3D denoised completion network for deep single-pixel reconstruction of hyperspectral images

Valeriya Pronina, Antonio Lorente Mur, Juan F. P. J. Abascal, Françoise Peyrin, Dmitry Dylov, Nicolas Ducros

Research Article | Vol. 29, No. 24 / 22 Nov 2021 / Optics Express 39559
Optics EXPRESS

3D denoised completion network for deep single-pixel reconstruction of hyperspectral images

VALERIYA PRONINA,[1,3] ANTONIO LORENTE MUR,[2,3] JUAN F. P. J. ABASCAL,[2] FRANÇOISE PEYRIN,[2] DMITRY V. DYLOV,[1,4,5] AND NICOLAS DUCROS[2,4,6]

[1] *Skolkovo Institute of Science and Technology, Bolshoy Boulevard 30, bld. 1, Moscow 121205, Russia*
[2] *Univ Lyon, INSA-Lyon, Université Claude Bernard Lyon 1, UJM Saint-Etienne, CNRS, Inserm, CREATIS UMR 5220, U1206, F-69621 Lyon, France*
[3] *Co-first authors*
[4] *Co-senior authors*
[5] *d.dylov@skoltech.ru*
[6] *nicolas.ducros@insa-lyon.fr*

Abstract: Single-pixel imaging acquires an image by measuring its coefficients in a transform domain, thanks to a spatial light modulator. However, as measurements are sequential, only a few coefficients can be measured in the real-time applications. Therefore, single-pixel reconstruction is usually an underdetermined inverse problem that requires regularization to obtain an appropriate solution. Combined with a spectral detector, the concept of single-pixel imaging allows for hyperspectral imaging. While each channel can be reconstructed independently, we propose to exploit the spectral redundancy between channels to regularize the reconstruction problem. In particular, we introduce a denoised completion network that includes 3D convolution filters. Contrary to black-box approaches, our network combines the classical Tikhonov theory with the deep learning methodology, leading to an explainable network. Considering both simulated and experimental data, we demonstrate that the proposed approach yields hyperspectral images with higher quantitative metrics than the approaches developed for grayscale images.

1. Introduction

Single-pixel imaging is a computational technique that can reconstruct an image from a single point detector [1]. It has found applications in fluorescence microscopy [2], image-guided surgery [3], diffuse optical tomography [4], short-wave infrared imaging [5], imaging through scattering media [6], 3D imaging [7]. Considering a spectral detector, the concept of single-pixel imaging extends to generic imaging. Multispectral imaging was demonstrated with a dispersive unit followed by photomultiplier tubes [8,9]. Hyperspectral single-pixel imaging has been demonstrated with a compact fiber spectrometer [10,11] or exploiting a Michelson interferometer [12]. These works focus on improvement of hyperspectral images (HSI) acquisition strategies and leave the solution of image reconstruction to classical optimization schemes.

Single-pixel measurements can be modelled as dot products between an underlying image and some two-dimensional patterns that are implemented through a spatial light modulator (SLM) [13]. To limit the acquisition times, it is highly desirable to reduce the number of the light patterns, which leads to an under-determined ill-posed inverse problem. Classical approaches to solve such problems are based on compressive sensing [14]; however, deep learning (DL) has also proven efficient in many optical problems, including image deblurring [15], image reconstruction [16], image denoising [17], and spectral and lifetime unmixing [18,19]. Not surprisingly, single-pixel imaging has also benefited from DL. In [20], a neural network was

#443134 https://doi.org/10.1364/OE.443134
Journal © 2021 Received 13 Sep 2021; revised 26 Oct 2021; accepted 31 Oct 2021; published 12 Nov 2021

Optics EXPRESS

applied to reconstruct and improve the solution obtained with computational ghost imaging (CGI), a method closely related to single-pixel imaging [21]. In [22], Rizvi *et al.* proposed to couple the real-time Fourier single-pixel imaging scheme with deep convolutional autoencoder network to improve the image quality and the reconstruction speed. In [23], a deep convolutional autoencoder network was shown to outperform compressed sensing in image reconstruction task. Hoshi *et al.* [24] proposed to exploit the correlation between measured single-pixel data and to use a recurrent neural network for image recovery. In [25], a compressed sensing generative adversarial network was proposed for reconstruction of heavily downsampled images. Contrary to the empirical design of the network in [23], [26] introduced an architecture where the first layer was shown to estimate the conditional expectation of the image given noiseless measurements. In [27], the network was generalized to the case of noisy data and to various noise levels.

A deep learning based solution for recovery of single-pixel HSI was proposed by Arias *et al.* in [28]. The approach is two-step: during the first stage, HSIs are reconstructed from compressive samples in a band-wise manner via ISTA-Net, while the second stage incorporates a merge-and-run deep neural network for image refinement. Recently, DL-based methods also showed to be an effective instrument in tasks of hyperspectral image denoising. Given the high dimensionality of the hyperspectral data, many researchers naturally resorted to 3D neural networks to solve the HSI restoration problem [29–31]. However, while the other computational hyperspectral approaches (*e.g.*, CASSI [32] or CS MUSI [33]) exploit full 3D convolutional networks to engage the hyperspectral dimension, the existing *single-pixel* reconstruction networks are still limited to 2D convolutions, with the problem of hyperspectral reconstruction still awaiting to be addressed.

In this paper, we propose deep learning networks for hyperspectral single-pixel imaging. Contrary to black-box approaches, we propose interpretable networks where layers provide solution of classical inverse problems and generalize the reconstruction network proposed in [27] to the more complex problem of HSI single-pixel imaging. Our networks combine the classical optimization scheme for denoising in the measurement domain with the deep learning approach for projection from the measurement domain to the image domain, and a final regularization of the solution in the image domain. To the best of our knowledge, this is the first effort to apply generalized Tikhonov regularization and to develop a 3D reconstructor for the single-pixel HSI. In addition, since there is a lack of HSI datasets to train such networks, we propose a method to simulate a HSI dataset from RGB images dataset. Finally, the proposed networks are evaluated both on simulated and experimental HSI data. Upon the acceptance of the manuscript, our hyperspectral reconstruction network will be integrated into the open source `spyrit` package [34], and the pipeline for HSI synthesis from RGB images will be integrated into the open source `spihim` package [35].

2. Single-pixel imaging

2.1. Image acquisition

Single-pixel imaging can be modeled as $\boldsymbol{m} = \alpha \boldsymbol{H}_1 \boldsymbol{f}$, where $\boldsymbol{m} \in \mathbb{R}^M$ represents the raw measurements, $\boldsymbol{H}_1 \in \mathbb{R}^{M \times N}$, with $M<N$, are the patterns uploaded onto a SLM, $\boldsymbol{f} \in [0;1]^N$ is the unknown image, and α is the unknown image intensity (in photons). The patterns $\boldsymbol{H}_1$ are traditionally chosen in an orthogonal basis $\boldsymbol{H} \in \mathbb{R}^{N \times N}$, with the classical choices including Fourier, discrete cosine transform, wavelets, and Hadamard bases. They can be either chosen before the acquisition [36] or sampled adaptively during the acquisition [37]. In practice, the raw measurements are corrupted by the mixture of Poisson, and Gaussian noise [38,39], where the Poisson noise is signal-dependent and arises due to the discrete nature of the electronic charge, and the Gaussian noise occurs due to the fluctuations in the circuit and is signal-independent. Because the SLMs, such as digital micromirror devices (DMDs), cannot implement the negative values, it is necessary to decompose the patterns $\boldsymbol{H}_1$ into a set of patterns with only the positive

Optics EXPRESS

values [40], leading to the measurement of

$$\boldsymbol{m}^{\alpha}_{+,-} \sim K\,\mathcal{P}(\alpha \boldsymbol{H}_1^{+,-}\boldsymbol{f}) + \mathcal{N}(\mu_{\text{dark}}, \sigma^2_{\text{dark}}). \tag{1}$$

Here, $\mathcal{P}$ denotes the Poisson distribution, while $\boldsymbol{H}_1^+$ and $\boldsymbol{H}_1^-$ are the positive and the negative parts of $\boldsymbol{H}_1$, respectively, K is the calibration coefficient that represents the overall system gain, μ_{dark} and σ^2_{dark} are the mean and the variance of the dark current. One simple, yet efficient, differential approach is to measure the fraction

$$\boldsymbol{m}^{\alpha} = \frac{\boldsymbol{m}^{\alpha}_{+} - \boldsymbol{m}^{\alpha}_{-}}{K\alpha}, \tag{2}$$

where the normalization factor $K\alpha$ is introduced to get the images with intensities in $[0, 1]$.

2.2. Deep image reconstruction

Following the recent widespread success of deep neural networks, the reconstruction of $\boldsymbol{f}$ was proposed to be performed with the non-linear models [26]

$$\hat{\boldsymbol{f}} = \mathcal{G}_{\boldsymbol{\theta}}(\boldsymbol{m}^{\alpha}), \tag{3}$$

where $\mathcal{G}$ is a neural network, the parameters of which are denoted as $\boldsymbol{\theta}$. Given an image database consisting of S measurement-image pairs $\{\boldsymbol{m}^{\alpha,(s)}, \boldsymbol{f}^{(s)}\}$, $1 \le s \le S$, the network parameters $\boldsymbol{\theta}$ are optimized during a training phase by minimizing the so-called loss function

$$\min_{\boldsymbol{\theta}} \sum_{s=1}^{S} \|\boldsymbol{f}^{(s)} - \mathcal{G}_{\boldsymbol{\theta}}(\boldsymbol{m}^{\alpha,(s)})\|_2^2 + \mathcal{J}(\boldsymbol{\theta}), \tag{4}$$

where $\mathcal{J}$ is typically chosen to penalize large network parameters and reduce the freedom in the neural network. Traditionally, neural networks are a composition of cascaded layers

$$\mathcal{G}_{\boldsymbol{\theta}} = \mathcal{G}^{L}_{\boldsymbol{\theta}} \circ \ldots \circ \mathcal{G}^{1}_{\boldsymbol{\theta}}, \tag{5}$$

where $\mathcal{G}^{\ell}_{\boldsymbol{\theta}}$, $1 \le \ell \le L$ is the ℓ-th (nonlinear) layer of the network and $\circ$ is the function composition.

2.3. Denoised completion (Tikhonov) network

The denoised completion network (DC-Net) establishes a link between traditional and deep image reconstruction methods by freezing the first layer of the network such that it provides the best linear estimator of the minimum mean squared error solution, also known as the Tikhonov solution. An overview of the DC-Net architecture is given in Fig. 1(a).

As described in [27], the Tikhonov solution is given by $\tilde{\boldsymbol{f}} = \boldsymbol{H}^{\top}\tilde{\boldsymbol{y}}$, where $\tilde{\boldsymbol{y}}(\boldsymbol{m}^{\alpha}) = [\boldsymbol{y}_1{}^{\top}, \boldsymbol{y}_2{}^{\top}]^{\top}$ can be computed in two steps as

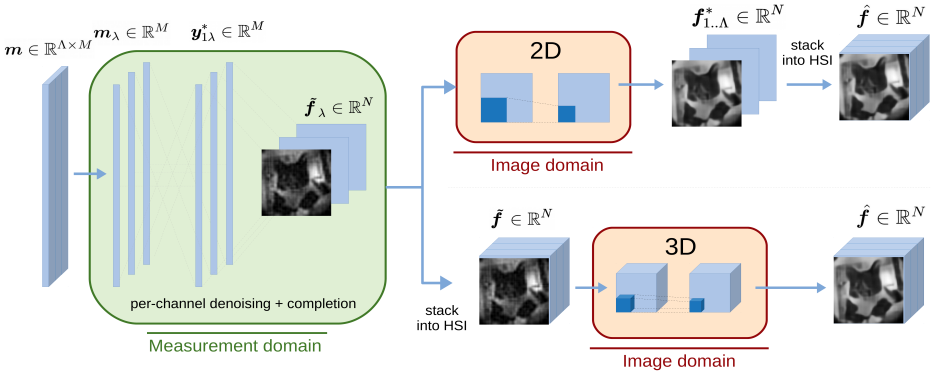
$$\boldsymbol{y}_1 = \boldsymbol{\mu}_1 + \boldsymbol{\Sigma}_1[\boldsymbol{\Sigma}_1 + \boldsymbol{\Sigma}_{\alpha}]^{-1}(\boldsymbol{m}^{\alpha} - \boldsymbol{\mu}_1), \tag{6a}$$

$$\boldsymbol{y}_2 = \boldsymbol{\mu}_2 + \boldsymbol{\Sigma}_{21}\boldsymbol{\Sigma}_1^{-1}[\boldsymbol{y}_1 - \boldsymbol{\mu}_1]. \tag{6b}$$

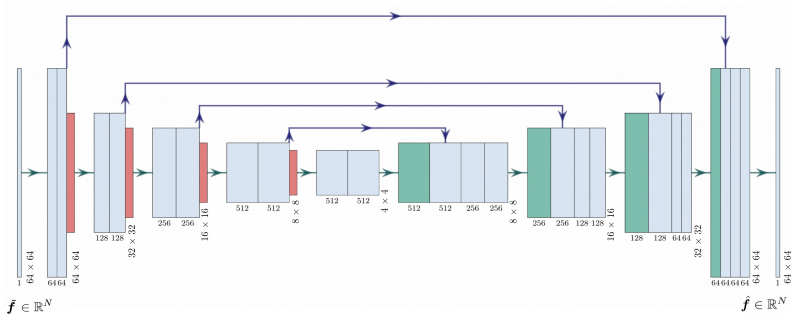
Here, $\boldsymbol{\mu}_1 \in \mathbb{R}^M$ and $\boldsymbol{\Sigma}_1 \in \mathbb{R}^{M\times M}$ are the expected value and covariance of the measured coefficients, respectively, $\boldsymbol{\Sigma}_{\alpha} \in \mathbb{R}^{M\times M}$ is the noise covariance, $\boldsymbol{\mu}_2 \in \mathbb{R}^{N-M}$ is the expected value of the missing coefficients, and $\boldsymbol{\Sigma}_{21} \in \mathbb{R}^{(N-M)\times M}$ is the covariance between the missing and measured coefficients. The computation of these quantities is described in 3.3.

The first step given by Eq. (6a) can be interpreted as the denoising of the raw measurements while the second step given by Eq. (6b) can be interpreted as the completion of the missing coefficients with relevant values. This choice for the first layer allows also for faster training compared to alternatives based on pseudo-inverse or back projection [26].

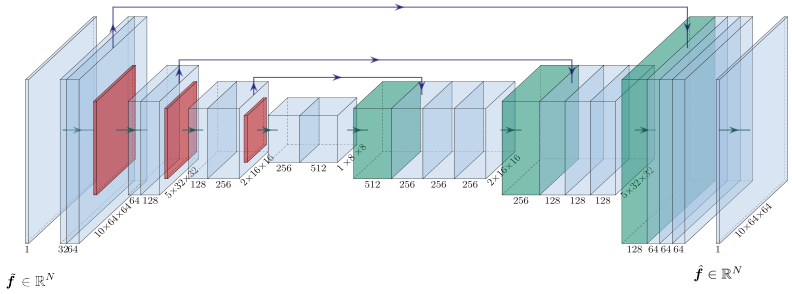
Optics EXPRESS



(a)



(b)



(c)

Fig. 1. Overview of the deep reconstruction networks. (a) Architecture of the 2D and 3D denoised completion networks (DC-Net). Denoising and completion occur in the measurement domain, for all channels independently in both the 2D and 3D cases. In the 2D case (top part of the diagram), the learnable layers ($\mathcal{G}^L_{\boldsymbol{\theta}} \circ \ldots \circ \mathcal{G}^2_{\boldsymbol{\theta}}$) correspond to 2D UNet; in the 3D case (bottom part of the diagram), to a 3D UNet. (b) Architecture of the 2D UNet. (c) Architecture of the 3D UNet.

Optics EXPRESS

3. Proposed deep hyperspectral reconstruction

In the specific case of hyperspectral imaging, the (noiseless) image formation process can be written as

$$\boldsymbol{m}_\lambda^\alpha = \alpha \boldsymbol{H}_1 \boldsymbol{f}_\lambda, \quad 1 \leq \lambda \leq \Lambda, \tag{7}$$

where $\boldsymbol{m}_\lambda$ and $\boldsymbol{f}_\lambda$ are the λ-th channel of a measurement and its corresponding image, respectively. We also denote the full measurements vector by $\boldsymbol{m} = [\boldsymbol{m}_1{}^\top \ldots \boldsymbol{m}_\Lambda{}^\top]^\top$ and the full hyperspectral image (HSI) by $\boldsymbol{f} = [\boldsymbol{f}_1{}^\top \ldots \boldsymbol{f}_\Lambda{}^\top]^\top$.

In the next sections, we explore the approaches for reconstruction of the hyperspectral $\boldsymbol{f}$ from the measurement vector $\boldsymbol{m}$. Our final architecture of the proposed 3D denoised completion network for the reconstruction of a single-pixel hyperspectral images is presented in Fig. 1(a).

3.1. 2D denoised completion network

A naive approach consists in processing each spectral channel independently, which can be summarized by

$$\tilde{\boldsymbol{f}}_\lambda = \mathcal{G}_{\boldsymbol{\theta}}^1(\boldsymbol{m}_\lambda^\alpha), \quad 1 \leq \lambda \leq \Lambda, \tag{8a}$$

$$\hat{\boldsymbol{f}_\lambda} = (\mathcal{G}_{\boldsymbol{\theta}}^L \circ \ldots \circ \mathcal{G}_{\boldsymbol{\theta}}^2)(\tilde{\boldsymbol{f}}_\lambda), \quad 1 \leq \lambda \leq \Lambda, \tag{8b}$$

where $\mathcal{G}^1$ implements the denoised completion solution given by Eq. (6). Note that each spectral channel is processed independently. In particular, the layers of $(\mathcal{G}_{\boldsymbol{\theta}}^L \circ \ldots \circ \mathcal{G}_{\boldsymbol{\theta}}^2)$ rely on 2D convolution kernels that act in the spatial domain only and that are trained using grayscale images.

3.2. 3D denoised completion network

To exploit the correlation between the spectral channels of a hyperspectral image, we introduce a 3D neural network. After the reconstruction in a per-channel manner using Eq. (6), we introduce higher-order convolutions into the image domain

$$\tilde{\boldsymbol{f}}_\lambda = \mathcal{G}_{\boldsymbol{\theta}}^1(\boldsymbol{m}_\lambda^\alpha), \quad 1 \leq \lambda \leq \Lambda, \tag{9a}$$

$$\hat{\boldsymbol{f}} = \mathcal{G}_{\boldsymbol{\theta}}(\tilde{\boldsymbol{f}}). \tag{9b}$$

where the learnable layers of $\mathcal{G}_{\boldsymbol{\theta}}$ act across all of the three dimensions of a hypercube, regularizing both the spatial and spectral dimensions contrary to the 2D network.

3.3. Estimation of the mean and covariance matrices

The first steps of the denoised completion network, given by Eq. (6), require the mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ of the data, as well as the noise covariance $\boldsymbol{\Sigma}_\alpha$. As described in [27], we compute the mean $\boldsymbol{\mu}$ and covariance $\boldsymbol{\Sigma}$ from the STL-10 database prior to the training phase and estimate the noise covariance from the raw measurements. Here, we have

$$\boldsymbol{\Sigma}_{\alpha_\lambda} = \mathrm{Diag}(\boldsymbol{\sigma}_{\alpha_\lambda}^2) = \frac{1}{K\alpha_\lambda^2}\mathrm{Diag}\left(\boldsymbol{m}_{+\lambda}^\alpha + \boldsymbol{m}_{-\lambda}^\alpha\right) - \rho\frac{2\mu_{\mathrm{dark}}}{\alpha_\lambda^2 K} + \rho\frac{2\sigma_{\mathrm{dark}}^2}{\alpha_\lambda^2 K^2}, \tag{10}$$

where $\mathrm{Diag}(\boldsymbol{\sigma}_{\alpha_\lambda}^2)$ refers to the diagonal matrix with the diagonal coefficients being the elements of the noise variance $\boldsymbol{\sigma}_{\alpha_\lambda}^2$, which depends on the spectral channel, as different channels of the same measurement contain different numbers of photons. The experimental parameters μ_{dark}, σ_{dark} and K can be estimated as described in [39]. The parameter ρ corresponds to the number of raw channels that are binned to produce each of the Λ spectral channels.

Optics EXPRESS

4. Experiments

We compare the proposed 3D denoised completion network (DC-UNet-3) against the 2D denoised completion network (DC-UNet-2) and the Tikhonov solution given by Eq. (6). We also investigate the property of the 2D UNet to obtain multichannel output from multichannel input by applying a convolutional kernel with the same number of input channels as the input data. With this method, hereafter named as DC-UNet-2ch, we consider regularization across both spectral and spatial dimensions via the 2D multichannel convolutional kernels. Finally, we consider the method described in [22] (DL-FSPI). The DL-FSPI framework is developed for 2D Fourier single-pixel images, but we use the reconstruction encoder-decoder network proposed by the authors as the postprocessing step and build our denoised completion network upon it. This way we are able to compare our proposed network with another approach and explore different architecture for the regularization stage. We dub this method as DC-DL-FSPI.

For DC-UNet-2 and DC-UNet-2ch, we use a customized 2D UNet architecture (with 13,394,177 learned parameters) based on the one described in [41] and presented in Fig. 1(b). For DC-UNet-3, we employ a 3D UNet architecture (with 16,320,257 learned parameters) described in [42] and presented in Fig. 1(c). The number of learned parameters in the encoder-decoder DL-FSPI network is 317,121.

In detail, the 2D UNet consists of a contracting path and an expansive path. Starting from initial convolution layer with a number of feature channels equal to 64, with each max-pooling procedure, except the last one, this number increases by the factor of 2, reaching the value of 512 at the last convolutional layer of the contracting path. During expansion, the number of feature channels decreases accordingly. The steps of the expanding path of the 2D UNet consist of bilinear upsampling of the feature maps, concatenation with feature maps from the contracting path and two convolutional layers. Each convolutional layer, except the last one, employs kernel of size 3×3 and is followed by batch normalization and a ReLU layer. The last layer, that produces the network output, consists of one convolutional layer with the kernel of size 1×1.

The implemented 3D UNet repeats the 2D UNet structure, however, each convolutional layer, except the last one, employs kernel of size $3 \times 3 \times 3$. The last layer, that produces the network output, consists of one convolutional layer with the kernel of size $1 \times 1 \times 1$. Starting from initial convolution layer with the number of feature channels equal to 32, this number doubles within each convolution layer, reaching the value of 512 in the last layer of the contracting path. In this network, the nearest neighbor algorithm is used for upsampling.

The architecture of the network used for DC-DL-FSPI is described in details in [22].

4.1. Training of the networks

To train our networks, we use the STL-10 database [43] which consists of 105,000 color images corresponding to the 'train' and 'unlabeled' subsets. The original 96×96 images are resized to 64×64 using bicubic transform. We create a dataset of hyperspectral images from the RGB images $[\boldsymbol{f}^{\mathrm{r}}, \boldsymbol{f}^{\mathrm{g}}, \boldsymbol{f}^{\mathrm{b}}] \in \mathbb{R}^{N \times 3}$ by combining the three channels for every wavelength λ as

$$\boldsymbol{f}_{\lambda} = \boldsymbol{f}^{\mathrm{r}} \eta_{\lambda}^{\mathrm{r}} + \boldsymbol{f}^{\mathrm{g}} \eta_{\lambda}^{\mathrm{g}} + \boldsymbol{f}^{\mathrm{b}} \eta_{\lambda}^{\mathrm{b}} \tag{11}$$

where $\eta_{\lambda}^{\mathrm{r}} \in \mathbb{R}$, $\eta_{\lambda}^{\mathrm{g}} \in \mathbb{R}$, and $\eta_{\lambda}^{\mathrm{b}} \in \mathbb{R}$ are spectral weights, chosen here to represent the spectral sensitivity of the long-, medium- and short-wavelength human photo-receptors, respectively [44,45]. As such, we create HSIs with $\Lambda = 10$ spectral channels.

During training, we simulate the raw data from the hyperspectral dataset according to Eq. (1), where the system gain is set to one and the dark noise is set to zero for simplicity. However, the actual system gain and dark noise of our system are taken into account through Eq. (10) for the reconstruction of experimental data.

We implement the full network training using PyTorch [46] and optimize it using Adam [47], with an initial learning rate of 10^{-3} that is divided by 5 every 20 epochs. We use a maximum of

Optics EXPRESS

100 epochs. We trained our proposed 3D reconstructor with an NVIDIA Quadro RTX 6000, and training process took 28 h 51 min on 4 GPUs. The two dimensional neural networks were trained with a NVIDIA GeForce GTX 1080Ti, and training took 10 h 35 min on 2 GPUs for DC-UNet-2 and DC-UNet-2ch and 7 h 50 min on 2 GPUs for DC-DL-FSPI. To accelerate and stabilize the training process [48], we normalize our images, both grayscale and hyperspectral, in the range [-1;1].

4.2. Experimental data

We consider the single-pixel camera described in [11] and the open single-pixel hyperspectral imaging (SPIHIM) dataset [35]. The SPIHIM dataset contains different objects acquired using 4096 Hadamard patterns of dimension $N = 64 \times 64$ for 2048 spectral channels in the range [317, 1064]. The objects are available for different light intensities corresponding to different neutral densities placed between the illumination lamp and the object.

In this work, we select four grayscale objects: the STL-10 cat, the Siemens star resolution target [49], the off-centered Siemens star target, and the LED lamp. For each object, we obtain a high noise dataset by placing a neutral optical density behind the lamp to reduce the light flux. We consider an optical density of 0.6 for the STL-10 cat and of 1.3 for the other objects. We choose an integration time of 8 ms for the centered star target and to 4 ms for all other objects. We also consider a color Siemens star resolution target where the colors are taken from the hue color wheel.

Finally, we down-sample the full measurement vector by retaining only $M = 1024$ measurements for all objects. We only retain the spectral channels in the visible range [390, 700] nm that we merge into $\Lambda = 10$ new spectral channels by summing $\rho = 77$ original channels, *i.e.*, each new spectral channel contains 77 raw spectral channels. The ground truth images are obtained from the full measurement vectors with the lowest neutral optical density, either 0 or 0.3.

4.3. Evaluation metrics

To assess the performance of the reconstruction methods quantitatively, we use the standard peak signal-to-noise ratio (PSNR), structural similarity index measure (SSIM) metrics, implemented in scikit-image library [50], the feature similarity index measure (FSIM) [51], and the spectral angle mapper (SAM) [52] metrics implemented in image-quality-assessment library [53]. For experimental data, we estimate the ground truth as $\boldsymbol{f} = \boldsymbol{H}^{\top}\boldsymbol{y}_{\mathrm{gt}}$, where $\boldsymbol{y}_{\mathrm{gt}}$ is a fully sampled (*i.e.*, $M = N$) measurement vector with high signal-to-noise ratio. Before computing metrics, we normalize the ground truth images in the range [-1;1]. To compute the FSIM and the SAM, we normalize both the ground truth and all the computed images in the range [0;255].

We compute the PSNR, SSIM, and FSIM for hyperspectral images $\hat{\boldsymbol{f}}$ as the mean across all channels. The SAM is computed as the spectral angle between each pixel of the reconstructed image and ground truth and then is averaged across all the pixels.

5. Results and discussion

5.1. Reconstruction from simulated data

We evaluate the Tikhonov, DC-UNet-2, DC-UNet-2ch, DC-DL-FSPI, and DC-UNet-3 methods on the (unseen) images of the test set of our hyperspectral database, for three different image intensities α (150, 200 and 500 photons) corresponding to different noise levels. Note that in the case of DC-UNet-2 the input to the postprocessing network has the shape of $1 \times 64 \times 64$, while in the case of DC-UNet-2ch the input to the postprocessing network has the shape of $10 \times 64 \times 64$. Here the first number corresponds to the number of input channels and the last two values correspond to the spatial size. In Table 1, we report the PSNR, SSIM, FSIM, and SAM obtained from the networks that are trained for the corresponding intensities α. In Fig. 2

Optics EXPRESS

we display ground truths and reconstruction results along with the per-channel SSIM for the first and the last channels of simulated hyperspectral cubes acquired with $\alpha = 150$.

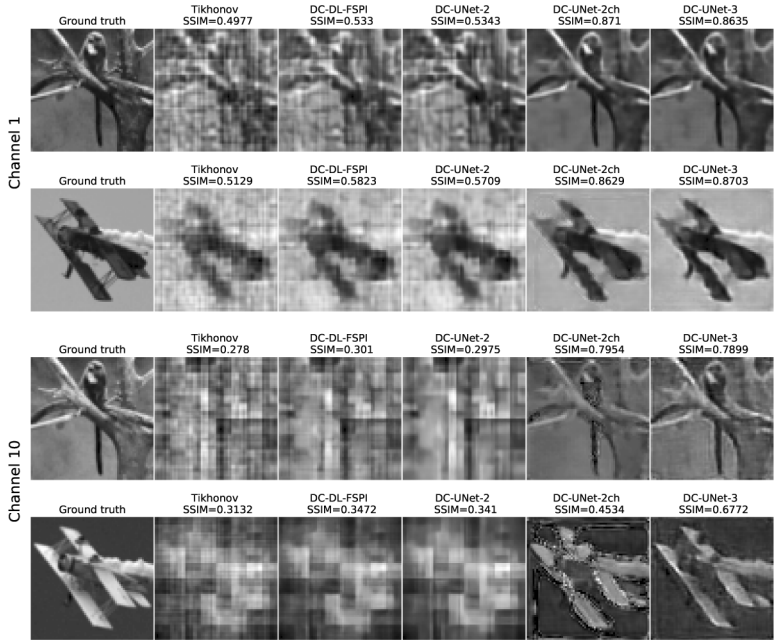


Fig. 2. Reconstruction of the simulated acquisition of two hyperspectral cubes with $\alpha = 150$. Two top rows: reconstructions of the 1st channel; two bottom rows: reconstructions of the 10th channel. We compare the reconstruction results using five methods: Tikhonov, DC-DL-FSPI, DC-UNet-2, DC-UNet-2ch, DC-UNet-3.

Table 1. PSNR, SSIM, FSIM, and SAM of simulated reconstructed images under various noise levels (the entire HSI $\alpha = 150, 200, 500$) on the STL-10 dataset. 'Tikhonov' refers to the generalised Tikhonov regularisation (Eq. (6)), 'DC-DL-FSPI' refers to the denoised completion network with a posprocessing network from [22] applied per channel, 'DC-UNet-2' refers to the denoised completion network with the UNet using 2D convolutions applied per channel (Sec.2.3), 'DC-UNet-2ch' refers to the denoised completion network with the UNet using 2D convolutions applied to entire HSI, 'DC-UNet-3' refers to the denoised completion network with the UNet using 3D convolutions (Sec.3.2). The highest metrics are highlighted in bold.

α		Tikhonov	DC-DL-FSPI	DC-UNet-2	DC-UNet-2ch	DC-UNet-3
150	PSNR	21.91 ± 2.06	23.06 ± 1.64	22.90 ± 2.47	27.49 ± 1.97	**27.70 ± 1.98**
	SSIM	.6699 ± .0724	.7201 ± .0609	.7311 ± .0626	.8861 ± .0366	**.8908 ± .0366**
	FSIM	.6987 ± .0460	.7325 ± .0313	.7481 ± .0322	.8389 ± .0230	**.8407 ± .0229**
	SAM	.0896 ± .0465	.0799 ± .0354	.0685 ± .0286	**.0228 ± .0117**	.0244 ± .0138
200	PSNR	22.37 ± 2.17	23.50 ± 1.68	23.33 ± 2.58	28.14 ± 2.09	**28.68 ± 2.20**
	SSIM	.6947 ± .0711	.7409 ± .0592	.7504 ± .0612	.9001 ± .0335	**.9111 ± .0357**
	FSIM	.7141 ± .0448	.7452 ± .0307	.7610 ± .0321	.8512 ± .0240	**.8571 ± .0250**
	SAM	.0833 ± .0446	.0748 ± .0343	.0643 ± .0281	.0221 ± .0125	**.0206 ± .0107**
500	PSNR	23.66 ± 2.53	24.72 ± 1.85	24.58 ± 2.95	28.22 ± 2.11	**29.31 ± 2.26**
	SSIM	.7639 ± .0631	.7608 ± .0611	.8065 ± .0559	.9008 ± .0303	**.9239 ± .0313**
	FSIM	.7579 ± .0389	.7798 ± .0290	.7988 ± .0296	.8524 ± .0212	**.8692 ± .0231**
	SAM	.0641 ± .0388	.0605 ± .0317	.0488 ± .0201	.0201 ± .0117	**.0186 ± .0109**

Optics EXPRESS

We observe that both methods that consider spatial-spectral information (DC-UNet-2ch and DC-UNet-3) achieve significantly higher reconstruction metrics w.r.t. the methods that take into account spatial information only (Tikhonov, DC-DL-FSPI, and DC-UNet-2). Achieving the best reconstruction metrics, DC-UNet-3 outperforms DC-UNet-2 and DC-DL-FSPI by at least 4.6 dBs and DC-UNet-2ch by at least 0.2 dBs.

From the results presented in Fig. 2 we observe that while Tikhonov, DC-DL-FSPI, and DC-UNet-2 fail to reconstruct information from the first and the last channels of HSI, introduction of spatial-spectral correlation helps DC-UNet-2ch and DC-UNet-3 to recover these channels. Inclusion of higher order convolutions allows DC-UNet-3 to perform the best reconstruction of image details.

5.2. Reconstruction from experimental data

Figure 3 and Fig. 4 illustrate the performance of the five methods for recovering the four experimental datasets described in 4.2. We plot the spectral curves of a small region in Fig. 5. We observe that the inclusion of the artifact correction network in the image domain produces less noisy and smoother images than those provided by the Tikhonov method. Since DC-DL-FSPI and DC-UNet-2 act in the spatial domain, they tend to improve images w.r.t. the Tikhonov method in each spectral channel separately. Despite some remaining artifacts, DC-UNet-2ch and

Fig. 3. Reconstruction of the experimental acquisition of two objects: STL-10 cat (columns 1–3) and off-centered Siemens star (columns 4–6). The top row displays the ground truth images for the 2nd, 5th and 9th channels. We compare the reconstruction results using five methods: Tikhonov, DC-DL-FSPI, DC-UNet-2, DC-UNet-2ch, DC-UNet-3.

DC-UNet-3 allow to restore sharper image details in all spectral channels with DC-UNet-3 being clearly superior. This is especially noticeable on the border channels – cat face in STL-10 cat,

Optics EXPRESS

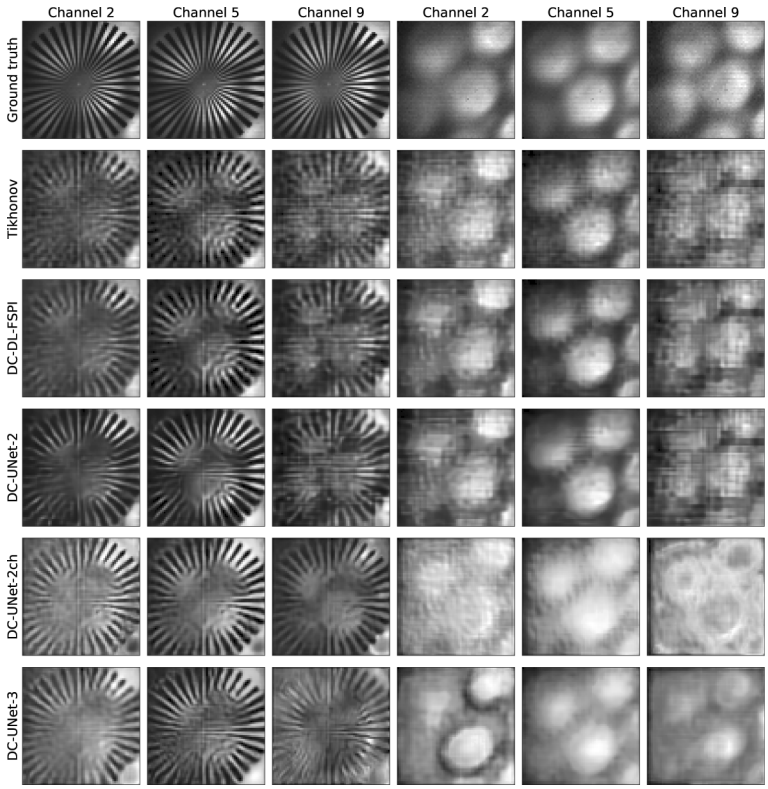


Fig. 4. Reconstruction of the experimental acquisition of two objects: centered Siemens star (columns 1–3) and LED lamp (columns 4–6). The top row displays the ground truth images for the 2nd, 5th and 9th channels. We compare the reconstruction results using five methods: Tikhonov, DC-DL-FSPI, DC-UNet-2, DC-UNet-2ch, DC-UNet-3.

letters in the off-centered Siemens star resolution target, and the Siemens star beams are closer to those of the ground truth images. This is achieved thanks to the spatial-spectral regularization nature and higher order convolutions of the postprocessing 3D network.

From the analysis of the spectral curves presented in Fig. 5 we observe, that while per-channel based methods (Tikhonov, DC-DL-FSPI, and DC-UNet-2) are closer to the ground truth images in terms of reflectance values in the border channels, DC-UNet-3 demonstrates closer values to the ground truths in all other channels. Notable that per-channel based methods (Tikhonov, DC-DL-FSPI, and DC-UNet-2) have the identical shape of the curve. From the first row, first column, and the third row, second column in Fig. 5 it is seen that DC-UNet-3 tends to show the most accurate repetition of the ground truth spectrum shape. DC-UNet-2ch, in turn, is often far from the ground truth (Fig. 5, first row, second row, first column, third row, first column).

DC-UNet-3 achieves superior results in terms of noise reduction compared to all other methods and better preservation of details compared to DC-UNet-2ch. We observe that taking into account spectral correlation between adjacent channels helps to recover channels that initially contain less information. This can be seen clearly from the off-centered Siemens star resolution target, where the inscription is preserved in all spectral channels in contradistinction to the results obtained with Tikhonov, DC-UNet-2, and DC-DL-FSPI.

Figure 6 presents the reconstruction of the color Siemens star target recovery using Tikhonov, DC-DL-FSPI, DC-UNet-2, DC-UNet-2ch, and DC-UNet-3 methods. The Tikhonov solution presents some artifacts such as the square-like structures on the frontier of the branches that stems

Optics EXPRESS

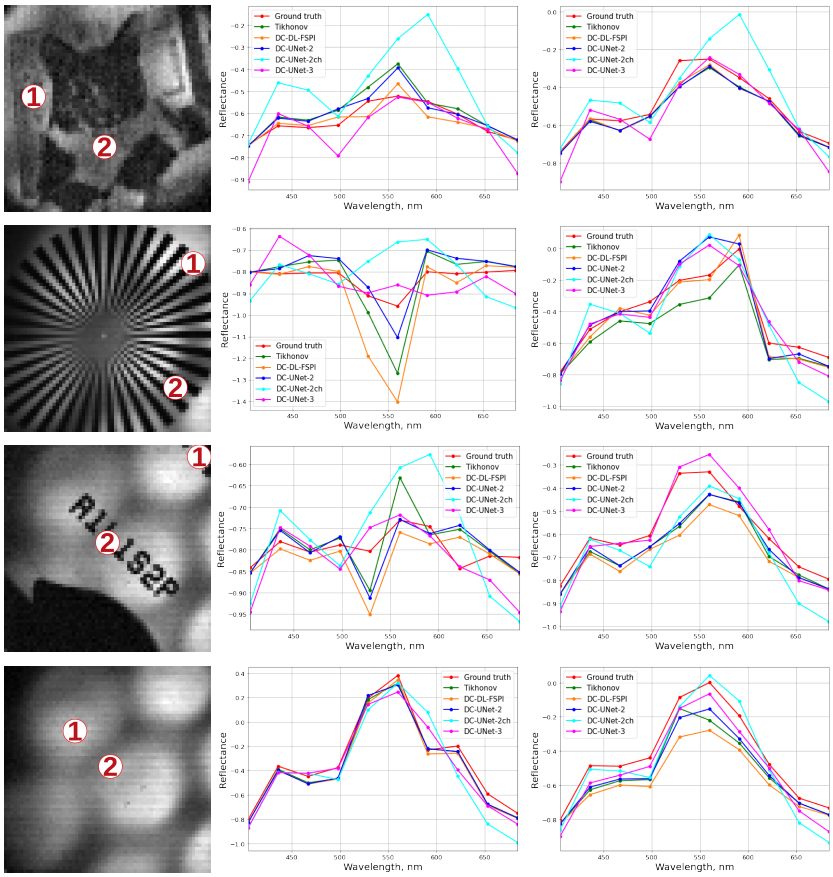


Fig. 5. Spectral curves of the experimental reconstructions. Left column: 5th channels of the ground truth images. Middle column: spectral curves for five methods and the reference, corresponding to the point 1 on the image. Right column: spectral curves for five methods and the reference, corresponding to the point 2 on the image.

from downsampling in the Hadamard basis, or the grain-like appearance from the photon-based noise, whereas reconstruction with all other methods results in smoother images. Additionally, DC-DL-FSPI tends to provide less noisy images than DC-UNet-2. Compared to DC-DL-FSPI and DC-UNet-2, DC-UNet-2ch and DC-UNet-3 are able to retain certain details that are lost when removing the artifacts. This can be observed in 'Channel 4' and 'Channel 6', where the upper branches of the star were mostly erased and lost their alignment when reconstructed using channel-based methods, but are mostly intact after recovery with methods considering the entire HSI. Furthermore, in 'Channel 9', DC-DL-FSPI and DC-UNet-2 almost erased the bottom left branches, as opposed to the results provided by DC-UNet3. In turn, DC-UNet-2ch, introduced unwanted artifacts into all channels. This indicates that DC-UNet-3 shows itself to be the best at exploiting redundancy across spectral channels. To get a visual comparison of the hyperspectral images, we also compute a pseudo color image as described in [54]. Comparing the pseudo color images in the bottom row of Fig. 6, one can see that all five methods succeed in color rendering. However, all methods give a slight shift in the color spectrum. The red branches reconstructed via Tikhonov, DC-DL-FSPI, and DC-UNet-2 appear to be biased towards red wavelengths, while the recovery with DC-UNet-2ch and DC-UNet-3 causes the red branches to

Optics EXPRESS

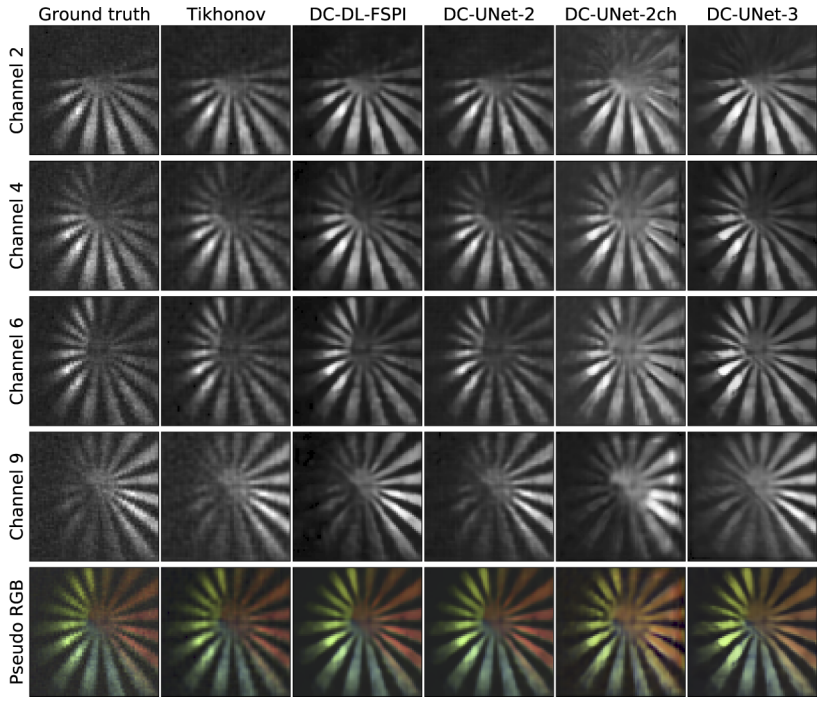


Fig. 6. Reconstruction of the experimental color Siemens target with five methods: Tikhonov, DC-DL-FSPI, DC-UNet-2, DC-UNet-2ch, and DC-UNet-3. Results are shown for the 2nd, 4th, 6th, and 9th channels. Bottom row displays pseudo color image computed with the corresponding method recovery result according to [54].

shift towards orange wavelengths. A similar situation is observed in the blue branches of the star, where Tikhonov, DC-DL-FSPI, and DC-UNet-2 applications result into the shift towards the lower blue wavelengths, while reconstruction via DC-UNet-2ch and DC-UNet-3 shifts the branches into the upper-blue or lower-green wavelengths. However, in terms of spatial resolution, the branches produced by DC-UNet-3 appear to be of the highest resolution, in particular in the upper left branches.

5.3. Limitations

The proposed three-dimensional network regularizes the reconstruction in both the spatial and the spectral dimensions, being capable of filling in the missing information and of improving the reconstruction quality. Still, there are some limitations to the proposed reconstruction pipeline.

The first limitation stems from a small number of spectral channels used in our study. Given that the conventional hyperspectral applications may require additional channels, one should carefully explore whether the spectral correlations between the channels can still be exploited when there is a number of them. Second, in our work we considered small images of size 64×64. While there is no specific limitation on application of the proposed method to larger images, training of the network in this case might require additional GPU memory, as the covariance matrix scales with the square of the image size. This issue may be addressed by discarding the rows and columns corresponding to small correlations. Another limitation is associated with the choice of the architecture of the neural network. Here, we focused on the UNet because it has demonstrated its effectiveness in many problems across many disciplines. However, this architecture could prove too heavy if larger convolutional filters are required (to exploit

Optics EXPRESS

additional channels or dimensions, *e.g.*, third spatial coordinate or time). In this case, the choice of other architectures may be more sensible, *e.g.* the one proposed in [22] and explored within DC-DL-FSPI. Lastly, the proposed DC-UNet-3 tends to correlate the adjacent spectral channels during the reconstruction. While this feature can be considered as an advantage for an object with a smooth spectra (*e.g.*, the cat's face or the off-centered Siemens target in Fig. 4), it could be a disadvantage if there are spectral discontinuities (*e.g.*, the color Siemens target in Fig. (6)). Training a 3D DC-Net using a hyperspectral image database with distinct spectral features should alleviate this issue.

6. Conclusion

We proposed the 3D deep neural network for reconstructing hyperspectral images from a small number of noisy single-pixel measurements. The layers of the proposed network combine the classical Tikhonov solution with learnable convolution filters. To exploit the correlation between the spectral channels, it considers the filters that act in both the spatial and the spectral dimensions. Our experiments show that the 3D UNet improves the reconstruction quality compared to the 2D UNet that acts in the image domain only and the 2D UNet that is applied in spatial and spectral domains. The proposed 3D network demonstrates superior performance in terms of quantitative metrics as well as visually. This is particularly evident in the channels with low photon counts, *i.e.*, with the overwhelming noise and lack of information. In the future work, we plan to consider spectral discontinuities, additional dimensions, and other architectures of the base network.

Funding. Agence Nationale de la Recherche (ANR-11-LABX-0063, ANR-17-CE19-0003).

Acknowledgments. This material is based upon work done on the PILoT facility (PILoT, INSA LYON,Bât. Blaise Pascal, 7 Av. Jean Capelle 69621 Villeurbanne). The authors acknowledge the support of Ostrogradski Scholarship from the Embassy of France in the Russian Federation awarded to Valeriya Pronina.

Disclosures. The authors declare that they have no conflicts of interest.

Data availability. Data underlying the results presented in this paper are available in Refs. [35,43].

References

1. G. M. Gibson, S. D. Johnson, and M. J. Padgett, "Single-pixel imaging 12 years on: a review," Opt. Express **28**(19), 28190–28208 (2020).
2. V. Studer, J. Bobin, M. Chahid, H. S. Mousavi, E. Candes, and M. Dahan, "Compressive fluorescence microscopy for biological and hyperspectral imaging," Proc. Natl. Acad. Sci. **109**(26), E1679–E1687 (2012).
3. E. Aguénounon, F. Dadouche, W. Uhring, N. Ducros, and S. Gioux, "Single snapshot imaging of optical properties using a single-pixel camera: a simulation study," J. Biomed. Opt. **24**(7), 071612 (2019).
4. Q. Pian, R. Yao, N. Sinsuebphon, and X. Intes, "Compressive hyperspectral time-resolved wide-field fluorescence lifetime imaging," Nat. Photonics **11**(7), 411–414 (2017).
5. Y. Zhang, G. Gibson, M. Edgar, G. Hammond, and M. Padgett, "Dual-band single-pixel telescope," Opt. Express **28**(12), 18180–18188 (2020).
6. F. Li, M. Zhao, Z. Tian, F. Willomitzer, and O. Cossairt, "Compressive ghost imaging through scattering media with deep learning," Opt. Express **28**(12), 17395–17408 (2020).
7. Z. Zhang, S. Liu, J. Peng, M. Yao, G. Zheng, and J. Zhong, "Simultaneous spatial, spectral, and 3d compressive imaging via efficient fourier single-pixel measurements," Optica **5**(3), 315–319 (2018).
8. F. Rousset, N. Ducros, F. Peyrin, G. Valentini, C. D'Andrea, and A. Farina, "Time-resolved multispectral imaging based on an adaptive single-pixel camera," Opt. Express **26**(8), 10550–10558 (2018).
9. C. Tao, H. Zhu, X. Wang, S. Zheng, Q. Xie, C. Wang, R. Wu, and Z. Zheng, "Compressive single-pixel hyperspectral imaging using rgb sensors," Opt. Express **29**(7), 11207–11220 (2021).
10. J. Peller, F. Farahi, and S. Trammell, "Hyperspectral imaging system based on a single-pixel camera design for detecting differences in tissue properties," Appl. Opt. **57**(27), 7651–7658 (2018).
11. A. Lorente Mur, B. Montcel, F. Peyrin, and N. Ducros, "Deep neural networks for single-pixel compressive video reconstruction," Proc. SPIE **11351**, 113510S (2020).
12. S. Jin, W. Hui, Y. Wang, K. Huang, Q. Shi, C. Ying, D. Liu, Q. Ye, W. Zhou, and J. Tian, "Hyperspectral imaging using the single-pixel fourier transform technique," Sci. Rep. **7**(1), 45209 (2017).
13. M. Edgar, G. Gibson, and M. Padgett, "Principles and prospects for single-pixel imaging," Nat. Photonics **13**(1), 13–20 (2019).
14. M. F. Duarte, M. Davenport, D. Takhar, J. Laska, T. Sun, K. Kelly, and R. Baraniuk, "Single-pixel imaging via compressive sampling," IEEE Signal Process. Mag. **25**(2), 83–91 (2008).

Optics EXPRESS

15. V. Pronina, F. Kokkinos, D. V. Dylov, and S. Lefkimmiatis, "Microscopy image restoration with deep wiener-kolmogorov filters," in *Computer Vision – ECCV 2020*, (Springer International Publishing, Cham, 2020), pp. 185–201.
16. L. Wang, C. Sun, Y. Fu, M. H. Kim, and H. Huang, "Hyperspectral image reconstruction using a deep spatial-spectral prior," in *Proceedings of IEEE/CVF Conference on Computer Vision and Pattern Recognition*, (CVPR, 2019), pp. 8024–8033.
17. C. Jiang, H. Zhang, L. Zhang, H. Shen, and Q. Yuan, "Hyperspectral image denoising with a combined spatial and spectral weighted hyperspectral total variation model," Can. J. Remote Sens. **42**(1), 53–72 (2016).
18. R. Yao, M. Ochoa, P. Yan, and X. Intes, "Net-flics: fast quantitative wide-field fluorescence lifetime imaging with compressed sensing – a deep learning approach," Light: Sci. Appl. **8**(1), 26 (2019).
19. J. T. Smith, M. Ochoa, and X. Intes, "Unmix-me: spectral and lifetime fluorescence unmixing via deep learning," Biomed. Opt. Express **11**(7), 3857–3874 (2020).
20. M. Lyu, W. Wang, H. Wang, H. Wang, G. Li, N. Chen, and G. Situ, "Deep-learning-based ghost imaging," Sci. Rep. **7**(1), 17865 (2017).
21. M.-J. Sun and J.-M. Zhang, "Single-pixel imaging and its application in three-dimensional reconstruction: A brief review," Sensors **19**(3), 732 (2019).
22. S. Rizvi, J. Cao, K. Zhang, and Q. Hao, "Improving imaging quality of real-time fourier single-pixel imaging via deep learning," Sensors **19**(19), 4190 (2019).
23. C. F. Higham, R. Murray-Smith, M. Padgett, and M. P. Edgar, "Deep learning for real-time single-pixel video," Sci. Rep. **8**(1), 2369 (2018).
24. I. Hoshi, T. Shimobaba, T. Kakue, and T. Ito, "Single-pixel imaging using a recurrent neural network combined with convolutional layers," Opt. Express **28**(23), 34069–34078 (2020).
25. J. Li, Y. Li, J. Li, Q. Zhang, and J. Li, "Single-pixel compressive optical image hiding based on conditional generative adversarial network," Opt. Express **28**(15), 22992–23002 (2020).
26. N. Ducros, A. Mur, and F. Peyrin, "A completion network for reconstruction from compressed acquisition," in *Proccedings of IEEE 17th International Symposium on Biomedical Imaging*, (ISBI, 2020), pp. 619–623.
27. A. Lorente Mur, P. Leclerc, F. Peyrin, and N. Ducros, "Single-pixel image reconstruction from experimental data using neural networks," Opt. Express **29**(11), 17097–17110 (2021).
28. F. Arias, H. Sierra, and E. Arzuaga, "A framework for an artificial neural network enabled single pixel hyperspectral imager," in *Proceedings of 10th Workshop on Hyperspectral Imaging and Signal Processing: Evolution in Remote Sensing*, (WHISPERS, 2019), pp. 1–5.
29. O. Sidorov and J. Y. Hardeberg, "Deep hyperspectral prior: Single-image denoising, inpainting, super-resolution," in *Proceedings of the IEEE/CVF International Conference on Computer Vision Workshops*, (ICCV, 2019), pp. 3844–3851.
30. Q. Shi, X. Tang, T. Yang, R. Liu, and L. Zhang, "Hyperspectral image denoising using a 3-d attention denoising network," *IEEE Transactions on Geoscience and Remote Sensing* pp. 1–16 (2021).
31. K. Wei, Y. Fu, and H. Huang, "3-d quasi-recurrent neural network for hyperspectral image denoising," IEEE Trans. Neural Netw. Learning Syst. **32**(1), 363–375 (2021).
32. L. Wang, T. Zhang, Y. Fu, and H. Huang, "Hyperreconnet: Joint coded aperture optimization and image reconstruction for compressive hyperspectral imaging," IEEE Trans. on Image Process. **28**(5), 2257–2270 (2019).
33. D. Gedalin, Y. Oiknine, and A. Stern, "Deepcubenet: reconstruction of spectrally compressive sensed hyperspectral images with deep neural networks," Opt. Express **27**(24), 35811–35822 (2019).
34. A. Lorente Mur, F. Peyrin, and N. Ducros, "Single-Pixel pYthon Image Reconstruction Toolbox (spyrit) Version 1.0," https://github.com/todo (2020).
35. A. Lorente Mur and N. Ducros, "Single-Pixel Hyperspectral Imaging dataset Version 1.0," Github, 2020, https://gitlab.in2p3.fr/nicolas.ducros/spihim.
36. L. Baldassarre, Y. Li, J. Scarlett, B. Gözcü, I. Bogunovic, and V. Cevher, "Learning-based compressive subsampling," IEEE J. Sel. Top. Signal Process. **10**(4), 809–822 (2016).
37. F. Rousset, N. Ducros, A. Farina, G. Valentini, C. D'Andrea, and F. Peyrin, "Adaptive basis scan by wavelet prediction for single-pixel imaging," IEEE Trans. Comput. Imaging **3**(1), 36–46 (2017).
38. A. Foi, M. Trimeche, V. Katkovnik, and K. Egiazarian, "Practical poissonian-gaussian noise modeling and fitting for single-image raw-data," IEEE Trans. on Image Process. **17**(10), 1737–1754 (2008).
39. M. Rosenberger, C. Zhang, P. Votyakov, M. Preissler, and G. Notni, "Emva 1288 camera characterisation and the influences of radiometric camera characteristics on geometric measurements," ACTA IMEKO **5**(4), 81 (2016).
40. A. Lorente Mur, M. Ochoa, J. E. Cohen, X. Intes, and N. Ducros, "Handling negative patterns for fast single-pixel lifetime imaging," Proc. SPIE **10862**, 108620A (2019).
41. O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015*, N. Navab, J. Hornegger, W. M. Wells, and A. F. Frangi, eds. (Springer International Publishing, Cham, 2015), pp. 234–241.
42. Ö. Çiçek, A. Abdulkadir, S. S. Lienkamp, T. Brox, and O. Ronneberger, "3d u-net: Learning dense volumetric segmentation from sparse annotation," in *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2016*, S. Ourselin, L. Joskowicz, M. R. Sabuncu, G. Unal, and W. Wells, eds. (Springer International Publishing, Cham, 2016), pp. 424–432.

Optics EXPRESS

43. A. Coates, A. Ng, and H. Lee, "An analysis of single-layer networks in unsupervised feature learning," in *Proceedings of the Fourteenth International Conference on Artificial Intelligence and Statistics*, vol. 15 of *Proceedings of Machine Learning Research*, 2011), pp. 215–223.
44. A. Stockman and L. Sharpe, "Cone spectral sensitivities and color matching," in *Color vision: From Genes to Perception*, K. Gegenfurtner and L. T. Sharpe, eds. (Cambridge University, 1999), chap. 2, pp. 53–87.
45. D. Foster, "Chromatic function of the cone," in *Encyclopedia of the Eye*, D. A. Dartt, J. C. Besharse, R. Dana, and P. Bex, eds. (Academic, 2010), pp. 266–274.
46. A. Paszke, S. Gross, F. Massa, A. Lerer, J. Bradbury, G. Chanan, T. Killeen, Z. Lin, N. Gimelshein, L. Antiga, A. Desmaison, A. Köpf, E. Yang, Z. DeVito, M. Raison, A. Tejani, S. Chilamkurthy, B. Steiner, L. Fang, J. Bai, and S. Chintala, "Pytorch: An imperative style, high-performance deep learning library," in *Advances in Neural Information Processing Systems 32*, (Curran Associates, Inc., 2019), pp. 8024–8035.
47. D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," arXiv:1412.6980 [cs.LG] (2015).
48. Y. L. Cun, I. Kanter, and S. A. Solla, "Eigenvalues of covariance matrices: Application to neural-network learning," Phys. Rev. Lett. **66**(18), 2396–2399 (1991).
49. International Organization for Standardization, Geneva, CH, *Photography – Electronic still picture imaging – Resolution and spatial frequency responses* (2019), https://www.iso.org/standard/71696.html.
50. S. van der Walt, J. L. Schönberger, J. Nunez-Iglesias, F. Boulogne, J. D. Warner, N. Yager, E. Gouillart, and T. Yu, the scikit-image contributors, "scikit-image: image processing in Python," PeerJ **2**, e453 (2014).
51. L. Zhang, L. Zhang, X. Mou, and D. Zhang, "Fsim: A feature similarity index for image quality assessment," IEEE Trans. on Image Process. **20**(8), 2378–2386 (2011).
52. R. H. Yuhas, A. F. H. Goetz, and J. W. Boardman, "Discrimination among semi-arid landscape endmembers using the spectral angle mapper (sam) algorithm," in *Summaries of the 4th JPL Airborne Earth Science Workshop*, (JPL Publication, 1992), pp. 147–149.
53. M. U. Müller, N. Ekhtiari, R. M. Almeida, and C. Rieke, "Super-resolution of multispectral satellite images using convolutional neural networks,"in *ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences*, (ISPRS, 2020), pp. 33–40.
54. M. Magnusson, J. Sigurdsson, S. E. Armansson, M. O. Ulfarsson, H. Deborah, and J. R. Sveinsson, "Creating rgb images from hyperspectral images using a color matching function," in *Proceedings of 2020 IEEE International Geoscience and Remote Sensing Symposium*, (IGARSS, 2020), pp. 2045–2048.