An approximate Bayesian formulation for deep image reconstruction in the presence of signal-dependent noise

Luis Amador, Elie Bretin, Nicolas Ducros

AN APPROXIMATE BAYESIAN FORMULATION FOR DEEP IMAGE RECONSTRUCTION IN THE PRESENCE OF SIGNAL-DEPENDENT NOISE

Luis Amador, Elie Bretin, Nicolas Ducros

Univ Lyon, INSA-Lyon, UCB Lyon 1,CNRS, Inserm, CREATIS UMR 5220, U1206, Lyon, France
Univ Lyon, INSA de Lyon, CNRS UMR 5208, Institut Camille Jordan, F-69621 Villeurbanne, France

ABSTRACT

Recently, a variety of unrolled networks have been proposed for image reconstruction. These can be interpreted as parameter-optimized algorithms that incorporate steps that are traditionally encountered during the optimization of a hand-crafted objective or in the Bayesian formulation. Here, we address the problem of training such networks in the presence of signal-dependent noise, which is more realistic that the common additive Gaussian noise; however, it is also much more computationally demanding, when possible. In particular, we focus on the deep expectation-maximization network and describe how to approximate the Bayesian denoised completion step to reduce the computational cost and memory requirements while limiting the reconstruction error. We present reconstruction results from simulated data at different noise levels. Our network yields higher reconstruction peak signal-to-noise ratios than other similar approaches. In particular, our network shows greater robustness in the practical case where the noise level is unknown or is badly estimated.

***Index Terms*—** Image reconstruction, deep learning, unfolded network, signal-dependent noise.

1. INTRODUCTION

Many biomedical modalities are ill-posed image reconstruction problems that require prior information to stabilize the solution, such as computerized tomography, magnetic resonance, and optical microscopy. In particular, single-pixel imaging is aimed at recovering an image from a few point measurements that correspond to the dot products between an image and a set of Hadamard functions that are implemented through a spatial light modulator [1]. This is a typical under-determined inverse problem subject to Poisson noise. Single-pixel imaging has been successfully applied to fluorescence microscopy, hyperspectral imaging, image-guided surgery, fluorescence lifetime imaging, and other such techniques.

Image reconstruction problems have long benefited from the theory of compressed sensing, but recent advances in deep learning have revolutionized the field (see [2], and references therein). In particular, unrolled networks that rely on the computation of traditional solutions (e.g., pseudo inverse, ℓ_1-penalized, maximum a posteriori) can be interpreted as iterative schemes that are optimized with respect to a particular task and database [3]. Unrolled networks can be used to efficiently solve a variety of problems, which includes image reconstruction problems in magnetic resonance imaging [4], computed tomography [5], and optical microscopy [6], and they are often based on convolutional neural networks [7], [8].

Inspired by the expectation-maximization (EM) algorithm [9], the deep EM network was recently proposed to solve inverse problems corrupted by signal-dependent noise [10], such as Poisson corrupted measurements encountered in computational optics. However, signal-dependent noise enforces the resolution of a different linear system for each image in the training database, which prevents pre-computation and poses a severe computational issue. In practice, approximate (e.g., diagonal) resolutions are implemented, which leads to sub-optimal image quality.

Our contribution to this problem is to introduce a novel approximate solution to the Bayesian reconstruction problem, where the computational complexity is compatible with the training phase that requires its evaluation for a large number of images. Having trained a deep EM net using this approximation, we observe improved quality compared to existing methods, such as [10]. In Section 2, we describe the forward model and the solution of the inverse problem using the deep EM algorithm. In Section 3, we detail the proposed Taylor approximation of the Bayesian solution. In Section 4, we describe how the network is implemented and trained. In Section 5, we report on and analyze our reconstruction results.

2. COMPUTATIONAL OPTICS

The philosophy of computational optics is to recover an image $\boldsymbol{f} \in [0,1]^N$ from hardware measurements using software $\boldsymbol{m} = \boldsymbol{H}_1 \boldsymbol{f} \in \mathbb{R}^M$ where $\boldsymbol{H}_1 \in \mathbb{R}^{M \times N}$ is a linear operator.

2.1. Forward problem

The system matrix $\boldsymbol{H}_1$ collects the patterns that are sequentially uploaded onto a spatial light modulator. We consider the case $M < N$, where we acquire fewer measurements than pixels in the image. The patterns are traditionally chosen on a Hadamard basis $\boldsymbol{H} \in \mathbb{R}^{N\times N}$ [11]; i.e., $\boldsymbol{H}_1 = \boldsymbol{S}\boldsymbol{H}$ with $\boldsymbol{S} = [\boldsymbol{I}_M, \boldsymbol{0}]$ and $\boldsymbol{H} = [\boldsymbol{H}_1; \boldsymbol{H}_2]$. The acquisition is corrupted by Poisson noise [12]

$$\hat{\mathbf{m}}^{\alpha,\pm} \sim \mathcal{P}(\alpha \boldsymbol{H}_1^{\pm} \boldsymbol{f}), \tag{1}$$

where α is the intensity (in photons) of the image. Note that experimental implementation of the negative values of $\boldsymbol{H}_1$ is achieved through the use of the positive patterns $\boldsymbol{H}_1^+$ and $\boldsymbol{H}_1^-$, such that $\boldsymbol{H}_1 = \boldsymbol{H}_1^+ - \boldsymbol{H}_1^-$ (see [13] for details). We finally introduce the normalized Hadamard measurements $\mathbf{m}^\alpha = \frac{1}{\alpha}(\hat{\mathbf{m}}^{\alpha,+} - \hat{\mathbf{m}}^{\alpha,-})$, such that $\mathbb{E}\left(\mathbf{m}^\alpha\right) = \boldsymbol{H}_1\boldsymbol{f}$. Note that the amplitude of the normalized measurements remains unchanged for different intensities of α

2.2. Inverse problem

We aim to compute the maximum a-posteriori solution of our problem

$$\underset{\boldsymbol{f}}{\operatorname{argmax}} \quad \log p(\boldsymbol{m}^\alpha|\boldsymbol{f}) + \log p(\boldsymbol{f}), \tag{2}$$

where the conditional probability function $p(\boldsymbol{m}^\alpha|\boldsymbol{f})$ represents the noise model, and the probability function $p(\boldsymbol{f})$ represents the prior knowledge about the unknown image.

2.2.1. Bayesian Denoised Completion

Under the assumptions $p(\boldsymbol{m}^\alpha|\boldsymbol{f}) \propto \exp -\frac{1}{2}\|\boldsymbol{H}_1\boldsymbol{f} - \boldsymbol{m}^\alpha\|^2_{\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}^{-1}}$ and $p(\boldsymbol{f}) \propto \exp -\frac{1}{2}\|\boldsymbol{H}\boldsymbol{f}\|^2_{\boldsymbol{\Sigma}^{-1}}$. where $\boldsymbol{\Sigma}_\alpha$ represents the noise covariance matrix, and $\boldsymbol{\Sigma} = \begin{bmatrix} \boldsymbol{\Sigma}_1 & \boldsymbol{\Sigma}_{21}^\top \\ \boldsymbol{\Sigma}_{21} & \boldsymbol{\Sigma}_2 \end{bmatrix}$ is the Hadamard image covariance, where $\boldsymbol{\Sigma}_1$, $\boldsymbol{\Sigma}_2$ and $\boldsymbol{\Sigma}_{12}$ are the blocks of the covariance matrix $\mathsf{Cov}\left(\boldsymbol{H}\boldsymbol{f}\right) = \mathsf{Cov}\left([\boldsymbol{H_1}; \boldsymbol{H_2}]\boldsymbol{f}\right)$ that can be precomputed as defined in [14]. It follows that $\boldsymbol{f} = \mathcal{G}^\alpha_{\text{Bay}}(\boldsymbol{m}^\alpha)$, where

$$\mathcal{G}^\alpha_{\text{Bay}}(\boldsymbol{m}^\alpha) = \boldsymbol{H}^\top \begin{bmatrix} \boldsymbol{I}_M \\ \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_1^{-1} \end{bmatrix} \boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_\alpha + \boldsymbol{\Sigma}_1)^{-1}\boldsymbol{m}^\alpha. \tag{3}$$

2.2.2. Deep expectation maximization

Deep-learning-based methods look for a reconstruction mapping $\mathcal{G}_{\boldsymbol{\theta}}$, where $\boldsymbol{\theta}$ represents the parameters of the deep neural network that are optimized during the training phase, by minimizing the mean squared error $\frac{1}{S}\sum_{s=1}^{S}\|\mathcal{G}_{\boldsymbol{\theta}}(\boldsymbol{m}^\alpha_{(s)}) - \boldsymbol{f}_{(s)}\|_2^2$, over an image database $\{\boldsymbol{f}_{(s)}, \boldsymbol{m}^\alpha_{(s)}\}_{1\le s\le S}$.

To solve the maximum a-posteriori problem, we adopt the deep expectation (EM) architecture $\boldsymbol{f}^{(k)} = \mathcal{G}^\alpha_{\boldsymbol{\theta},k}(\boldsymbol{m}^\alpha)$, where the k-th iteration $\boldsymbol{f}^{(k)}$ is defined recursively by [10]

$$\bar{\boldsymbol{f}}^{(k)} = \boldsymbol{f}^{(k)} + \mathcal{G}^\alpha_{\text{Bay}}(\boldsymbol{m}^\alpha - \boldsymbol{H}_1\boldsymbol{f}^{(k)}) \tag{4a}$$

$$\boldsymbol{f}^{(k+1)} = \mathcal{D}_{\boldsymbol{\theta}}(\bar{\boldsymbol{f}}^{(k)}) \tag{4b}$$

with the initialization $\boldsymbol{f}^{(0)} = \boldsymbol{0}$. Here, $\mathcal{D}_{\boldsymbol{\theta}}$ is a neural network (e.g., convolutional neural network) that corrects artifacts in the image domain.

2.2.3. Practical issues

The computational cost of the training phase is dominated by the evaluation of Equation (3), and in particular by

$$\boldsymbol{P}_\alpha = \boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_\alpha + \boldsymbol{\Sigma}_1)^{-1}, \tag{5}$$

where the noise covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$ can be reasonably estimated from the raw data [14]. However, as $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$ depends on the image and the intensity, no precomputation is possible, and a different linear system should be solved for each of the S images in the training database. To alleviate this problem, [10] proposed to replace $\boldsymbol{P}_\alpha$ in Equation (3) by the diagonal approximation

$$\boldsymbol{P}_{\alpha,1} = \mathsf{diag}\left(\boldsymbol{\Sigma}_1\right)\left[\boldsymbol{\Sigma}_\alpha + \mathsf{diag}\left(\boldsymbol{\Sigma}_1\right)\right]^{-1}. \tag{6}$$

As the noise covariance matrix $\boldsymbol{\Sigma}_{\boldsymbol{\alpha}}$ is diagonal for independent measurements, the computation of Equation (6) is trivial. Despite the simplicity and effectiveness of this approximation, it leads to degraded image quality, as will be shown in Section 5.

3. PROPOSED METHOD

Instead of solving S different linear systems, our idea is to introduce an approximation $\boldsymbol{P}_\alpha \simeq \boldsymbol{Q}_\alpha + \boldsymbol{R}_\alpha$, where $\boldsymbol{Q}_\alpha$ is independent of $\boldsymbol{f}$ and can be precomputed, and $\boldsymbol{R}_\alpha$ is a correction term for $\boldsymbol{Q}_\alpha$, which depends on $\boldsymbol{f}$, but requires no system resolution (or matrix inversion) during training.

3.1. Taylor approximation

To do so, we introduce an approximate noise covariance matrix $\widetilde{\boldsymbol{\Sigma}}_\alpha$, independent of $\boldsymbol{f}$ and such that $(\boldsymbol{\Sigma}_1 + \widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1}(\boldsymbol{\Sigma}_\alpha - \widetilde{\boldsymbol{\Sigma}}_\alpha)$ is small. The choice of $\widetilde{\boldsymbol{\Sigma}}_\alpha$ is discussed in Section 3.2. Then, we note that

$$\begin{aligned}\boldsymbol{P}_\alpha &= \boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_\alpha + \widetilde{\boldsymbol{\Sigma}}_\alpha - \widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1} \\ &= \boldsymbol{\Sigma}_1[(\boldsymbol{\Sigma}_1 + \widetilde{\boldsymbol{\Sigma}}_\alpha)(\boldsymbol{I}_M + (\boldsymbol{\Sigma}_1 + \widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1}(\boldsymbol{\Sigma}_\alpha - \widetilde{\boldsymbol{\Sigma}}_\alpha))]^{-1} \\ &= \boldsymbol{\Sigma}_1(\boldsymbol{I}_M + (\boldsymbol{\Sigma}_1 + \widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1}(\boldsymbol{\Sigma}_\alpha - \widetilde{\boldsymbol{\Sigma}}_\alpha))^{-1}(\boldsymbol{\Sigma}_1 + \widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1}.\end{aligned}$$

The first order Taylor approximation gives $\boldsymbol{P}_\alpha \simeq \boldsymbol{P}_{\alpha,2}$, defined by $\boldsymbol{P}_{\alpha,2} = \boldsymbol{\Sigma}_1[\boldsymbol{I}_M - (\boldsymbol{\Sigma}_1 + \widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1}(\boldsymbol{\Sigma}_\alpha - \widetilde{\boldsymbol{\Sigma}}_\alpha)](\boldsymbol{\Sigma}_1 +$

$\widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1}$, which is of the form $\boldsymbol{P}_{\alpha,2} = \boldsymbol{Q}_\alpha + \boldsymbol{R}_\alpha$, where we identify

$$\boldsymbol{Q}_\alpha = \boldsymbol{\Sigma}_1(\boldsymbol{\Sigma}_1 + \widetilde{\boldsymbol{\Sigma}}_\alpha)^{-1} \tag{7a}$$

$$\boldsymbol{R}_\alpha = -\boldsymbol{Q}_\alpha(\boldsymbol{\Sigma}_\alpha - \widetilde{\boldsymbol{\Sigma}}_\alpha)\boldsymbol{\Sigma}_1^{-1}\boldsymbol{Q}_\alpha \tag{7b}$$

Note that $\boldsymbol{Q}_\alpha$ can be computed once and for all before the learning phase, while the computation of $\boldsymbol{R}_\alpha$ only requires three matrix products (including one with a diagonal matrix) for a given $\widetilde{\boldsymbol{\Sigma}}_\alpha$. Note that this does not require any matrix inversion.

3.2. Computation of the approximate noise covariance

The noise covariance matrix $\boldsymbol{\Sigma}_\alpha$ is directly linked to to the total image intensity. Indeed, by deriving the variance of the raw measurements, we obtain $\boldsymbol{\Sigma}_\alpha = \boldsymbol{I}_M \frac{1}{\alpha}\left(\boldsymbol{H}_1^+ \boldsymbol{f} + \boldsymbol{H}_1^- \boldsymbol{f}\right)$. Then, using the properties of the Hadamard matrix $\boldsymbol{H}_1^+ + \boldsymbol{H}_1^- = \mathbf{1}$, we have $\boldsymbol{\Sigma}_\alpha = \boldsymbol{I}_M \frac{1}{\alpha}\sum_{n=1}^N \boldsymbol{f}_n \leq \boldsymbol{I}_M \frac{N}{\alpha}$, where the upper bound is reached when $\widetilde{\boldsymbol{\Sigma}}_\alpha = \boldsymbol{I}_M \frac{N}{\alpha}$.

4. EXPERIMENTS

In our experiments, we choose $M = 1024$ Hadamard patterns of size $N = 64 \times 64$ pixels. We train two variants of the deep EM network (i.e., $K = 1$ and $K = 5$ iterations) in an end-to-end fashion, choosing $\mathcal{D}_{\boldsymbol{\theta}}$ as the convolutional neural network described in [15]. All of the networks are trained using 105,000 images (i.e., the 'unlabeled' and 'train' subsets of the STL10 database; [16]); 8,000 images are used for the test (i.e., the 'test' subset of STL10). The original 96×96 images were cropped to 64×64, and normalized between -1 and 1. We implement our methods using the Python SPyRiT package [17], which is based on Pytorch [18]. For training, we consider the ADAM optimizer for 30 epochs. The step size is initialized to 10^{-3} and is reduced by half every 10 epochs. The weight decay regularization parameter is set to 10^{-8}. The number of learned parameters is 8,129. The training phase takes 38 min for $K = 1$, and 1 h 40 min for $K = 5$ on a NVIDIA GeForce GTX 1660 Ti, for both the diagonal and proposed approximations

5. RESULTS AND DISCUSSION

5.1. Influence of approximations in Bayesian denoised completion

First, we compare the image reconstructions from the Bayesian denoised completion of Equation (3) using the exact formulation corresponding to $\boldsymbol{P}_\alpha$ given by Equation (5), the diagonal approximation given by Equation (6), and the proposed approximation given by Equation (7). We consider four different noise levels corresponding to four different image intensities α. Increasing image intensities correspond to decreasing signal-to-noise ratios.

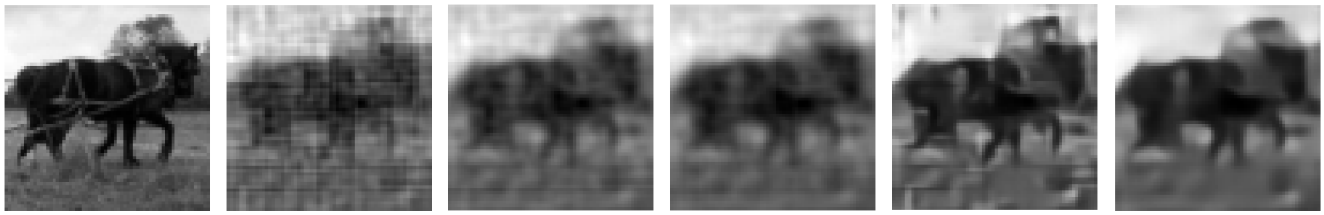
We report the means and standard deviation of the reconstruction PSNR computed over the test database in the left half of Table 1, and display the corresponding reconstructed images in Fig. 1b-d. We observe that the diagonal approximation leads to reduced image quality compared to the exact solution, in particular at the high noise levels that correspond to low image intensities. For instance, at $\alpha = 5$ photons, the mean PSNR degradation is 1 dB. For the top row image displayed in Fig. 1, the PSNR degradation reaches 1.32 dB and the image reconstruction with the diagonal approximation shows severe artifacts. However, the proposed approximation performs very similarly to the exact Bayesian inversion, both in terms of mean PSNRs (e.g., degradation equal to or less than 0.78 dB) and visual quality.

5.2. Advantage of EM deep reconstruction

In Table 1, we report the reconstruction quality metrics over the stl-10 test dataset for the different reconstruction networks using the diagonal approximation and the proposed approximation. First, we observe that deep reconstruction outperforms Bayesian denoised completion, regardless the implementation of Equation (3) as exact or approximated (i.e., diagonal or proposed). For instance, using the diagonal approximation, the mean PSNR is improved by 20.45-19.27 = 1.18 dB at 2 photons, by 22.18 - 20.73 = 1.45 dB at 5 photons, and by 23.58-21.97 = 1.61 dB at 10 photons. Secondly, among the deep methods, the networks trained with the proposed approximation slightly outperform the networks trained using the diagonal approximation. For instance, at 5 photons, the improvement is 22.37 - 22.18 = 0.19 dB using $K = 1$ iterations. The reconstruction quality can be further improved by 22.48 - 22.37 = 0.11 dB using $K = 5$ iterations in the EM-Net.

5.3. Robustness of the proposed method

Our objective here is to assess the robustness of the different reconstruction methods when the estimation of the light intensity α (i.e., the noise level) is either biased or does not fit with the training noise level of the available networks. More specifically, in Table 2, we propose to reconstruct an image acquired at α in $\{2, 5, 10\}$ photons with networks trained at $\alpha = 50$ photons. As expected, the reconstruction quality is lower when the training noise level deviates from the acquisition noise level. More surprisingly, at high noise levels (i.e., low light intensities), the degradation is more limited using the proposed method. Indeed, at $\alpha = 5$ photons, the mean PSNR of the EM-Net reconstruction is reduced by $22.18 - 21.62 = 0.56$ dB using the diagonal approximation, while the reduction is only $22.37 - 22.00 = 0.37$ dB when using the proposed approximation with $K = 1$. The proposed approximation with $K = 5$ has a lower mean PSNR; however,the PSNR standard deviation is improved. Fig. 1 shows



(a) Ground Truth. (b) PSNR = 21.08 dB (c) PSNR = 22.40 dB (d) PSNR = 22.43 dB (e) PSNR = 21.82 dB (f) PSNR = 22.88 dB

Fig. 1: Reconstructions on the stl-10 test dataset for the different methods assuming $\alpha = 5$ ph. (a) Ground truth. (b-d) Bayesian denoised completion $\mathcal{G}^{\alpha}_{\text{Bay}}$ using the diagonal approximation (b), the exact solution (c), and the proposed approximation (d). (e, f) EM-Net reconstruction for $K = 5$ iterations using the diagonal approximation (e) and the proposed method (f) for networks trained using α =50 ph. PSNR, peak signal-to-noise ratio.

		Bayesian denoised completion			EM-Net			
α		Diag	Exact	Proposed	Diag (K=1)	Proposed (K=1)	Diag (K=5)	Proposed (K=5)
2	PSNR	19.27 ± 1.45	20.88 ± 1.67	$\mathbf{20.10 \pm 1.62}$	20.45 ± 1.70	$\mathbf{20.67 \pm 1.70}$	20.70 ± 1.63	20.75 ± 1.67
	SSIM	0.781 ± 0.137	0.831 ± 0.141	$\mathbf{0.809 \pm 0.138}$	0.817 ± 0.139	$\mathbf{0.825 \pm 0.141}$	0.825 ± 0.145	0.827 ± 0.141
5	PSNR	20.73 ± 1.37	21.73 ± 1.43	$\mathbf{21.61 \pm 1.57}$	22.18 ± 1.71	$\mathbf{22.37 \pm 1.70}$	22.51 ± 1.66	22.48 ± 1.70
	SSIM	0.851 ± 0.098	0.877 ± 0.093	$\mathbf{0.873 \pm 0.093}$	0.883 ± 0.096	$\mathbf{0.889 \pm 0.094}$	0.889 ± 0.095	0.888 ± 0.097
10	PSNR	21.97 ± 1.31	22.94 ± 1.42	$\mathbf{22.82 \pm 1.58}$	23.58 ± 1.80	$\mathbf{23.71 \pm 1.80}$	23.92 ± 1.71	23.83 ± 1.74
	SSIM	0.892 ± 0.072	0.910 ± 0.067	$\mathbf{0.908 \pm 0.066}$	0.918 ± 0.067	$\mathbf{0.920 \pm 0.064}$	0.923 ± 0.065	0.921 ± 0.066
50	PSNR	25.25 ± 1.37	25.84 ± 1.52	$\mathbf{25.77 \pm 1.66}$	26.82 ± 1.91	$\mathbf{26.82 \pm 1.94}$	27.06 ± 1.89	27.00 ± 1.90
	SSIM	0.951 ± 0.035	0.957 ± 0.032	$\mathbf{0.956 \pm 0.032}$	0.961 ± 0.033	$\mathbf{0.963 \pm 0.029}$	0.965 ± 0.029	0.964 ± 0.030

Table 1: Peak signal-to-noise ratio (PSNR) and structural similarity (SSIM) of the reconstructions over the stl-10 test set. The Bayesian denoised completion reconstruction corresponds to Equation (3) using the diagonal ('Diag') approximation given by Equation (6), the exact ('Exact') inversion given by Equation (5), and the proposed ('Proposed') approximation given by Equation (7). The EM-Net reconstruction corresponds to Equation (4) using either the diagonal ('Diag') or proposed ('Proposed') approximation, for $K = 1$ and $K = 5$ iterations. Here, all images are reconstructed with $M = 1024$, and the networks are trained with the image intensity α used during acquisition.

α		Diag (K=1)	Proposed (K=1)	Proposed (K=5)
2	PSNR	19.86 ± 1.70	$\mathbf{20.33 \pm 1.76}$	$\mathbf{19.86 \pm 1.46}$
	SSIM	0.795 ± 0.139	$\mathbf{0.812 \pm 0.141}$	$\mathbf{0.817 \pm 0.143}$
5	PSNR	21.62 ± 1.68	$\mathbf{22.00 \pm 1.77}$	$\mathbf{21.87 \pm 1.52}$
	SSIM	0.869 ± 0.095	$\mathbf{0.878 \pm 0.094}$	$\mathbf{0.884 \pm 0.094}$
10	PSNR	23.16 ± 1.69	$\mathbf{23.40 \pm 1.78}$	$\mathbf{23.48 \pm 1.60}$
	SSIM	0.910 ± 0.066	$\mathbf{0.915 \pm 0.065}$	$\mathbf{0.920 \pm 0.065}$

Table 2: Robustness of the EM-Net reconstruction for the acquisition of noise levels that differ from the training noise level. Details as for the legend to Table 1, except that all of the networks are trained with $\alpha = 50$ ph. PSNR, peak signal-to-noise ratio; SSIM structural similarity.

an example of the deep EM reconstruction obtained for an image intensity (i.e., $\alpha = 5$ ph) that differs from that used during training ($\alpha = 50$ ph). We observe that the proposed approximation preserves the image regularity, whereas the network accentuates artifacts from the diagonal approximation, which results in a drop in PSNR, for both the $K = 1$ and $K = 5$ variants.

6. CONCLUSION

We consider the problem of reconstruction of an image from few measurements corrupted by signal-dependent noise. In particular, we focus on an unrolled network that is based on a Bayesian formulation of this inverse problem. It requires the resolution of many different linear systems during the training phase, which raises memory and computational cost issues. While a diagonal approximation that does not suffer from these limitations has been proposed, it leads to degraded image quality. In this work, we propose a novel approximation, which has the same advantages as the diagonal approximation, while leading to improved image quality. Coupled with neural networks, this improves both the quality and the robustness of the reconstruction method. Our approach is well suited to photonic applications where measurements are corrupted by Poisson noise. In future work, we will assess the method for experimental acquisitions for hyperspectral imaging.

7. ACKNOWLEDGMENTS

This work was supported by the French National Research Agency (ANR), under Grant ANR-17-CE19-0003 (ARMONI Project), and performed within the framework of the LABEX MILYON (ANR-10-LABX-0070) and the LABEX PRIMES (ANR-11-LABX-0063) of Université de Lyon.

8. REFERENCES

[1] Matthew P. Edgar et al., "Principles and prospects for single-pixel imaging," Nature Photonics, vol. 13, no. 1, pp. 13–20, Jan. 2019.

[2] Simon Arridge et al., "Solving inverse problems using data-driven models," Acta Numerica, vol. 28, pp. 1–174, 2019.

[3] Vishal Monga, Yuelong Li, and Yonina C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," IEEE Signal Processing Magazine, vol. 38, no. 2, pp. 18–44, 2021.

[4] H. K. Aggarwal et al., "Modl: Model-based deep learning architecture for inverse problems," IEEE Transactions on Medical Imaging, vol. 38, no. 2, pp. 394–405, 2019.

[5] H. Gupta et al., "Cnn-based projected gradient descent for consistent ct image reconstruction," IEEE Transactions on Medical Imaging, vol. 37, no. 6, pp. 1440–1453, June 2018.

[6] M. R. Kellman, E. Bostan, N. A. Repina, and L. Waller, "Physics-based learned design: Optimized coded-illumination for quantitative phase imaging," IEEE Transactions on Computational Imaging, vol. 5, no. 3, pp. 344–353, 2019.

[7] J. Adler and O. Öktem, "Learned primal-dual reconstruction," IEEE Transactions on Medical Imaging, vol. 37, no. 6, pp. 1322–1332, 2018.

[8] M. T. McCann et al., "Convolutional neural networks for inverse problems in imaging: A review," IEEE Signal Processing Magazine, vol. 34, no. 6, pp. 85–95, 2017.

[9] J. A. Fessler et al., "Space-alternating generalized expectation-maximization algorithm," IEEE Transactions on Signal Processing, vol. 42, no. 10, pp. 2664–2677, 1994.

[10] Antonio Lorente Mur, Paul Bataille, Françoise Peyrin, and Nicolas Ducros, "Deep expectation-maximization for image reconstruction from under-sampled poisson data," in 2021 IEEE 18th International Symposium on Biomedical Imaging (ISBI), 2021, pp. 1535–1539.

[11] M. Ochoa et al., "Assessing patterns for compressive fluorescence lifetime imaging," Opt. Lett., vol. 43, no. 18, pp. 4370–4373, Sep 2018.

[12] A. Foi et al., "Practical poissonian-gaussian noise modeling and fitting for single-image raw-data," IEEE Transactions on Image Processing, vol. 17, no. 10, pp. 1737–1754, Oct. 2008.

[13] Antonio Lorente Mur et al., "Handling negative patterns for fast single-pixel lifetime imaging," in SPIE Photonics : Molecular-Guided Surgery: Molecules, Devices, and Applications V, 2019, vol. 10862.

[14] A. Lorente Mur, P. Leclerc, F. Peyrin, and N. Ducros, "Single-pixel image reconstruction from experimental data using neural networks," in Optics Express, 2021, vol. 29, pp. 17097–17110.

[15] N. Ducros et al., "A completion network for reconstruction from compressed acquisition," in 2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI), 2020, pp. 619–623.

[16] Adam Coates et al., "An analysis of single-layer networks in unsupervised feature learning," in Proceedings of the fourteenth international conference on artificial intelligence and statistics, 2011, pp. 215–223.

[17] A. Lorente Mur, F. Peyrin, and N. Ducros, "Single-Pixel pYthon Image Reconstruction Toolbox (spyrit) Version 1.2," https://github.com/openspyrit/spyrit, 2022.

[18] Adam Paszke et al., "Pytorch: An imperative style, high-performance deep learning library," in Advances in Neural Information Processing Systems 32, pp. 8024–8035. Curran Associates, Inc., 2019.