Yair Ben Sahel, John P. Bryan, Brian Cleary,
Samouil L. Farhi, and Yonina C. Eldar

Deep Unrolled Recovery in Sparse Biological Imaging

Achieving fast, accurate results

Deep algorithm unrolling has emerged as a powerful, model-based approach to developing deep architectures that combine the interpretability of iterative algorithms with the performance gains of supervised deep learning, especially in cases of sparse optimization. This framework is well suited to applications in biological imaging, where physics-based models exist to describe the measurement process and the information to be recovered is often highly structured. Here we review the method of deep unrolling and show how it improves source localization in several biological imaging settings.

Introduction

Biological imaging that precisely labels microscopic structures offers a range of insights, from single-molecule localization microscopy (SMLM) [1] visualizing the microarchitecture of the cellular cytoskeleton, to single-molecule fluorescence in situ hybridization [2], revealing the spatial distribution of gene expression, and synaptic immunofluorescence [3], showing the distribution of neuronal connections in the brain. To succeed, each of these imaging methods is accompanied by analysis tools that identify and localize the desired signal.

In localization problems, regardless of the imaging system used, analysis pipelines attempt to undo the blurring effect of its imperfect impulse response, or point-spread function (PSF). The PSF always has finite width, leading to limited image resolution. If small objects, like the two thin microtubules seen in Figure 1, are nearer to each other than the width of the PSF, they may be difficult to distinguish. The biological signal in the field of view (FoV) of the imaging system can be represented as a high-resolution matrix, where each element's value represents the intensity of the signal at that physical location. The imaging process may be thought of as a 2D convolution of the "true" object being imaged and an array representing the PSF. The goal of a localization algorithm is to "undo" this convolution.

The localization problem is much more tractable if the images have a predictable structure because the solution space can be constrained. For example, some biological images comprise

45

similarly sized cells, elongated fibers of known width, or small, scattered fluorescent spots representing individual molecules with dimensions below the diffraction limit. This knowledge can be combined with an understanding of the physical parameters of the imaging system, such as the numerical aperture and magnification, in analysis pipelines, to identify cell centers [4], trace long fibers [1], or localize fluorophores, which have been bound to biologically relevant molecules [5]. The results of these algorithms may then be used to achieve higher-level biological goals, from generating tissue atlases, to determining genetic expression patterns or diagnosing pathologies.

Here we focus specifically on biological localization problems on sparse images, meaning that the high-resolution information to be recovered has relatively few nonzero values. Some signals, such as fluorescently tagged messenger RNA (mRNA), clusters of proteins, or microbubbles in ultrasound imaging are naturally sparse. In other cases, experimental techniques may be used to induce sparsity [1], or sparsity in bases other than the spatial domain may be exploited.

Iterative optimization techniques have emerged as one of the most powerful approaches for localizing sparse signal emitters. For example, in SMLM, a superresolution technique that relies on subpixel localization of scattered fluorophores, the original peak-finding algorithms have been outperformed by approaches using iterative convex optimization-based algorithms in terms of localization accuracy, signal-to-noise ratio, and resolution [7]. These advantages of iterative optimization techniques for sparse recovery go beyond SMLM to other biological imaging problems, providing high-accuracy localization in many settings [8]–[10]. However, they also suffer from some disadvantages. They require adjustment of optimization parameters and explicit knowledge of the impulse response of the imaging system, which restricts their use when the imaging system is not well characterized. They are computationally expensive and converge slowly, which limits their use in real-time, live-cell imaging. Finally, they are relatively inflexible: a given algorithm is designed to take advantage of a particular structure (here, signal sparsity) but ignores other context, which may be important (e.g., cell size or density).

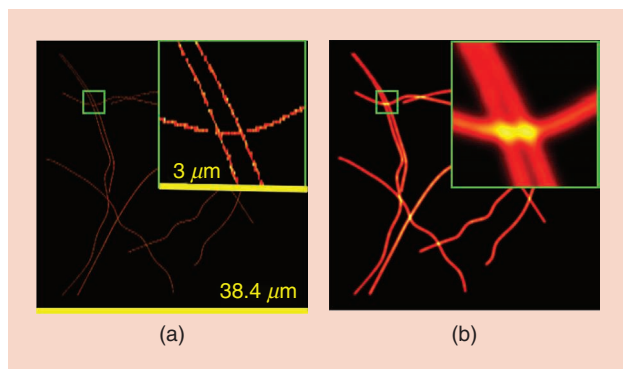Many of these disadvantages can be overcome by replacing the iterations of these algorithms with trained neural networks, which perform the same mathematical operation, a process known as *algorithmic unrolling* [11] (alternatively, *unfolding*). By doing so, parameters which would have to be specified explicitly or tuned empirically are learned automatically, and relevant context ignored by the algorithm may be incorporated into the learned model. Since its introduction a decade ago, a wide variety of techniques have been adapted using learned unrolling, enabling improvements in performance across a variety of settings [12].

In the rest of this article, we review how learned unrolling is applied to the localization of sparse sources in biological imaging data. We first formulate biological localization as a sparse recovery problem and discuss the advantages and disadvantages of iterative approaches to sparse recovery. We then describe how algorithmic unrolling addresses some of the shortcomings and how the general sparse recovery problem may be adapted to the unrolling framework. Next we show in detail how unrolling has been used to achieve fast, accurate superresolution in the optical microscopy technique SMLM. We then review a number of additional biological imaging analysis problems to which unrolling has been applied to improve performance: ultrasound localization microscopy (ULM), light-field microscopy (LFM), and cell-center localization in fluorescence microscopy. Throughout, we discuss a number of additional data analysis problems in sparse optical microscopy and propose that algorithmic unrolling be applied to achieve the same benefits obtained in the reviewed techniques.

Sparse recovery in biological imaging

The localization of biological objects, from microtubules to neural synapses, can be approached effectively as a convex optimization problem. For ease of notation, we first reframe the imaging process, typically thought of as a convolution, as a matrix-vector multiplication. The high-resolution signal is "vectorized" and then multiplied by a matrix representing the PSF. We note that the problem can be formulated and solved with 2D convolutions equally well, and the techniques described throughout the article may be applied to a 2D formulation, as in [6] and [10].

Formally, we consider the FoV as a high-resolution square grid with side length n_h. The total number of locations in this grid is $N_h = n_h^2$. This grid is vectorized to form a vector: $\mathbf{x} \in \mathbb{R}^{N_h}$. The locations of emitters in the sample may be modeled by assigning each element of $\mathbf{x}$ a value related to the number of photons emitted from that location within the FoV. If the FoV is imaged using a sensor with N_l pixels (with $N_h > N_l$), then we can model the imaging process as multiplication by matrix $\mathbf{A} \in \mathbb{R}^{N_l \times N_h}$, in which element (i, j) is the proportion of a signal emitted from location j on the high-resolution grid that will be detected at pixel i of the sensor. Thus defined, the columns of $\mathbf{A}$ represent the PSF of the imaging system such that column j of $\mathbf{A}$ is the PSF of the system for a point source at location j. The (vectorized) measured image is then $\mathbf{y} = \mathbf{Ax}$, with $\mathbf{y} \in \mathbb{R}^{N_l}$. The goal of the analysis pipeline is to infer the value of $\mathbf{x}$, given $\mathbf{y}$ and $\mathbf{A}$.

This inference problem can be formulated as a least-squares optimization problem: we seek to find

FIGURE 1. A simulated microscopic imaging of microtubules from [6]. (a) A true localization of the microtubule structure. (b) An imaged microtubule "blurred" by the microscope PSF.

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\| \mathbf{y} - \mathbf{A}\mathbf{x} \right\|_2^2. \tag{1}$$

Even if $\mathbf{A}$ is known perfectly, as long as $N_h > N_l$, $\mathbf{A}$ will have a nontrivial null space so that the optimization problem is underdetermined. Leveraging knowledge of the biological structure of $\mathbf{x}$ can resolve this issue. If, as discussed previously, it is known that $\mathbf{x}$ is sparse, then we may choose a sparse optimization technique, such as the well-known LASSO [13], [14] to recover $\mathbf{x}$:

$$\hat{\mathbf{x}} = \underset{\mathbf{x}}{\operatorname{argmin}} \left\| \mathbf{y} - \mathbf{A}\mathbf{x} \right\|_2^2 + \lambda \left\| \mathbf{x} \right\|_1. \tag{2}$$

In particular, by correctly tuning λ, minimizer $\hat{\mathbf{x}}$ of (2) will provide accurate locations of each signal-emitting object in the FoV.

Algorithmic unrolling for sparse localization

Once a problem is framed as a sparse optimization of the form (2), a number of algorithms may be used to find the minimizer $\hat{\mathbf{x}}$. The examples include the alternating direction method of multipliers (ADMMs), [15] iterative shrinkage-thresholding algorithm (ISTA) [16], and half-quadratic splitting algorithm (HQS) [17]. These methods converge to the correct minimizer $\hat{\mathbf{x}}$ but have some limiting disadvantages: slow convergence, the requirement of parameter tuning and explicit knowledge of the imaging system [6], and mathematical inflexibility.

Deep learning approaches have overcome some of these disadvantages. In analyzing SMLM data, we find that convolutional neural network (CNN) models have achieved fast, accurate superresolution [18], able to improve recovery by incorporating structures not specified by the user. Deep learning, however, comes with disadvantages of its own. In particular, deep learning is typically thought of as a black-box process: it is difficult to interpret the way the model transforms the input to obtain a result. Because of this, when inaccurate results are produced, it can be difficult to understand how to improve the model. Typical deep learning approaches are strongly dependent on the available training data, causing a lack of model robustness to new examples. Finally, when using generic network architectures, many layers and parameters are typically required for good performance.

In 2010, Gregor and LeCun proposed a method to create neural networks based on iterative methods used for sparse recovery [11], known as *algorithm unrolling*. The goal is to take advantage of both the interpretability of iterative techniques and the flexibility of learned methods. In learned unrolling, the transformation applied to the input by each iteration of the algorithm is replaced with a neural network layer, which applies the same type of function: for instance, matrix multiplication can be replaced by a fully connected layer, and thresholding can be replaced by an activation function with a learnable threshold, representing an appropriate regularizer. These iteration layers are concatenated together, and the resulting, model-based neural network is optimized using supervised learning, with training data consisting of paired examples of the signal vector $\mathbf{x}$ and measurement vector $\mathbf{y}$ from

(2). The loss function used for optimization may be customized depending on the task but is often based on the mean-squared error [11]. Training data may be obtained, for example, from measurement simulations with known ground truth, as in [6]. A forward pass through the optimized network will then perform the same operations as the iterative algorithm, with the parameters of each transformation optimized to map training input $\mathbf{y}$ to its paired signal $\mathbf{x}$.

Gregor and LeCun applied the unrolling framework to ISTA, calling the ISTA-inspired network *learned ISTA* (*LISTA*). For a given number of iterations/layers, the trained LISTA network obtains lower prediction error than ISTA and even achieves faster convergence and higher accuracy than the accelerated version of ISTA, that is, the fast iterative shrinkage-thresholding algorithm (FISTA) [11]. In "From the Iterative Shrinkage-Thresholding Algorithm to the Learned Iterative Shrinkage-Thresholding Algorithm," we detail the process of constructing the LISTA network based on ISTA.

This framework provides several key advantages. First, algorithm parameters, such as λ in (2), are learned automatically. Second, with an unrolled model, the part of the model corresponding to $\mathbf{A}$ in (2) is learned, removing the need to explicitly model the PSF. Inherent properties of the PSF, such as locality, can be incorporated by using convolutional layers in the model. Finally, although these iterative algorithms are designed to solve a specific problem, i.e., sparse recovery, the approach is general, solving all the problems of this type equally well. With a neural network, the model can learn to analyze data which may have additional structure not explainable by sparsity, thereby obtaining higher-accuracy results more quickly. Because the underlying structure of the algorithm remains intact, the network is also less prone to overfitting, which improves robustness.

The unrolling framework also has a few drawbacks. Its data-driven approach requires a substantial quantity of training data, which may be difficult to obtain. If the data used to train the network are generated differently from those being analyzed (for example, if a different microscope is used with a substantially different PSF), the recovery performance will degrade. However, it has been found that learned unrolled networks are much more robust than traditional learned neural networks to changes in the distribution of signal (for example, studying a different type of subcellular structure [6]) and require many fewer training data. The learned weights of the network may also be less interpretable than an algorithm's iterative step in which each component has explicit physical meaning. In [6], however, it is shown that the LISTA-based neural network used for the superresolution task learns transformations which are closely related to the operation of the iterative step (for instance, convolution filters are learned with shapes similar to the PSF) and are easier to understand than those learned by the purely data-driven approach of typical neural networks.

It is also important to note that, being partially data driven and using a small, fixed number of iterations, learned unrolled networks no longer explicitly solve the optimization problems

on which they are based (such as sparse recovery). Although the structure of the algorithm is maintained, the transformation applied by the network does not follow the exact steps that are guaranteed to find the minimizer of (2) and may not be applied in the image domain. So, even though learned unrolled networks have been shown to be successful in localizing spatially sparse sources, they do not necessarily explicitly solve the sparse recovery problem.

Throughout the rest of the article, we concentrate on applications of deep unrolling to the recovery of sparse biological data, specifically using unrolled networks based on ISTA. Importantly, the learned unrolling strategy is not restricted to ISTA, nor is it restricted to problems with a sparse prior: any algorithm for which the iterative step may be carried out by a learnable neural network layer may be unrolled. Gregor and LeCun developed a learned version of the coordinate descent algorithm, finding that the learned

version again obtained much lower prediction error than the iterative version [11]. Other authors have applied the unrolling framework to a variety of algorithms for biological data processing tasks, including ADMM [19] and robust principal component analysis [20], which were shown to obtain lower errors in magnetic resonance imaging signal recovery and ultrasound clutter suppression, respectively, in less time than then-state-of-the-art algorithms, consistent with learned unrolled networks converging more quickly. Outside of the realm of biology, unrolling of the HQS algorithm has been shown to achieve both high-quality denoising [21] and superresolution in natural images [17]. Many additional example applications are provided in a recent review [12].

Unrolling in optical localization microscopy

In the following sections, we focus on the domain of optical localization microscopy. First, we give a detailed example of

---

From the Iterative Shrinkage-Thresholding Algorithm to the Learned Iterative Shrinkage-Thresholding Algorithm

Here we detail the iterative shrinkage-thresholding algorithm (ISTA) and use it as a case study to describe the process of algorithm unrolling. Given a problem of the form (2), ISTA estimates $\mathbf{x}$, taking as inputs the measurement matrix $\mathbf{A}$, measurement vector $\mathbf{y}$, regularization parameter λ, and L, a Lipschitz constant of $\nabla \|\mathbf{Ax} - \mathbf{y}\|_2^2$.

Algorithm 1. ISTA

Require: $\mathbf{y}, \mathbf{A}, \lambda, L,$ number of iterations $k_{\max}$
Ensure: $\hat{\mathbf{x}}$
1: $\hat{\mathbf{x}}_1 = 0, \quad k = 1.$
2: **while** $k < k_{\max}$ **do**
3: $\quad \hat{\mathbf{x}}_{k+1} = \mathcal{T}_{\frac{\lambda}{L}}(\hat{\mathbf{x}}_k - 2L\mathbf{A}^\mathbf{T}(\mathbf{A}\hat{\mathbf{x}}_k - \mathbf{y}))$
4: $\quad k \leftarrow k + 1$
5: **end while**
6: $\hat{\mathbf{x}} = \hat{\mathbf{x}}_{k_{\max}}$

Here $\mathcal{T}_{\lambda/L}$ represents the soft thresholding operator after which ISTA is named:

$$\mathcal{T}_\alpha(x) = \max\{|x| - \alpha, 0\} \cdot \text{sgn}(x), \quad \text{(S1)}$$

where $\text{sgn}(\cdot)$ is the sign operator:

$$\text{sgn}(x) = \begin{cases} -1, & x < 0 \\ 1, & x > 0. \end{cases} \quad \text{(S2)}$$

The iterative step of ISTA is given in line 3 of Algorithm 1. The argument of $\mathcal{T}_{\lambda/L}(\cdot)$ in the iterative step can be rewritten as the sum of matrix-vector products with $\mathbf{y}$ and $\mathbf{x}_k$:

$$\mathbf{x}_k - 2L\mathbf{A}^T(\mathbf{Ax}_k - \mathbf{y})) = 2L\mathbf{A}^T\mathbf{y} + (\mathbf{I} - 2L\mathbf{A}^T\mathbf{A})\mathbf{x}_k$$
$$= \mathbf{W}_{0k}\mathbf{y} + \mathbf{W}_k\mathbf{x}_k. \quad \text{(S3)}$$

This step can be modeled by the sum of fully connected neural network layers and an activation function with learned threshold, as depicted in Figure S1. Depending on the structure of $\mathbf{A}$, convolutional layers may be chosen to preserve locality and reduce computational load. By stringing several of these layers together, the resulting deep neural network, i.e., the learned iterative shrinkage-thresholding algorithm, has the same form as the operation performed by running ISTA over multiple iterations. Different from ISTA, the weights of each layer are trained independently, providing greater flexibility, and the number of layers is chosen as a fixed, small number.
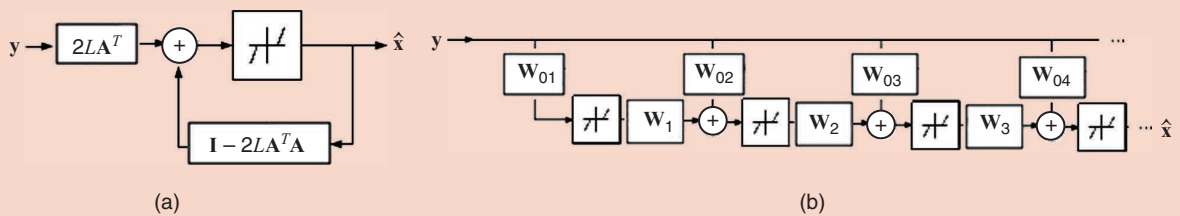

(a)   (b)

FIGURE S1. The unrolling of ISTA into the learned iterative shrinkage-thresholding algorithm (LISTA). (a): A diagram of the operation of ISTA as a feedback loop. (b) A diagram of LISTA; the matrix multiplications by $2L\mathbf{A}^T$ are replaced by weight matrices, $\mathbf{W}_{0k}$, and multiplication by $\mathbf{I} - 2L\mathbf{A}^T\mathbf{A}$ is replaced by $\mathbf{W}_k$, which, along with $\mathbf{W}_{0k}$, may be optimized with supervised learning.

how unrolling enhances the capabilities of one optical imaging technique: SMLM. Then we discuss how the concept of unrolling can be applied to other sparse biological optical imaging problems.

Unrolling in single-molecule localization microscopy

The visualization of subcellular features and organelles within biological cells requires imaging techniques with nanometer resolution. In the case of optical imaging systems, from the 19th century until the recent development of superresolution microscopy, the resolution limit was considered to be set by Abbe's diffraction limit for a microscope:

$$d = \frac{\beta}{2NA}, \tag{3}$$

where d is the minimal distance, below which two point sources cannot be distinguished; β is the wavelength of the emitted photons; and NA is the numerical aperture of the microscope. In fluorescence microscopy, the sample is stained with fluorophores, which can be excited with one color of light, and emit photons of a higher wavelength for subsequent detection. As most cells are not naturally fluorescent, this allows for specific imaging of the stained biomolecules. If the number of photons emitted is sufficiently high and the background is sufficiently low, single molecules can be detected in this way. However, biological structures of interest are typically made of multitudes of the same biomolecule type in close apposition, obscuring details finer than the diffraction limit of the emitted photons when all fluorophores are emitting at the same time.

One may overcome the diffraction limit by distinguishing between the photons coming from two neighboring fluorophores [22]. One way to distinguish neighboring molecules is by utilizing photo-activated or photo-switching fluorophores to separate fluorescent emission in time; this is the basis for SMLM techniques such as photo-activated localization microscopy (PALM) and stochastic optical reconstruction microscopy (STORM) [23], [24]. Optical, physical, or chemical means are used to ensure that at any given moment only a small subset of all fluorophores are emitting photons. Then a large number of diffraction-limited images is collected, each containing just a few active, isolated fluorophores. The imaging sequence is long enough such that each fluorophore is stochastically activated from a nonemissive state to a bright state, and back to a nonemissive (or bleached) state. During each cycle, the density of activated molecules is kept low enough so that emission profiles of individual fluorophores do not overlap.

High-resolution fluorophore localization can be framed as a linear inverse problem. Let us denote the collected sequence of diffraction-limited frames as $\mathbf{Y} \in \mathbb{R}^{M^2 \times T}$, where every column is the M^2 vector stacking of the corresponding $M \times M$ frame. Our goal is to reconstruct an image of size $N \times N$, consisting of fluorophore locations on a fine grid ($N > M$). We can model the generation of $\mathbf{Y}$ as

$$\mathbf{Y} = \mathbf{AX}, \tag{4}$$

where $\mathbf{X} \in \mathbb{R}^{N^2 \times T}$ is the sequence of vector-stacked high-resolution frames, and the nonzero entries in each frame (i.e., columns in the matrix) correspond to the locations of activated fluorophores. Matrix $\mathbf{A} \in \mathbb{R}^{M^2 \times N^2}$ is the measurement matrix, where each column of $\mathbf{A}$ is defined as *the system's PSF shifted by a single pixel on the high-resolution grid.*

The simplest way to retrieve $\mathbf{X}$ without leveraging knowledge of its biological structure is by fitting the observed emission profile, $\mathbf{Y}$, to the PSF of the system, which is typically modeled as a Gaussian function in 2D. This results in localizations with precision greater than the diffraction limit (accurate up to a few to tens of nanometers, versus a diffraction limit of 200 nm), allowing for imaging at a molecular scale within cells. Figure 2 illustrates the enhanced resolution of SMLM: STORM reveals the underlying structure of a circular DNA construct, which was completely unseen in its diffraction-limited images.

Although they achieve excellent resolution, standard SMLM methods have one main drawback: they require lengthy imaging times to achieve full coverage of the imaged specimen on the one hand, and minimal overlap between PSFs on the other.
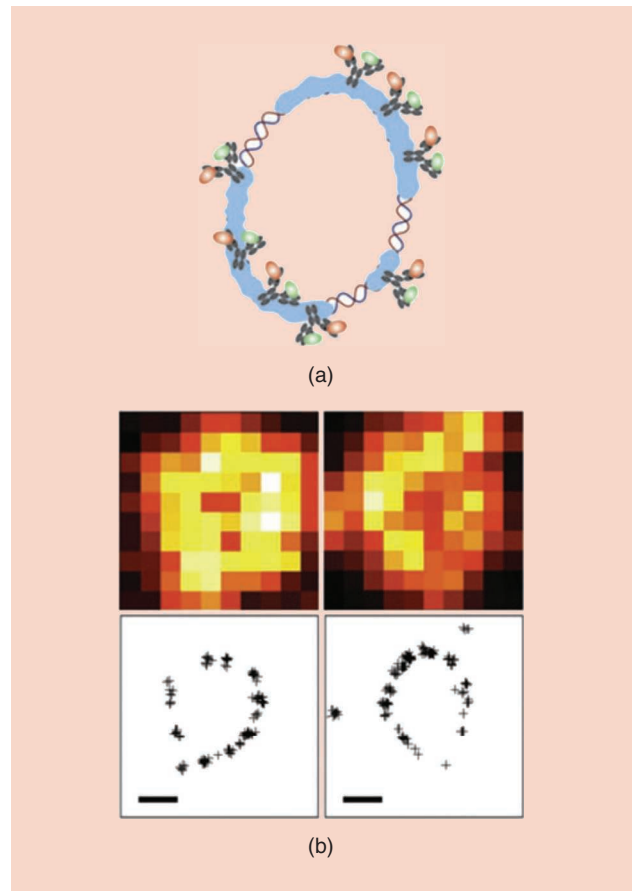

FIGURE 2. The sample experimental results from [24], comparing diffraction-limited and STORM-generated images of RecA-coated circular plasmid DNA. (a) An illustration of the DNA construct, with linked fluorophores (via immunohistochemistry). (b) The diffraction-limited frames taken by a total internal reflection microscope (top), and the reconstructed STORM images of the same frames (bottom). Scale bars: 300 nm.

Thus, in its classical form, this technique has low temporal resolution, preventing its application to fast-changing specimens in live-cell imaging. To circumvent the long acquisition periods required for SMLM methods, a variety of techniques have emerged, which enable the use of a smaller number of frames for reconstructing the 2D superresolved image [18], [25], [27]–[29]. These techniques take advantage of prior information regarding either the optical setup, geometry of the sample, or statistics of the emitters. One such technique is sparsity-based super-resolution correlation microscopy (SPARCOM) [25], [30], which exploits sparsity in the correlation domain while assuming that the blinking emitters are uncorrelated over time and space. This allows reformulation of the localization task as a sparse recovery problem, which can be solved using ISTA (see "Learned Sparsity-Based Super-Resolution Correlation Microscopy" and Figure S2).

SPARCOM yields excellent results when compared to a standard STORM reconstruction (using ThunderSTORM [26]), as illustrated in Figure 3. SPARCOM achieves similar spatial resolution with as few as 361 and even 60 frames, compared with the 12,000 frames needed for ThunderSTORM to produce a reliable recovery, corresponding to a 33- or 200-times faster acquisition rate when using SPARCOM. Thus, SPARCOM improves temporal resolution while retaining the spatial resolution of PALM/STORM. Gaining these benefits comes with tradeoffs: SPARCOM requires prior knowledge of the PSF of the optical setup for the calculation of the measurement matrix, which is not always available, and a careful choice of regularization factor λ, which is generally done heuristically.

As shown in the previous section, these shortcomings can be overcome by learning from data using an algorithm unrolling approach. This was done recently by Dardikman-Yoffe and Eldar [6], who introduced learned SPARCOM (LSPARCOM), a deep network with 10 layers resulting from unrolling SPARCOM, as detailed in "Learned Sparsity-Based Super-Resolution Correlation Microscopy."

The results shown in Figure 4 illustrate that inference from 10 folds of LSPARCOM is comparable to running SPARCOM for 100 iterations with a carefully chosen regularization parameter. Both methods succeed in reconstructing the underlying tubulin structure from a sequence of 350 high-density frames. Moreover, if a shorter, denser sequence is constructed by summing groups of 14 frames of the original sequence, the SPARCOM reconstruction's resolution degrades while the LSPARCOM reconstruction remains excellent. Thus, even with 25 extremely dense frames as input, LSPARCOM yields excellent reconstruction of subwavelength features, which allows for substantially higher temporal resolution compared to the hundreds of frames needed for SPARCOM. LSPARCOM is also faster to use, with an approximate five-times improvement over SPARCOM in execution time [6]. LSPARCOM enables efficient and accurate imaging well below the diffraction limit, without prior knowledge regarding the PSF or imaging parameters.

Given its enhanced capabilities, LSPARCOM has great potential for the localization of biological structures. Meeting the temporal and spatial resolutions' criteria for imaging of dynamic cellular processes at a molecular scale, it might replace its iterative counterpart as a robust, efficient method for live-cell imaging. The success of LSPARCOM further suggests that unrolling may benefit other sparse biological imaging problems, as we discuss in the following sections.

Optical microscopy extensions
In the previous section, learned unrolling was shown to achieve fast, highly accurate results in SMLM. Next we touch on two other applications that may benefit from unrolling: imaging transcriptomics (IT) and synapse detection.

Imaging transcriptomics
IT is a family of fluorescence microscopy techniques that studies the spatial distribution of mRNA transcripts in cells. This can enable the classification of individual cells by their gene expression in the context of their location in a tissue, yielding insight about the function of the whole system [5] or revealing subcellular spatial organization of mRNA transcripts. Many IT methods are based on single-molecule fluorescence in situ hybridization (smFISH) [2] in which fluorophore-labeled
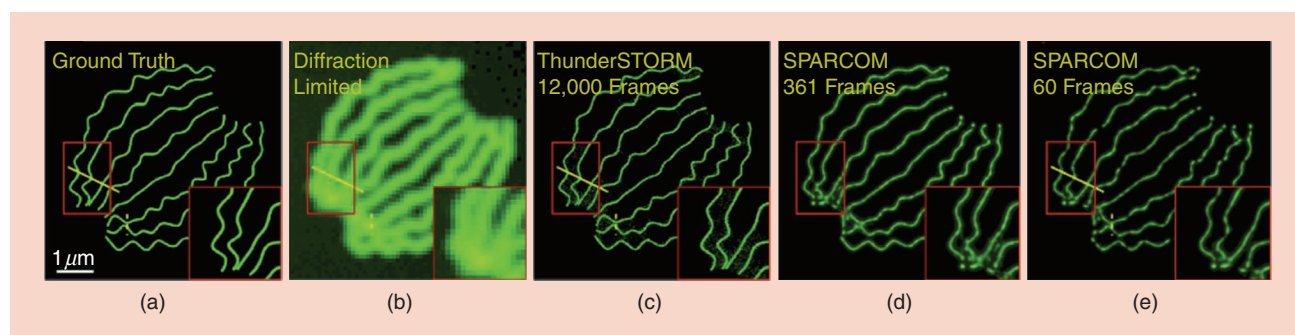

FIGURE 3. The results from [25], showing the simulation and reconstruction of microtubules from a motion picture [7] of 361 high-density frames. (a) The simulated ground truth of the image with subwavelength features. (b) A diffraction-limited image obtained by summing all the 361 frames in the movie. (c) A single-molecule localization reconstruction from a low-density movie of 12,000 frames of the same simulated microtubules and the same number of emitters (the image is constructed using the ThunderSTORM plug-in [26] for ImageJ). SPARCOM recoveries for movies with 361 and 60 high-density frames (the simulated microtubules and number of emitters is the same) are given in (d) and (e), respectively.

In SPARCOM, we start by observing the temporal covariance matrices of $\mathbf{X}$ and $\mathbf{Y}$, $\mathbf{M}_X$ and $\mathbf{M}_Y$. According to (4), we can write the following:

$$\mathbf{M}_Y = \mathbf{A}\mathbf{M}_X\mathbf{A}^T. \tag{S4}$$

We assume that different emitters are uncorrelated over time and space. Thus, $\mathbf{M}_X$ is a diagonal matrix, where each entry on its diagonal, $\mathbf{m}$, represents the variance of the emitter fluctuation on a high-resolution grid. As a nonzero variance can only exist where there is fluctuation in emission, the support of the diagonal corresponds to the emitters' locations on the high-resolution grid. Therefore, recovering $\mathbf{m}$ and reshaping it as a matrix yields the desired high-resolution image. For this purpose, let us rewrite (S4) as

$$\mathbf{M}_Y = \sum_{i=1}^{N^2} \mathbf{A}_i \mathbf{A}_i^T \mathbf{m}_i, \tag{S5}$$

where $\mathbf{A}_i$ is the ith column in $\mathbf{A}$, and $\mathbf{m}_i$ is the ith entry in $\mathbf{m}$. Following (2), we can exploit the sparsity of emitters and compute $\mathbf{m}$ by solving the following sparse recovery problem:

$$\min_{\mathbf{m}\geq 0} \lambda \|\mathbf{m}\|_1 + \frac{1}{2}\left\|\mathbf{M}_Y - \sum_{i=1}^{N^2} \mathbf{A}_i \mathbf{A}_i^T \mathbf{m}_i\right\|_2^2, \tag{S6}$$

where $\lambda \geq 0$ is the regularization parameter. The iterative shrinkage-thresholding algorithm can be used to solve this optimization problem, as shown in Figure S2.

To apply unrolling to SPARCOM, we need to replace the operations performed in a single iteration with neural network layers and choose the input of the unrolled algorithm. The unrolling process is illustrated in Figure S2: to start, $\mathbf{G}$, the $N \times N$, matrix-shaped resized version of the diagonal of $\mathbf{M}_Y$, is taken as input. The matrix-multiplication operations performed in each iteration are replaced with convolutional filters $W_p^{(k)}$, $k = 0, \ldots, 9$, and the positive soft-thresholding operator is replaced with a differentiable, sigmoid-based approximation of the positive hard-thresholding operator [31], denoted as $S_{\alpha_0,\beta_0}^+(\cdot)$. The unrolling process results in LSPARCOM, a deep neural network that acts as the operation performed by running SPARCOM over multiple iterations. LSPARCOM can be trained on a single sequence of frames taken from one field of view with a known, underlying structure, which can be generated using simulations (like the one offered by ThunderSTORM [26]). The model is then trained on overlapping small patches taken from multiple frames of that sequence.
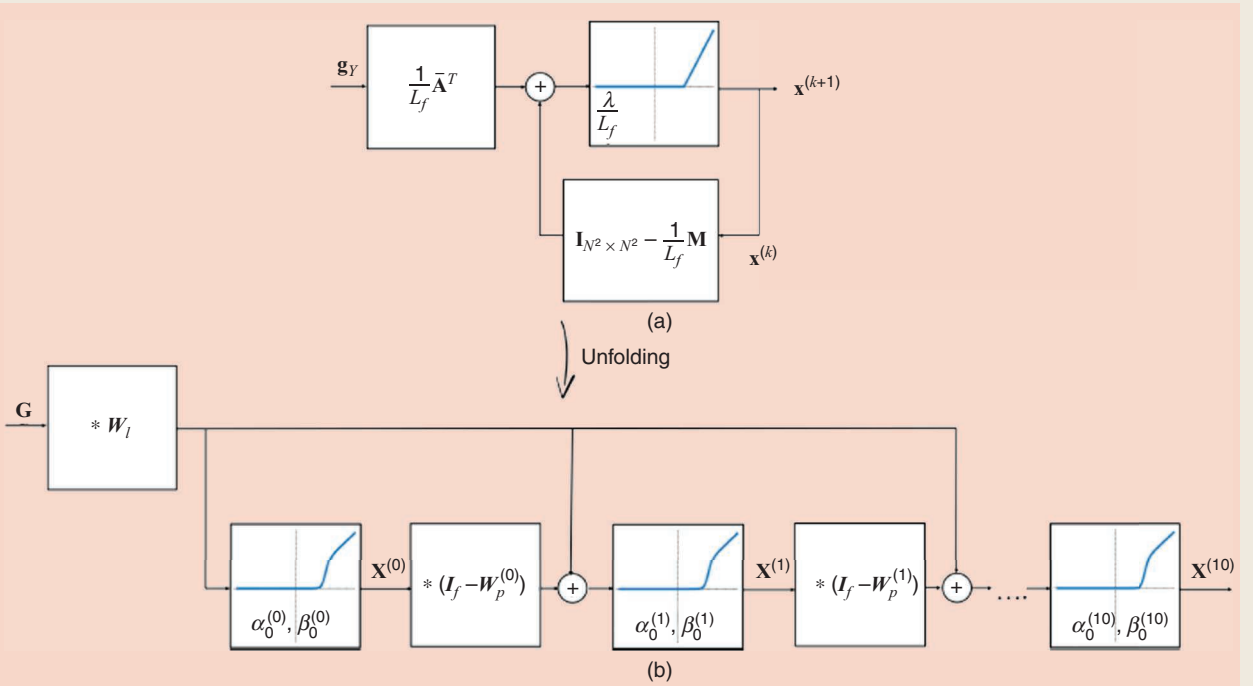

FIGURE S2. The unrolling of SPARCOM to LSPARCOM from [6]. (a) A block diagram of SPARCOM (via the iterative shrinkage-thresholding algorithm), recovering the vector-stacked superresolved image $\mathbf{x}^{(k)}$ (which corresponds to $\mathbf{m}$). The input is $\mathbf{g}_Y$, the diagonal of $\mathbf{M}_Y$. The block with the blue graph is $\mathcal{T}_{\lambda/L_f}$, the positive, soft-thresholding operator with parameter λ/L_f, where L_f is the Lipschitz constant of the gradient of (S6). The other blocks denote matrix multiplication (from the left side), where $\bar{\mathbf{A}} = \mathbf{A}^2$ (element-wise power) and $\mathbf{M} = |\mathbf{A}^T\mathbf{A}|^2$ (absolute value and power operations performed element wise). (b) LSPARCOM, recovering the superresolved image $\mathbf{X}^{(k)}$. Input $\mathbf{G}$ is the matrix-shaped resized version of $\mathbf{g}_Y$. The blocks with the blue graph apply smooth-activation function $S_{\alpha_0,\beta_0}^+(\cdot)$ with two trainable parameters: $0 \geq \alpha_0^{(k)} \geq 1$, $\beta_0^{(k)}$, $k = 0, \ldots, 10$. The other blocks denote convolutional layers, where I_f is a nontrainable identity filter and W_i, $W_p^{(k)}$, and $k = 0, \ldots, 9$ are trainable filters.

probes bind to complementary regions of mRNA. Although smFISH localizes transcripts of one gene at a time, in many experiments, it is desirable to study multiple genes at once, up to tens of thousands. To achieve this goal, combinatorial IT techniques, like multiplexed error-robust fluorescence in situ hybridization (MERFISH) [5], assign a distinct binary barcode with length F to each transcript. The barcodes are chosen to be distinct entries in a "codebook": F rounds of FISH imaging are performed, with transcripts appearing as spots in round f if the fth bit of its barcode is 1, as depicted in Figure 5. By using this technique, up to 2^F genes may be studied in only F rounds of imaging.

Once these F rounds of imaging are performed, images are processed to produce a set of localizations for mRNA of each gene. The problem of translating images into such a list is a sparse recovery problem: the fluorophores, scattered sparsely across the sample, appear in the images, modulated by the codebook and blurred by the PSF of the microscope. The goal, similar to SMLM, is to locate these sparsely scattered fluorescent emitters. Currently used processing techniques analyze image data with a heuristic approach in which each location is separately checked for a signal. In [8], we formalized this system analogously to the sparse optimization problem (2), in a method called the *joint sparse method for IT*. For IT data with

F rounds of imaging, studying G genes, with N_h locations on the high-resolution location grid and N_l pixels in measurement images, we can set up an optimization problem similar to (2). We vectorize and concatenate images into matrix $\mathbf{Y} \in \mathbb{R}^{N_l \times F}$. Then we take $\mathbf{Y}$ to be generated as a product of three matrices:

$$\mathbf{Y} = \mathbf{AXC}. \tag{5}$$

Here, $\mathbf{A} \in \mathbb{R}^{N_l \times N_h}$ is the same as in (1), the columns of $\mathbf{X} \in \mathbb{R}^{N_h \times G}$ are signal vectors like $\mathbf{x}$ in (1), with each column representing a specific gene, and $\mathbf{C} \in \mathbb{R}^{G \times F}$ is the set of barcodes.

Given (5), we recover $\mathbf{X}$ from measurement $\mathbf{Y}$ and known matrices $\mathbf{A}$ and $\mathbf{C}$ using an optimization-based approach, constrained by assumptions of sparsity: that only one mRNA will be present at each location, and that relatively few mRNA will be present in the FoV. This constrained optimization problem can be solved with an iterative algorithm. In addition to being more interpretable than the currently used heuristic, this method has achieved more accurate mRNA localization, especially in low-magnification imaging. A natural extension of this formulation is the application of learned unrolling. Based on our experience with other applications, unrolling may improve performance, obtaining more accurate genetic expression levels in fewer iterations while eliminating
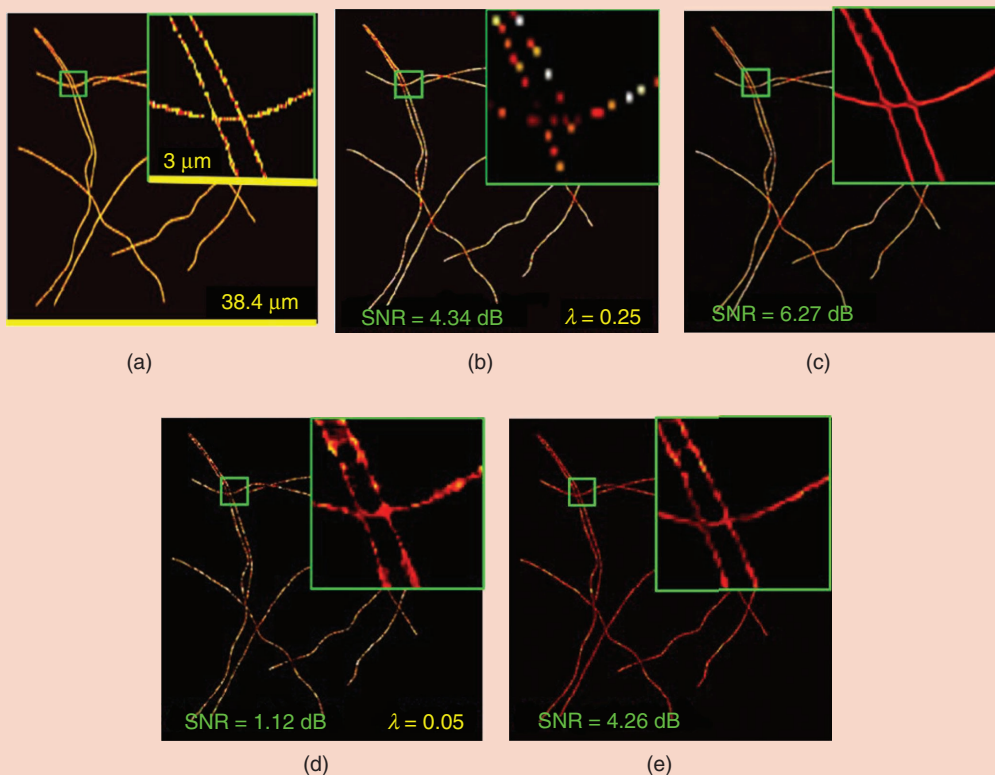

FIGURE 4. The sample results from [6], reconstructed from a simulated biological tubulins data set [7], composed of (b) and (c) 350 high-density or (d) and (e) 25 very high-density frames. (a) The simulated ground-truth tubulin structure. (b) and (d) The SPARCOM reconstruction, executed over 100 iterations with (b) $\lambda = 0.25$ for 350 frames and (d) $\lambda = 0.05$ for 25 frames. (c) and (e) The LSPARCOM reconstruction, given (c) 350 frames and (e) 25 frames as input. SNR: signal-to-noise ratio.

parameter-tuning requirements and explicit knowledge of the optical PSF **A**.

Synapse detection

Many other biological settings involve sparse emitters but have not necessarily been framed as sparse recovery problems to improve performance. For example, sparse biological images are encountered in neuronal synapse detection for characterization of the neurophysiological consequences of genetic and pharmacological perturbation screens. Synapses are localized by identifying positions in which fluorescently tagged pre- and postsynaptic proteins are located in close proximity. These proteins cluster into puncta, which appear as point sources, similar to fluorophores in SMLM data. A sparse analysis could identify subpixel locations for each punctum and, from the presence of both pre- and postsynaptic proteins, infer accurate synapse locations. Although sparse algorithm unrolling has not yet been applied in this context, model-based learning strategies have already shown good results for this setting [3].

Unrolling in other imaging modalities

Sparse emitters arise in other biological imaging modalities beyond epifluorescent microscopy, and the algorithmic unrolling method has achieved fast, highly accurate localization in several such settings. In the following sections, we review three such cases: ULM, LFM, and cell-center localization in nonspatially sparse histology images.

Unrolling in ultrasound localization microscopy

The attainable resolution of ultrasonography is fundamentally limited by wave diffraction, i.e., the minimum distance between separable scatters is one half of a wavelength. Due to this limit, conventional ultrasound techniques are bound to a tradeoff between resolution and penetration depth: increases in the transmit frequency shortens the wavelength (thus increasing resolution) but come at the cost of reduced penetration depth as higher frequency waves suffer from stronger absorption. This tradeoff particularly hinders deep, high-resolution microvascular imaging, which is crucial for many diagnostic applications.

A decade ago, this tradeoff was circumvented by the introduction of ULM [9], [32], which leverages the principles of SMLM and adapts these to ultrasound imaging. In SMLM, stochastic "blinking" of subsets of fluorophores is exploited to provide sparse point sources; in ULM, lipid-shelled gas microbubbles fulfill this role. A sequence of diffraction-limited ultrasonic scans is acquired, each containing just a few active isolated sources. Thus, each received image frame can be written as

$$\mathbf{y} = \mathbf{A}\mathbf{x} + \mathbf{w}, \tag{6}$$

where **x** is a vector that describes the sparse microbubble distribution on a high-resolution image grid, **y** is a vectorized image frame from the ultrasound sequence, **A** is the measurement matrix defined by the system's PSF, and **w** is a noise vector. As in SMLM, this enables precise localization of their centers on a subdiffraction grid. The accumulation of many such localizations over time yields a superresolved image. This approach achieves a resolution up to 10-times smaller than the wavelength [33], showing that ultrasonography at subdiffraction scale is possible.

Similar to SMLM, the quality of ULM imaging is dependent on the quantity of localized microbubbles and localization accuracy; thus, it gives rise to a new tradeoff between microbubble density and acquisition time. To achieve the desired signal sparsity for straightforward isolation of the backscattered echoes, ULM is typically performed using a very diluted solution of microbubbles. On regular ultrasound systems, this constraint leads to long acquisition times to cover the full vascular bed. Ultrafast plane-wave ultrasound (uULM) imaging has managed to lower the acquisition time [33] by taking many snapshots of individual microbubbles as they transport through the vasculature, thereby facilitating high-fidelity reconstruction of the larger vessels. Nevertheless, mapping the full capillary bed still requires microbubbles to pass through
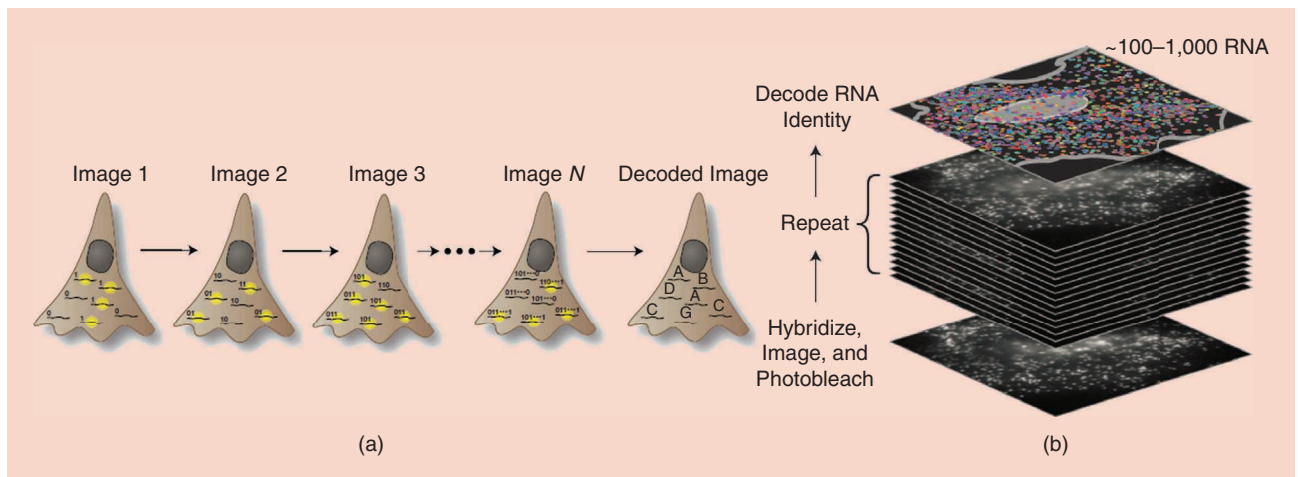

FIGURE 5. A depiction of MERFISH multiplexed IT from [5]. (a) Fluorophores are hybridized to mRNA transcripts if the bit of the associated barcode is equal to one, and if mRNA appears as a spot. (b) The acquisition and decoding of MERFISH data.

each capillary, capping acquisition time benefits to tens of minutes [34].

As with SMLM, uULM can be extended by using the sparsity of the measured signal (whether spatially sparse or in any transform domain [36]). Sparse recovery again enables improved localization precision and recall for high-microbubble concentrations [37]. Figure 6 depicts this, showing how the sparse recovery method produces fine visualization of the human prostate from a high-density sequence of in vivo ultrasound scans. However, as in the case of SPARCOM, solving the ULM sparse recovery problem requires iterative algorithms, such as ISTA. Unfortunately, as previously noted, these algorithms are not computationally efficient, and their effectiveness is strongly dependent on a good approximation of the system's PSF and careful tuning of the optimization parameters. With the unrolling approach, these challenges can be met in a fashion similar to the aforementioned SMLM (see "Deep Unrolled Ultrasound Localization Microscopy").

The tests on synthetic data show that deep unrolled ULM significantly outperforms standard ULM and sparse recovery through FISTA for high-microbubble concentrations (see Figure 7), offering better recall (measured by the recovered density) and lower localization error. In addition, when tested on in vivo ultrasound data, van Sloun et al. [38] observed that deep unrolled ULM yields superresolution images with higher fidelity, implying improved robustness and better generalization capabilities. Furthermore, Bar-Shira et al. [39] demonstrated how the use of deep unrolled ULM for in vivo human superresolution imaging allows for better diagnosis of breast pathologies. The unrolled method is also highly efficient, requiring slightly more than 1,000 flops and containing only
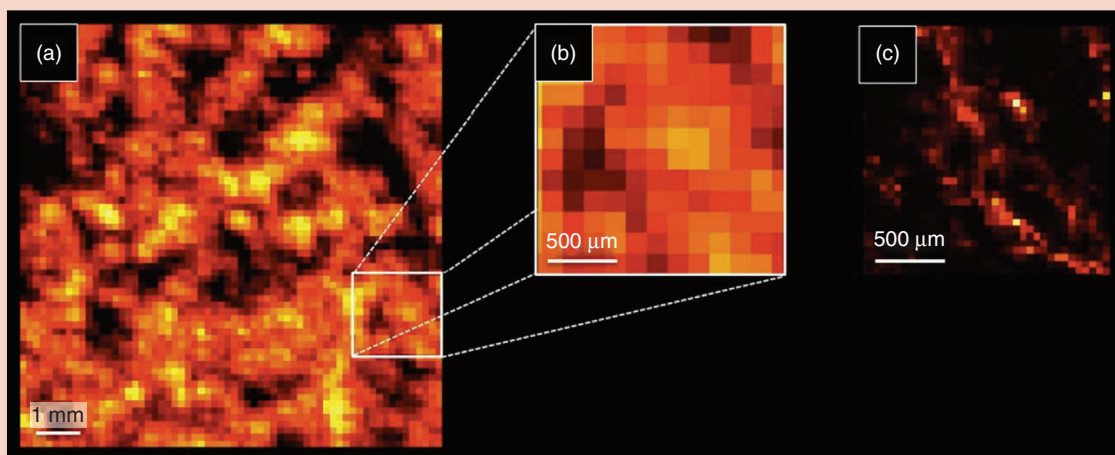

FIGURE 6. The sample results from [35]. Reconstructed, highly dense sequence of 300 frames, clinically acquired in vivo from a human prostate. (a) A maximum intensity projection image of the sequence, (b) a selected area in the image, and (c) the sparsity-driven, superresolution ultrasound on the same area.

Deep Unrolled Ultrasound Localization Microscopy

Under the assumption of spatial sparsity of microbubbles in the high-resolution grid, $\mathbf{x}$ [as defined in (6)] corresponds to the solution of the l_1-regularized inverse problem, which was previously presented in (2); thus, it can be computed using the iterative shrinkage-thresholding algorithm (ISTA). After estimating $\mathbf{x}$ for each frame, the estimates are summed across all frames to yield the final superresolution image, which describes the microbubble distribution throughout the entire sequence.

To apply unrolling in this case, the learned iterative shrinkage-thresholding algorithm can replace ISTA, as in the "Sparse Recovery in Biological Imaging" section. van Sloun et al. [38] have implemented such a model, resulting in a

10-layer, feedforward neural network. Each layer consists of trainable, 5×5 convolutional filters $\mathbf{W}_{0k}$ and $\mathbf{W}_k$, along with a trainable shrinkage parameter λ^k $(k = 0,...,9)$. The convolutional filters replace the fully connected layers, which appear in the original unrolled version of ISTA (see Figure S1). Replacing the proximal soft-thresholding operator T_λ [see (S1)] with a smooth sigmoid-based soft-thresholding operation [31], helped avoid vanishing gradients. Similar to LSPARCOM, this network is trained on simulated ultrasound scans of point sources, with a variety of point-spread function and noise realizations, Overlapping small patches are taken from multiple frames of each simulated scan sequence and given to the network as training samples.

506 parameters (compared to millions of parameters in other deep learning models).

In summary, deep unrolled ULM can be a method for efficient, robust, and parameter-free ultrasonic imaging, with comparable (or superior) resolution to that of other ULM methods (standard and sparsity based). Given the ability of deep unrolled ULM to perform precise reconstructions for high-microbubble concentrations, deep-tissue ultrasound imaging becomes feasible. High-microbubble concentrations dramatically shorten required acquisition times, which allows deep unrolled ULM to perform deep, high-resolution imaging much faster. Thus, intricate ultrasonography tasks like microvascular imaging, which have a key role in the noninvasive, in vivo diagnosis of many medical conditions such as cancer, arteriosclerosis, stroke, and diabetes, become simpler to execute.

Light-field microscopy

Another imaging domain dealing with spatially sparse data is LFM [40]. Obtaining 3D information from a single acquisition is valuable, enabling real-time volumetric neural imaging. LFM provides single-shot 3D imaging by placing a microlens array between the microscope objective and the camera sensor. This configuration captures both lateral and angular information from each light ray emitted from the sample, so deconvolution of the system's PSF produces 3D emitter locations.

As each spatial location is imaged in multiple pixels on the detector, LFM faces a tradeoff between depth and lateral resolution. If the sample is composed of spatially sparse emitters, localization on a high-resolution, 3D grid can be performed, as in SMLM and ULM.

The authors in [10] present the problem of localizing neurons in 3D space with LFM images: in LFM, neurons are small enough to be considered point sources and are distributed in a spatially sparse manner. To localize neurons, measured images are converted to a structure called an *epipolar plane image* (*EPI*); the system PSF in this domain varies strongly with depth, as shown in Figure 8. By performing sparse optimization, the authors are able to achieve fast, accurate neuron localization.

By framing 3D neuron localization as a sparse optimization problem, the problem is opened to unrolling. In [10], Song et al. first use a convolutional variant of ISTA (CISTA) to solve the localization problem, then create an unrolled network based on that algorithm, called *CISTA-net*. The unrolled network recovers neuron location with higher accuracy in all dimensions and performs the recovery task more than 10,000-times faster than ISTA. This increase in speed expands the applicability of LFM: it could enable, for instance, live, 3D imaging of whole nervous systems in small-model organisms like *C. elegans*, or of activity of large volumes of the mammalian cortex.
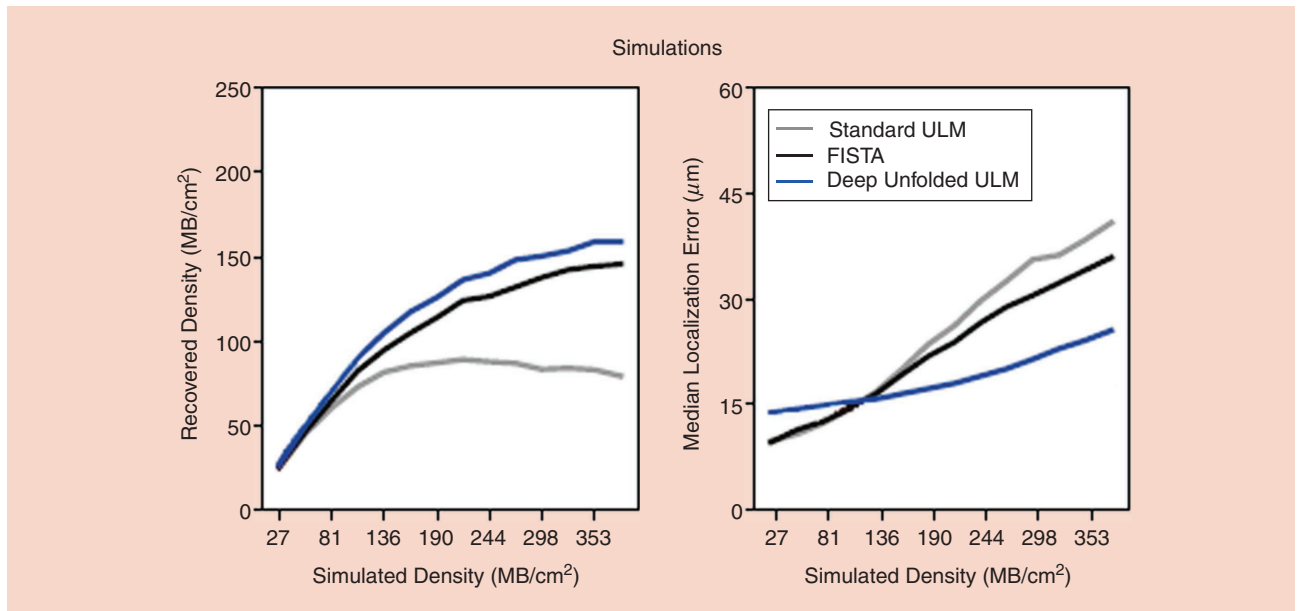

FIGURE 7. A performance comparison of standard ULM, sparse-recovery (FISTA), and deep unrolled ULM on simulations (taken from [38]).
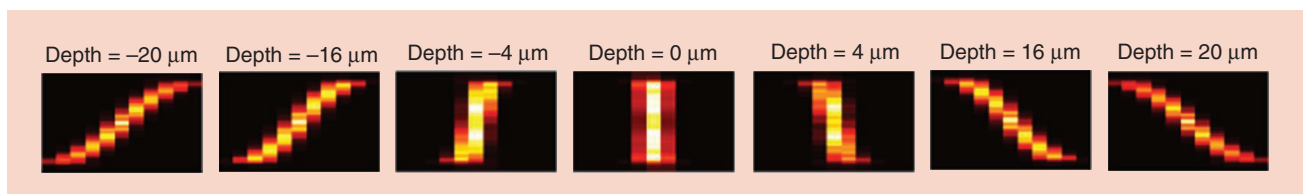

FIGURE 8. The epipolar plane images (EPIs) derived from LFM images of emitters at different depths (from [41] under Creative Commons License 4.0). By matching an observed EPI, the depth of sources may be determined.

Nonspatially sparse imaging

The most obvious way of thinking about sparse recovery in biological imaging is in the domain of spatially sparse sources, but other methods leverage sparse coding in other aspects, and unrolling can achieve accurate results in these situations as well. One example is cell-center localization in histology slides. Although cell centers are scattered sparsely in an FoV, cell shapes are irregular, so there is not a single "impulse response" transforming cell-center locations into images of cells, as in the sparse recovery form of (2). In [4], a traditional CNN is combined with a LISTA-like network to localize cell centers. In this framework, the locations of the centers of cells in a 2D FoV with dimension $h \times w$ are represented by binary matrix $\mathbf{X} \in \mathbb{R}^{h \times w}$. The matrix $\mathbf{X}$ is radon transformed to represent the cell centers in polar coordinates, $\mathbf{X}_p = \mathcal{R}f(X)$. A measurement matrix $\mathbf{A}$ is generated as a random Gaussian projection matrix, and the product $\mathbf{A}\mathbf{X_p} = \mathbf{Y}$ is formed. Xue et al. found that although $\mathbf{Y}$ cannot be measured directly, a CNN may be trained to infer $\mathbf{Y}$ from an image of the FoV. A two-stage neural network is designed: in the first stage, a traditional CNN transforms the images into an estimate $\hat{\mathbf{Y}}$ of the matrix $\mathbf{Y}$; in the second stage, a LISTA-like network is used to obtain an estimate $\hat{\mathbf{X}}_p$ of the sparse matrix $\mathbf{X_p}$, which, after an inverse Radon transformation, gives the cell-center locations $\mathbf{X}$. The network, called *end-to-end CNN and compressed sensing* (*ECNNCS*), is trained by penalizing the differences between both $\hat{\mathbf{y}}$ and $\mathbf{y}$, and $\hat{\mathbf{x}}$ and $\mathbf{x}$. The ECNNCS model achieved better localization accuracy than the state-of-the-art algorithms used as a comparison [4], showing that unrolling can improve performance outside problems of strict spatial sparsity.

Compressed in situ imaging (CISI) [42] is another imaging technique that leverages sparse recovery despite not using spatially sparse data and, accordingly, may benefit from algorithm unrolling. Like IT, CISI evaluates expression levels of genes at single-cell resolution. Different from IT, in CISI, data are not spatially sparse. Instead, CISI takes advantage of genetic coexpression patterns to infer single cells' transcriptomes from a few "composite measurements," which measure multiple genes at once. In CISI, single-cell transcriptomes are conceptualized as linear combinations of "modules," sparse linear combinations of coexpressed genes. The sparse recovery problem is to infer from composite measurements of genes the sparse set of active coexpression modules. Currently, modules are defined before the experiment, but with algorithmic unrolling, optimal coexpression modules could be learned, enabling improved transcriptome inference.

Conclusions

New biological imaging techniques are constantly being developed, and with them, computational pipelines to identify and characterize the imaged biological structures. We described a few of these techniques and their accompanying pipelines. In many cases, these techniques consist of heuristic strategies, which have limited accuracy and are difficult to interpret. As computational power continues to increase and the methods become more developed and powerful, interpretable processing techniques have been created by incorporating biological and physical assumptions into constrained optimization problems, solved with iterative methods. These methods, in turn, require parameter tuning and explicit knowledge of the experimental setup. A natural next step in pipeline development is model-based learning methods, including algorithmic unrolling. We have shown how, in many imaging modalities requiring source localization, unrolling achieves fast, accurate results with robust models, and proposed that unrolling be extended widely to other similar problems, including to methods involving biological structure other than sparsity. We hope this work will inspire methods extending unrolling to further biological imaging modalities and experimental settings leading to rapid, interpretable, high-performing physics-based methods.

Authors

Yair Ben Sahel (yair.bensahel@weizmann.ac.il) received his B.Sc. degree in physics from The Hebrew University of Jerusalem, Israel, in 2019. He is currently working toward his M.Sc. degree in mathematics and computer science at the Weizmann Institute of Science, Rehovot, 7610001, Israel. His research interests include machine learning and computational biology, focusing on learning-based algorithms for superresolution in biological imaging.

John P. Bryan (bryan@broadinstitute.org) received his B.S. and M.S. degrees in electrical engineering from Stanford University in 2019. He is currently working toward his Ph.D. degree, also in electrical engineering, at the Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, 02142, USA. He works in the Broad Institute of MIT and Harvard University, Cambridge, Massachusetts, 02142, USA, where he investigates applications of compressed sensing theory to analysis of microscopy data.

Brian Cleary (bcleary@broadinstitute.org) received his Ph.D. degree in computational and systems biology from the Massachusetts Institute of Technology (MIT). He is a computational systems biologist and Broad Fellow at the Broad Institute of MIT and Harvard University, Cambridge, Massachusetts, 02142, USA. His lab researches new ways to perform biological experiments using an algorithmic lens to jointly leverage biological structure and mathematical principles.

Samouil L. Farhi (sfarhi@broadinstitute.org) received his Ph.D. degree in chemical biology from Harvard University, where he developed methods for single photon all-optical neurophysiology. He now directs the Optical Profiling Platform at the Broad Institute of MIT and Harvard University, Cambridge, Massachusetts, 02142, USA. His lab's goal is to bring next generation imaging approaches to bear on large-scale biological problems, with a focus on profiling the spatial organization of tissues, high-content optical screens, and all-optical studies of the electrical properties of cells.

Yonina C. Eldar (yonina.eldar@weizmann.ac.il) is a professor in the Department of Math and Computer Science at the Weizmann Institute of Science, Rehovot, 7610001, Israel, where she heads the Center for Biomedical Engineering and Signal Processing. She is also a visiting professor at the Massachusetts Institute of Technology (MIT), Cambridge, Massachusetts, and at the Broad Institute of MIT and Harvard University, Cambridge, Massachusetts, 02142, USA, and is an adjunct professor at Duke University, Durham, North Carolina. She is a member of the Israel Academy of Sciences and Humanities and heads the Committee for Promoting Gender Fairness in Higher Education Institutions in Israel. She is the editor-in-chief of *Foundations and Trends in Signal Processing* and serves IEEE on several technical and award committees. She has received many awards for excellence in research and teaching, including the IEEE Signal Processing Society Technical Achievement Award, IEEE/AESS Fred Nathanson Memorial Radar Award, and IEEE Kiyo Tomiyasu Award. She is a Fellow of IEEE and of EURASIP.

References

[1] X. Zhuang, "Nano-imaging with STORM," *Nature Photon.*, vol. 3, no. 7, pp. 365–367, 2009, doi: 10.1038/nphoton.2009.101.

[2] A. M. Femino, F. S. Fay, K. Fogarty, and R. H. Singer, "Visualization of single RNA transcripts *in situ*," *Science*, vol. 280, no. 5363, pp. 585–590, 1998, doi: 10.1126/science.280.5363.585.

[3] V. Kulikov, S.-M. Guo, M. Stone, A. Goodman, A. Carpenter, M. Bathe, and V. Lempitsky, "DoGNet: A deep architecture for synapse detection in multiplexed fluorescence images," *PLoS Comput. Biol.*, vol. 15, no. 5, p. e1007012, 2019, doi: 10.1371/journal.pcbi.1007012.

[4] Y. Xue, G. Bigras, J. Hugh, and N. Ray, "Training convolutional neural networks and compressed sensing end-to-end for microscopy cell detection," *IEEE Trans. Med. Imag.*, vol. 38, no. 11, pp. 2632–2641, 2019, doi: 10.1109/TMI.2019.2907093.

[5] K. H. Chen, A. N. Boettiger, J. R. Moffitt, S. Wang, and X. Zhuang, "Spatially resolved, highly multiplexed RNA profiling in single cells," *Science*, vol. 348, no. 6233, p. 412, 2015, doi: 10.1126/science.aaa6090.

[6] G. Dardikman-Yoffe and Y. C. Eldar, "Learned SPARCOM: Unfolded deep super-resolution microscopy," *Opt. Exp.*, vol. 28, no. 19, pp. 27,736–27,763, Sep. 2020, doi: 10.1364/OE.401925.

[7] D. Sage, H. Kirshner, T. Pengo, N. Stuurman, J. Min, S. Manley, and M. Unser, "Quantitative evaluation of software packages for single-molecule localization microscopy," *Nature Methods*, vol. 12, no. 8, pp. 717–724, 2015, doi: 10.1038/nmeth.3442.

[8] J. Bryan, B. Cleary, S. Farhi, and Y.C. Eldar, "Sparse recovery of imaging transcriptomics data," in *Proc. Int. Symp. Biomed. Imag.*, 2021, pp. 802–806, doi: 10.1109/ISBI48211.2021.9433927.

[9] I. Grundberg *et al.*, "Super-resolution ultrasound imaging," *Ultrasound Med. Biol.*, vol. 46, no. 4, pp. 865–891, 2020, doi: 10.1016/j.ultrasmedbio.2019.11.013.

[10] P. Song, H. V. Jadan, C. L. Howe, P. Quicke, A. J. Foust, and P. L. Dragotti, "Model-inspired deep learning for light-field microscopy with application to neuron localization," 2021, arXiv:2103.06164.

[11] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proc. 27th Int. Conf. Mach. Learn.*, 2010, pp. 399–406.

[12] V. Monga, Y. Li, and Y. C. Eldar, "Algorithm unrolling: Interpretable, efficient deep learning for signal and image processing," *IEEE Signal Process. Mag.*, vol. 38, no. 2, pp. 18–44, 2021, doi: 10.1109/MSP.2020.3016905.

[13] R. Tibshirani, "Regression shrinkage and selection via the lasso," *J. Roy. Statist. Soc. Ser. B (Methodol.)*, vol. 58, no. 1, pp. 267–288, 1996, doi: 10.1111/j.2517-6161.1996.tb02080.x.

[14] Y. C. Eldar and G. Kutyniok, *Compressed Sensing: Theory and Applications*. Cambridge, U.K.: Cambridge Univ. Press, May 2012.

[15] S. Boyd, N. Parikh, E. Chu, B. Peleato, and J. Eckstein, "Distributed optimization and statistical learning via the alternating direction method of multipliers," *Found. Trends Mach. Learn.*, vol. 3, no. 1, pp. 1–122, 2011, doi: 10.1561/2200000016.

[16] I. Daubechies, M. Defrise, and C. De Mol, "An iterative thresholding algorithm for linear inverse problems with sparsity constraint," *Commun. Pure Appl. Math.*, vol. 57, no. 11, pp. 1413–1457, Nov. 2004, doi: 10.1002/cpa.20042.

[17] K. Zhang, L. V. Gool, and R. Timofte, "Deep unfolding network for image super-resolution," in *Proc. IEEE/CVF Conf. Comput. Vision Pattern Recognit.*, 2020, pp. 3217–3226, doi: 10.1109/CVPR42600.2020.00328.

[18] E. Nehme, L. E. Weiss, T. Michaeli, and Y. Shechtman, "Deep-STORM: Super-resolution single-molecule microscopy by deep learning," *Optica*, vol. 5, no. 4, pp. 458–464, 2018, doi: 10.1364/OPTICA.5.000458.

[19] Y. Yang, J. Sun, H. Li, and Z. Xu, "Deep ADMM-Net for compressive sensing MRI," in *Proc. 30th Int. Conf. Neural Inf. Process. Syst.*, NIPS'16, Red Hook, NY, USA: Curran Associates, Inc., 2016, pp. 10–18.

[20] O. Solomon, R. Cohen, Y. Zhang, Y. Yang, Q. He, J. Luo, R. J. van Sloun, and Y. C. Eldar, "Deep unfolded robust PCA with application to clutter suppression in ultrasound," *IEEE Trans. Med. Imag.*, vol. 39, no. 4, pp. 1051–1063, 2019, doi: 10.1109/TMI.2019.2941271.

[21] Y. Li, M. Tofighi, J. Geng, V. Monga, and Y. C. Eldar, "Efficient and interpretable deep blind image deblurring via algorithm unrolling," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 666–681, Jan. 2020, doi: 10.1109/TCI.2020.2964202.

[22] E. Betzig, "Proposed method for molecular optical imaging," *Opt. Lett.*, vol. 20, no. 3, pp. 237–239, 1995, doi: 10.1364/OL.20.000237.

[23] E. Betzig *et al.*, "Imaging intracellular fluorescent proteins at nanometer resolution," *Science*, vol. 313, no. 5793, pp. 1642–1645, Sep. 2006, doi: 10.1126/science.1127344.

[24] M. J. Rust, M. Bates, and X. Zhuang, "Sub-diffraction-limit imaging by stochastic optical reconstruction microscopy (STORM)," *Nature Methods*, vol. 3, no. 10, pp. 793–796, Aug. 2006, doi: 10.1038/nmeth929.

[25] O. Solomon, M. Mutzafi, M. Segev, and Y. C. Eldar, "Sparsity-based super-resolution microscopy from correlation information," *Opt. Exp.*, vol. 26, no. 14, pp. 18,238–18,269, Jun. 2018, doi: 10.1364/OE.26.018238.

[26] M. Ovesný, P. Křížek, J. Borkovec, Z. Svindrych, and G. M. Hagen, "ThunderSTORM: A comprehensive ImageJ plug-in for PALM and STORM data analysis and super-resolution imaging," *Bioinformatics*, vol. 30, no. 16, pp. 2389–2390, 2014, doi: 10.1093/bioinformatics/btu202.

[27] J. Min *et al.*, "FALCON: Fast and unbiased reconstruction of high-density super-resolution microscopy data," *Scientific Rep.*, vol. 4, no. 1, pp. 1–9, 2014, doi: 10.1038/srep04577.

[28] L. Zhu, W. Zhang, D. Elnatan, and B. Huang, "Faster STORM using compressed sensing," *Nature Methods*, vol. 9, no. 7, pp. 721–726, 2012, doi: 10.1038/nmeth.1978.

[29] T. Dertinger, R. Colyer, G. Iyer, S. Weiss, and J. Enderlein, "Fast, background-free, 3D super-resolution optical fluctuation imaging (SOFI)," *Proc. Nat. Acad. Sci. USA*, vol. 106, no. 52, pp. 22,287–22,292, 2009, doi: 10.1073/pnas.0907866106.

[30] O. Solomon, Y. C. Eldar, M. Mutzafi, and M. Segev, "SPARCOM: Sparsity based super-resolution correlation microscopy," *SIAM J. Imag. Sci.*, vol. 12, no. 1, pp. 392–419, Feb. 2019, doi: 10.1137/18M1174921.

[31] X.-P. Zhang, "Thresholding neural network for adaptive noise reduction," *IEEE Trans. Neural Netw.*, vol. 12, no. 3, pp. 567–584, 2001, doi: 10.1109/72.925559.

[32] Y. Desailly, O. Couture, M. Fink, and M. Tanter, "Sono-activated ultrasound localization microscopy," *Appl. Phys. Lett.*, vol. 103, no. 17, pp. 174,107, 2013, doi: 10.1063/1.4826597.

[33] C. Errico, J. Pierre, S. Pezet, Y. Desailly, Z. Lenkei, O. Couture, and M. Tanter, "Ultrafast ultrasound localization microscopy for deep super-resolution vascular imaging," *Nature*, vol. 527, no. 7579, pp. 499–502, 2015, doi: 10.1038/nature16066.

[34] V. Hingot, C. Errico, B. Heiles, L. Rahal, M. Tanter, and O. Couture, "Microvascular flow dictates the compromise between spatial resolution and acquisition time in ultrasound localization microscopy," *Scientific Rep.*, vol. 9, no. 1, p. 2456, 2019, doi: 10.1038/s41598-018-38349-x.

[35] R. J. G. van Sloun, O. Solomon, Y. C. Eldar, H. Wijkstra, and M. Mischi, "Sparsity-driven super-resolution in clinical contrast-enhanced ultrasound," in *Proc. IEEE Int. Ultrason. Symp.*, 2017, pp. 1–4, doi: 10.1109/ULTSYM.2017.8092945.

[36] A. Bar-Zion, O. Solomon, C. Tremblay-Darveau, D. Adam, and Y. C. Eldar, "SUSHI: Sparsity-based ultrasound super-resolution hemodynamic imaging," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 65, no. 12, pp. 2365–2380, 2018, doi: 10.1109/TUFFC.2018.2873380.

[37] O. Solomon, R. van Sloun, H. Wijkstra, M. Mischi, and Y. C. Eldar, "Exploiting flow dynamics for superresolution in contrast-enhanced ultrasound," *IEEE Trans. Ultrason., Ferroelectr., Freq. Control*, vol. 66, no. 10, pp. 1573–1586, 2019, doi: 10.1109/TUFFC.2019.2926062.

[38] R. van Sloun, R. Cohen, and Y. C. Eldar, "Deep learning in ultrasound imaging," *Proc. IEEE*, vol. 108, no. 1, pp. 11–29, 2020, doi: 10.1109/JPROC.2019.2932116.

[39] O. Bar-Shira, A. Grubstein, Y. Rapson, D. Suhami, E. Atar, K. Peri-Hanania, R. Rosen, and Y. C. Eldar, "Learned super resolution ultrasound for improved breast lesion characterization," in *Proc. Int. Conf. Med. Image Comput. Comput. Assisted Intervention*, 2021, pp. 1–10, doi: 10.1007/978-3-030-87234-2_11.

[40] M. Levoy, R. Ng, A. Adams, M. Footer, and M. Horowitz, "Light field microscopy," in *Proc. 2006 ACM SIGGRAPH Papers*, pp. 924–934, 2006, doi: 10.1145/1179352.1141976.

[41] P. Song, H. V. Jadan, C. L. Howe, P. Quicke, A. J. Foust, and P. L. Dragotti, "3D localization for light-field microscopy via convolutional sparse coding on epipolar images," *IEEE Trans. Comput. Imag.*, vol. 6, pp. 1017–1032, May 2020, doi: 10.1109/TCI.2020.2997301.

[42] B. Cleary *et al.*, "Compressed sensing for imaging transcriptomics," *bioRxiv*, Oct. 2020, doi: 10.1101/743039.

SP