, and Yu Bin Yang. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. In *Advances in Neural Information Processing Systems*, pages 2810–2818, mar 2016. 2

[31] Stanley Osher, Martin Burger, Donald Goldfarb, Jinjun Xu, and Wotao Yin. An iterative regularization method for total variation-based image restoration. *Multiscale Modeling and Simulation*, 4(2):460–489, jul 2005. 2

[32] Seonhee Park, Soohwan Yu, Byeongho Moon, Seungyong Ko, and Joonki Paik. Low-light image enhancement using variational optimization-based retinex model. *IEEE Transactions on Consumer Electronics*, 63(2):178–184, may 2017. 2

[33] Tobias Plötz and Stefan Roth. Benchmarking denoising algorithms with real photographs. In *Proceedings - 30th IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017*, volume 2017-Janua, pages 2750–2759. Institute of Electrical and Electronics Engineers Inc., jul 2017. 2

[34] Javier Portilla, Vasily Strela, Martin J. Wainwright, and Eero P. Simoncelli. Image denoising using scale mixtures of Gaussians in the wavelet domain. *IEEE Transactions on Image Processing*, 12(11):1338–1351, nov 2003. 2

[35] Tal Remez, Or Litany, Raja Giryes, and Alex M. Bronstein. Deep Convolutional Denoising of Low-Light Images. *ArXiv e-prints*, jan 2017. 1, 2

[36] Wenqi Ren, Sifei Liu, Lin Ma, Qianqian Xu, Xiangyu Xu, Xiaochun Cao, Junping Du, and Ming Hsuan Yang. Low-Light Image Enhancement via a Deep Hybrid Network. *IEEE Transactions on Image Processing*, 28(9):4364–4375, sep 2019. 2

[37] H. Riris, G. Neuman, J. Cavanaugh, X. Sun, P. Liiva, and M. Rodriguez. The Lunar Orbiter Laser Altimeter (LOLA) on NASA's Lunar Reconnaissance Orbiter (LRO) mission. In Naoto Kadowaki, editor, *International Conference on Space Optics — ICSO 2010*, volume 10565, page 77. SPIE, nov 2017. 5

[38] M. S. Robinson, S. M. Brylow, M. Tschimmel, D. Humm, S. J. Lawrence, P. C. Thomas, B. W. Denevi, E. Bowman-Cisneros, J. Zerr, M. A. Ravine, M. A. Caplinger, F. T. Ghaemi, J. A. Schaffner, M. C. Malin, P. Mahanti, A. Bartels, J. Anderson, T. N. Tran, E. M. Eliason, A. S. McEwen, E. Turtle, B. L. Jolliff, and H. Hiesinger. Lunar reconnaissance orbiter camera (LROC) instrument overview. *Space Science Reviews*, 150(1-4):81–124, jan 2010. 1, 2, 4

[39] Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. *ArXiv e-prints*, may 2015. 2, 6

[40] Leonid I. Rudin, Stanley Osher, and Emad Fatemi. Nonlinear total variation based noise removal algorithms. *Physica D: Nonlinear Phenomena*, 60(1-4):259–268, nov 1992. 2

[41] Joseph Salmon, Zachary Harmany, Charles Alban Deledalle, and Rebecca Willett. Poisson noise reduction with non-local PCA. *Journal of Mathematical Imaging and Vision*, 48(2):279–294, apr 2014. 2

[42] H. M. Sargeant, V. T. Bickel, C. I. Honniball, S. N. Martinez, A. Rogaski, S. K. Bell, E. C. Czaplinski, B. E. Farrant, E. M. Harrington, G. D. Tolometti, and D. A. Kring. Using Boulder Tracks as a Tool to Understand the Bearing Capacity of Permanently Shadowed Regions of the Moon. *Journal of Geophysical Research: Planets*, 125(2), feb 2020. 2

[43] Padhraic Smyth. On Loss Functions Which Minimize to Conditional Expected Values and Posterior Probabilities. *IEEE Transactions on Information Theory*, 39(4):1404–1408, jul 1993. 6

[44] Richard Szeliski. *Computer Vision: Algorithms and Applications*. Springer-Verlag, Berlin, Heidelberg, 1st edition, 2010. 2

[45] Ying Tai, Jian Yang, Xiaoming Liu, and Chunyan Xu. MemNet: A Persistent Memory Network for Image Restoration. In *Proceedings of the IEEE International Conference on Computer Vision*, volume 2017-Octob, pages 4549–4557. Institute of Electrical and Electronics Engineers Inc., dec 2017. 2

[46] Kaixuan Wei, Ying Fu, Jiaolong Yang, and Hua Huang. A Physics-based Noise Formation Model for Extreme Low-light Raw Denoising. In *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2020. 1, 2, 7

[47] Jun Xu, Yingkun Hou, Dongwei Ren, Li Liu, Fan Zhu, Mengyang Yu, Haoqian Wang, and Ling Shao. STAR: A Structure and Texture Aware Retinex Model. *IEEE Transactions on Image Processing*, 29:5022–5037, jun 2020. 2

[48] Ke Xu, Xin Yang, Baocai Yin, and Rynson W.H. Lau. Learning to restore low-light images via decomposition-and-enhancement. In *Proceedings of the IEEE Computer Soci-*

ety Conference on Computer Vision and Pattern Recognition, pages 2278–2287, 2020. 1, 2

[49] Kai Zhang, Wangmeng Zuo, Yunjin Chen, Deyu Meng, and Lei Zhang. Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Transactions on Image Processing*, 26(7):3142–3155, aug 2017. 2

[50] Yide Zhang, Yinhao Zhu, Evan Nichols, Qingfei Wang, Siyuan Zhang, Cody Smith, and Scott Howard. A poisson-gaussian denoising dataset with real fluorescence microscopy images. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2019-June:11702–11710, 2019. 1, 2