TOO FAR, TOO SMALL, TOO DARK, TOO FOGGY:
ON THE USE OF MACHINE LEARNING FOR COMPUTATIONAL IMAGING PROBLEMS

George Barbastathis, Massachusetts Institute of Technology &
Singapore-MIT Alliance for Research and Technology (SMART) Centre
gbarb@mit.edu
Mo Deng, Massachusetts Institute of Technology
Shuai Li, Massachusetts Institute of Technology
Alexandre Goy, Massachusetts Institute of Technology
Kwabena Arthur, Massachusetts Institute of Technology

Key Words: Computational imaging, inverse problems, machine learning.

Computational imaging system design involves the joint optimization of hardware and software to deliver high fidelity images to a human user or artificial intelligence (AI) algorithm. For example, in medical tomography CAT scanners produce non-invasively cross-sectional images of the patient's organs and then medical professionals or, increasingly, automated recognition systems perform diagnosis and decide upon a course of treatment. We refer to this operation of AI as *image interpretation.*

This talk is about a different paradigm where machine learning (ML) is used at the step of image formation itself, *i.e.* for *image reconstruction* rather than interpretation. The ML algorithm, typically implemented as a deep neural network (DNN), is programmed using physically generated or rigorously simulated examples of objects and their associated signals produced on the sensor (or camera.) The training phase consists of adjusting the connection weights of the DNN until it becomes possible, given the sensor signal from a hitherto unseen object, for the DNN to yield an accurate estimate of the object's spatial structure.

The ML approach to solving inverse problems in such fashion has its roots in optimization methods employed long before in computational imaging, compressed sensing and dictionaries in particular. By replacing the proximal gradient step of the optimization with a DNN [K. Gregor & Y. LeCun, ICML 2010], it becomes possible to learn priors other than sparsity, and restrict the object class almost arbitrarily to facilitate the solution of "hard" inverse problems, e.g. highly ill-posed and highly noisy at the same time. Moreover, execution becomes very fast because pre-trained DNNs mostly consist of forward computations which can easily be run at real time, whereas traditional compressed sensing optimization routines are generally iterative. DNN training is time consuming too, but it is only run up front while developing the algorithm; it is not a burden during operation. Unfortunately, however, with the DNN approach some of the nice properties of compressed sensing are lost, most notably convexity.

In this talk we will review these basic developments and then discuss in detail their application to the specific problem of phase retrieval in lensless (free-space propagation) or defocused imaging systems. More specifically, we will investigate the impact of the power spectral density of the training example database on the quality of the reconstructions. We will review a sequence of papers where we first ignored this problem [A. Sinha et al, *Optica* 4:1117, 2017], then improved it in an *ad hoc* way by pre-modulation of the training examples [Li Shuai et al, *Opt. Express* 26:29340, 2018] and finally devised a dual-band approach where the signal is first separated into its low- and high-frequency components, their respective reconstructions are obtained by two DNNs trained separately and then re-composed by a third "synthesizer" DNN [Deng Mo et al, arXiv:1811.07945]. We will explain why each new attempt improves resolution and overall fidelity through progressively more balanced treatment of the spatial frequency spectrum.

We will also discuss implications of this method for phase retrieval under extremely low-photon (too dark) conditions [A. Goy et al, *Phys. Rev. Lett.* 121:243902, 2018] other related inverse problems, e.g. super resolution (too far or too small), and imaging through diffusers (too foggy.)