Natural Image Noise Dataset

Benoit Brummer and Christophe De Vleeschouwer

Université Catholique de Louvain, Belgium

Introduction

Photographic image noise occurs when a camera is not given enough time to capture the information present in a scene with the given settings and the sensor sensitivity (ISO setting) is increased.

Photographic development involves the use of a denoising algorithm to detect the noise and reconstruct the latent clean image. Convolutional neural networks have been the focus of research aiming to solve the denoising problem, but they require large amount of paired data to learn a mapping. In the absence of such data, artificial noise is added to a clean image and the result is unsatisfactory in real life applications.

We introduce the Natural Image Noise Dataset; a photographic image dataset of clean-noisy scenes taken with different ISO settings that is large, diverse, and of high enough quality to train neural networks for image denoising tasks.

We obtain state-of-the-art results with a single trained generative model, using various network architectures (including RED-Net and U-Net). The models appears to generalize well to other cameras.

Methods

Capture

- Multiple images per static scene (average of 6)
  - some images feature multiple ground-truths (avoid overfitting)
- Avoided movement:
  - no moving subject, stable lighting, fixed camera (tripod, wireless control, no excessive wind)
- No overexposure in the ground-truth
- Fixed aperture, manual focus

The dataset currently consists of 616 images from 101 different scenes. A large subset of the ground-truth images is shown on Figure 1.

Software processing

- raw-to-sRGB development using darktable (open source development software) [4]
- standard development without sharpening (should be applied after denoising as it amplifies noise)
- matched exposure (using darktable's automatic exposure to match the same histogram percentile target)
- white balance copied over from a ground-truth image
- overexposure checked visually with the overexposure indicator tool
- final pixel alignment with hugin's align_image_stack tool
- images with significant overexposure in the ground-truth, detected changes across versions, or misalignments were cropped or discarded
- saved as high-quality JPEG or lossless PNG

Publication

The Natural Image Noise Dataset is published on Wikimedia Commons, an online repository of free-use images and other digital media.

- Collaborative: open for inclusion, modification, and discussion
- Community can visually assess the ground-truth quality on the "Quality Image candidates" page
- Images can be visualized prior to full-size download
- Dataset can be downloaded in subset
- All versions are preserved, therefore a specific dated snapshot of the dataset can always be downloaded

Availability

The NIND is available for download and third-party contributions on:
https://commons.wikimedia.org/wiki/Natural_Image_Noise_Dataset

References

[1] Benoit Brummer and Christophe De Vleeschouwer. "Natural Image Noise Dataset." In: 2019 IEEE Conference on Computer Vision and Pattern Recognition (CVPR). 2019

[2] O. Ronneberger, P.Fischer, and T. Brox. "U-Net: Convolutional Networks for Biomedical Image Segmentation." In: Medical Image Computing and Computer-Assisted Intervention (MICCAI)

[3] Phillip Isola, Jun-Yan Zhu, Tinghui Zhou, and Alexei A. Efros. "Image-to-Image Translation with Conditional Adversarial Networks." In: 2017 IEEE Conference on Computer Vision and Pattern Recognition, CVPR 2017

[4] darktable. https://www.darktable.org/ .

[5] Abdelrahman Abdelhamed, Stephen Lin, and Michael S. Brown. "A High-Quality Denoising Dataset for Smartphone Cameras." In: The IEEE Conference on Computer Vision and Pattern Recognition (CVPR). June 2018.

Sample images from the NIND: ground-truths and varying ISO

Figure 1: Sample images from the NIND (center) with visual noise comparison (edges)

Training CNN-based denoising models: results

We trained various network architectures (such as DnCNN, RED-Net, and U-Net [2]) using the NIND and found that most networks obtain state-of-the-art performance, alas some of them provide significantly faster training and inference. We ended up using U-Net for most of our experiments as it provides a useful inference speed of $\lesssim$ 0.75 s per megapixel on a GeForce GTX 1070.

The model's performance is shown on Figure 2 (SSIM), Figure 4 (visually), and Figure 3 (both). We also acquired images with a different camera to test the generalization performance and found that the model trained on one camera could generalize without perceptible performance hit compared with the model trained on both cameras. (Figure 5) We also tested whether adding different data could deteriorate performance by adding 320 images from the SIDD [5] and we found that the model could handle them as well (even though a model trained only on the SIDD did not perform as well on our DSLR-like images).

Figure 2:
Test set denoising performance of
U-Net model trained on NIND,
U-Net trained on artificial noise,
BM3D
On five test images

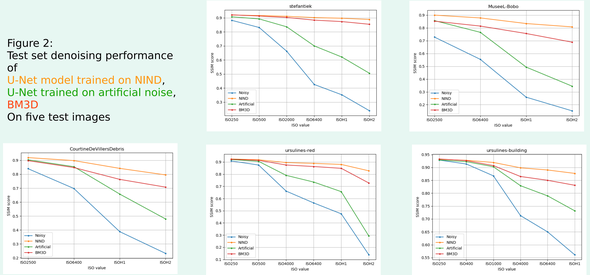

Figure 3: denoising stefantiek. 1: ISO200 ground-truth (1/20 s, SSIM: 1.00), 2: high ISO (1/1900 s, SSIM: 0.239), 3: 2 denoised using U-Net model trained with NIND (SSIM: 0.890), 4: 2 denoised using U-Net model trained with SIDD-320 (SSIM: 0.870), 5: 2 denoised using BM3D (SSIM: 0.854)

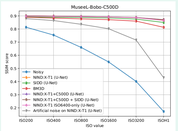

Figure 5: Generalization performance comparison: Image from a Canon 500D camera denoised with models trained on the NIND with only Fujifilm X-T1 pictures vs the entire dataset.

Figure 4: Visual comparison of a high-speed low-aperture ISO6400 image (top) denoised with a U-Net model trained with the NIND (middle) and with BM3D (bottom)

Contribution

- A rich dataset including diverse scenes captured with a variety of ISO values
- Denoising models trained on this dataset generalizes well to images captured with different sensors, as well as to dynamic scenes
- Dataset maintained on Wikimedia Commons with open access and curation
- Code available on https://github.com/trougnouf/mthesis-denoise (including a download script, cropping tool, testing and training with a single generator or (c)GAN discriminator(s))