# DEEP CONVOLUTIONAL ROBUST PCA WITH APPLICATION TO ULTRASOUND IMAGING

*Regev Cohen*[⋆] *Yi Zhang*[†] *Oren Solomon*[⋆] *Daniel Toberman*[⋆] *Liran Taieb*[⋆] *Ruud JG van Sloun*[‡] *Yonina C. Eldar*[⋆]

[⋆]Department of Electrical Engineering, Technion, Haifa 32000, Israel
[†]Department of Electronics Engineering, Tsinghua university, China
[‡]Department of Electrical Engineering, Eindhoven University of Technology, The Netherlands

## ABSTRACT

Sparse and low-rank decomposition, also known as robust principle component analysis, has been applied successfully in numerous applications. Typically, this approach leads to a minimization problem which is solved using iterative algorithms. Drawing inspiration from recurrent networks, in recent years deep-learning strategies have been extended to mimic the behavior of iterative algorithms, with reduced complexity. In this work, we propose an extension of these deep architectures to robust principle component analysis in which fully-connected layers are replaced with convolutional ones. This strategy offers spatial invariance and significant reduction in the number of learned parameters. We then apply the proposed method to contrast-enhanced ultrasound, in which low-rank tissue signal needs to be removed in order to visualize blood vessels. We demonstrate the effectiveness of our approach on simulations and *in-vivo* rat brain scans. The resulting images exhibit improved visual quality and contrast compared with images obtained by commonly practiced methods.

***Index Terms***— Ultrasound, Deep learning, Contrast agents, Unfolding, Low-rank decomposition.

## 1. INTRODUCTION

Robust principle component analysis (RPCA) [1, 2] is a powerful statistical tool for data analysis which is widely used in a variety of applications such as video surveillance [3], face recognition [4], and latent semantic indexing [5]. This modeling approach is based on decomposing the sample matrix into a sum of a low-rank component and a sparse component. In stochastic optimization, the recovery problem is formulated as a convex nuclear norm minimization which is typically solved by iterative algorithms such as principal component pursuit (PCP) [3], outlier pursuit [6], iterative reweighted least squares [7, 8], Bayesian RPCA [8] and iterative shrinkage and thresholding algorithm (ISTA) [9]. However, such techniques exhibit high complexity and slow run-time.

The rising interest in neural networks in recent years has led to the development of various deep-learning methodologies which mimic the behavior of the aforementioned algorithms [10] to reduce overall execution time and achieve similar performance with fixed numerical complexity. Unfolding, or unrolling an iterative algorithm was suggested by Gregor and LeCun [11] to accelerate algorithm convergence. Iterative algorithms provide a natural recurrent architecture, designed to solve a specific problem, such as sparse approximations [12], channel estimation [13] and more. The authors of [11] showed that by considering each iteration of an iterative algorithm as a layer in a deep network and subsequent concatenation of a few such layers, it is possible to train networks to achieve a dramatic improvement in convergence over iterative algorithms.

This project has received funding from the European Unions Horizon 2020 research and innovation program under grant No. 646804-ERC-COG-BNYQ, and from the Israel Science Foundation under grant No. 0100101.

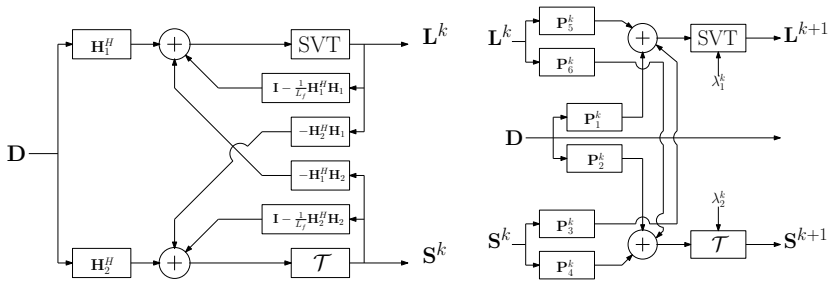
In the context of RPCA, a principled way to construct learnable pursuit architectures for structured sparse and robust low rank models was introduced in [10]. The proposed networks, derived from the iteration of proximal descent algorithms, were shown to faithfully approximate the solution of RPCA while demonstrating several orders of magnitude speed-up compared to standard optimization algorithms. Following the same paradigm, various solutions were introduced in recent years for different tasks such as clustering and classification [14], dynamic anomalography [15], online optimization [16] and more. However, this approach is based on a non-convex formulation of the nuclear norm in which the rank (or an upper bound of it) is assumed to be known a priori. This poses a network design limitation, as the rank can vary between different applications or even different realizations of the same application.

In this paper, we propose to combine RPCA with the power of complex convolutional neural networks [12, 17]. We introduce a network architecture based on unfolding a variation of ISTA for low-rank and sparse modeling where the thresholding parameters are learned and the rank is not known in advance. In addition, we replace the fully-connected layers, typically used, with convolutional kernels which offer spatial invariance and lead to significant reduction in the number of learned parameters. We demonstrate the effectiveness of our approach on contrast-enhanced ultrasound (CEUS) scans [18], for suppressing tissue clutter signal from the blood signal.

CEUS allows the detection and visualization of blood vessels whose different physical parameters are associated with different clinical conditions [19]. Gas microbubbles are used as ultrasound contrast agents (UCAs) which are administrated intravenously and flow throughout the vascular system due to their similarity in size to red blood cells [20]. A major challenge in CEUS is to suppress clutter signals stemming from stationary and slowly moving tissue as they introduce significant artifacts in blood flow imaging [21]. Most of the current solutions in the field of ultrasound imaging are based on singular value decomposition (SVD) of the spatiotemporal correlation matrix of successive temporal samples [22, 23]. However, similar to [10], this SVD-filtering uses a predefined rank which typically is unknown, and can vary between scans of the same organ of different patients.

To overcome this, we adopt RPCA for CEUS where we model the received signal as the composition of low-rank tissue signal and sparse UCA signal. We implement and apply fast ISTA (FISTA) [24] on *in-vivo* data to separate the UCA from the tissue signal. Then, we use the results of the latter along with additional simulated UCA/tissue acquisitions to train an unfolded network with fixed-length and complex convolutional layers. The performance of the network is tested on *in-vivo* rat brain scans and the results are compared with those achieved by the commonly practiced SVD-filtering and FISTA. Using only 10 layers, we achieve better contrast in vascular visualization than SVD-filtering and FISTA with 30000 iterations.

The rest of the paper is organized as follows.In Section 2 we introduce the problem formulation. Section 3 describes a solution via an iterative algorithm. In Section 4 we propose a convolution-




(a) Iterative algorithm for sparse and low-rank separation.

(b) $k$th layer from the corresponding unfolded algorithm.

**Fig. 1**: Architecture comparison between the iterative algorithm (panel (a)) and its unfolded counterpart (panel (b)). The learned network in panel (b) draws its architecture from the iterative algorithm, and is trained on examples from a given dataset. In both panels, $\mathbf{D}$ is the input measurements matrix, $\mathbf{S}_k$ and $\mathbf{L}_k$ are the estimated sparse and low-rank matrices in each iteration/layer, respectively.

based unfolded network for RPCA and its application to ultrasound imaging. Section 5 presents simulation and *in-vivo* results of both the iterative and unfolded algorithms. We conclude in Section 6.

Throughput the paper, $x$ represents a scalar, $\mathbf{x}$ a vector, $\mathbf{X}$ a matrix and $\mathbf{I}_{N\times N}$ is the $N \times N$ identity matrix. The notation $||\cdot||_p$ represents the standard $p$-norm and $||\cdot||_F$ is the Frobenius norm. Subscript $x_l$ denotes the $l$th element of $\mathbf{x}$ and $\mathbf{x}_l$ is the $l$th column of $\mathbf{X}$. Superscript $\mathbf{x}^{(p)}$ represents $\mathbf{x}$ at iteration $p$, $T^*$ denotes the adjoint of T, and $\bar{\mathbf{A}}$ is the complex conjugate of $\mathbf{A}$.

## 2. PROBLEM FORMULATION

We begin by describing the general acquisition model. Assume a measured data matrix $\mathbf{D}$ of dimensions $m \times n$, which is expressed as the following sum

$$\mathbf{D} = \mathbf{H}_1\mathbf{L} + \mathbf{H}_2\mathbf{S} + \mathbf{N}. \tag{1}$$

Here, $\mathbf{H}_1$ and $\mathbf{H}_2$ are known measurement matrices of dimensions $m \times k_1$ and $m \times k_2$, respectively, and $\mathbf{N}$ is a matrix of additive noise. The latter model asserts that the measurements can be decomposed into a low-rank part $\mathbf{L}$ and a sparse part $\mathbf{S}$. Our goal is to recover both $\mathbf{L}$ and $\mathbf{S}$ of (1).

Similar to [25], recovering $\mathbf{L}$ and $\mathbf{S}$ can be performed via minimization of

$$\min_{\mathbf{L},\mathbf{S}} \frac{1}{2}||\mathbf{D} - (\mathbf{H}_1\mathbf{L} + \mathbf{H}_2\mathbf{S})||_F^2 + \lambda_1||\mathbf{L}||_* + \lambda_2||\mathbf{S}||_{1,2}, \tag{2}$$

where $||\cdot||_*$ stands for the nuclear norm and $||\cdot||_{1,2}$ is the mixed $l_{1,2}$ norm. We use the mixed $l_{1,2}$ norm (though the standard $l_1$ norm can also be used), since in our application UCA positions correspond to the positions of blood vessels, which are assumed to be static, or change slightly during the acquisition period. Thus, they share a common support dictated by the blood vessels in which they flow. The nuclear norm is known to promote low-rank solutions, and is the convex relaxation of the non-convex rank minimization constraint [26]. By defining

$$\mathbf{X} = \begin{bmatrix} \mathbf{L} \\ \mathbf{S} \end{bmatrix},\ \mathbf{P}_1 = \begin{bmatrix} \mathbf{I} \\ \mathbf{0} \end{bmatrix},\ \mathbf{P}_2 = \begin{bmatrix} \mathbf{0} \\ \mathbf{I} \end{bmatrix}$$

and $\mathbf{A} = [\mathbf{H}_1, \mathbf{H}_2]$, (2) can be rephrased as

$$\min_{\mathbf{L},\mathbf{S}} \frac{1}{2}||\mathbf{D} - \mathbf{A}\mathbf{X}||_F^2 + h(\mathbf{X}), \tag{3}$$

where $h(\mathbf{X}) = \sum_{i=1}^{2} \lambda_i \rho_i(\mathbf{P}_i\mathbf{X})$ with $\rho_1 = ||\cdot||_*$ and $\rho_2 = ||\cdot||_{1,2}$.

Our goal is to minimize (3) using a fixed complexity deep network. To do so, we first describe an iterative solution to (3), which we then unfold as a deep network.

## 3. ITERATIVE SOLUTION VIA PROXIMAL DESCENT

The minimization problem (3) is a regularized least-squares problem, for which numerous numerical minimization algorithms exist. Specifically, (F)ISTA involves finding *Moreau's proximal* (prox) mapping [27] of $h$, defined as

$$\text{prox}_h(\mathbf{X}) = \underset{\mathbf{U}}{\text{argmin}} \left\{ h(\mathbf{U}) + \frac{1}{2}||\mathbf{U} - \mathbf{X}||_F^2 \right\}. \tag{4}$$

Plugging the definition of $\mathbf{X}$ into (4) yields

$$\text{prox}_h(\mathbf{X}) = \underset{\mathbf{U}_1,\mathbf{U}_2}{\text{argmin}} \Big\{ \lambda_1\rho_1(\mathbf{U}_1) + \frac{1}{2}||\mathbf{U}_1 - \mathbf{L}||_F^2 + \lambda_2\rho_2(\mathbf{U}_2) + \frac{1}{2}||\mathbf{U}_2 - \mathbf{S}||_F^2 \Big\},$$

for which it can be shown that since $\text{prox}_h(\mathbf{X})$ is separable in $\mathbf{L}$ and $\mathbf{S}$, it holds that

$$\text{prox}_h(\mathbf{X}) = \begin{bmatrix} \text{prox}_{\rho_1}(\mathbf{L}) \\ \text{prox}_{\rho_2}(\mathbf{S}) \end{bmatrix}. \tag{5}$$

The proximal mapping (5) is applied in each iteration to the gradient of the quadratic part of (3), given by

$$\frac{d}{d\mathbf{X}}\frac{1}{2}||\mathbf{D} - \mathbf{A}\mathbf{X}||_F^2 = \mathbf{A}^H(\mathbf{A}\mathbf{X} - \mathbf{D}),$$

and more specifically,

$$\begin{bmatrix} \frac{d}{d\mathbf{L}} \\ \frac{d}{d\mathbf{S}} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_1^H(\mathbf{H}_1\mathbf{L} + \mathbf{H}_2\mathbf{S} - \mathbf{D}) \\ \mathbf{H}_2^H(\mathbf{H}_1\mathbf{L} + \mathbf{H}_2\mathbf{S} - \mathbf{D}) \end{bmatrix}.$$

Similar to [28], the general iterative step of ISTA applied to minimizing (2) is given by

$$\begin{aligned} \mathbf{L}^{k+1} &= \text{SVT}_{\lambda_1/L_f}\left\{\left(\mathbf{I} - \tfrac{1}{L_f}\mathbf{H}_1^H\mathbf{H}_1\right)\mathbf{L}^k - \mathbf{H}_1^H\mathbf{H}_2\mathbf{S}^k + \mathbf{H}_1^H\mathbf{D}\right\} \\ \mathbf{S}^{k+1} &= \mathcal{T}_{\lambda_2/L_f}\left\{\left(\mathbf{I} - \tfrac{1}{L_f}\mathbf{H}_2^H\mathbf{H}_2\right)\mathbf{S}^k - \mathbf{H}_2^H\mathbf{H}_1\mathbf{L}^k + \mathbf{H}_2^H\mathbf{D}\right\} \end{aligned}, \tag{6}$$

where $L_f$ is the Lipschitz constant of the quadratic term of (3), given by the spectral norm of $\mathbf{A}^H\mathbf{A}$ and SVT$(\cdot)$ and $\mathcal{T}(\cdot)$ are the singular value thresholding [29] and mixed $l_{1/2}$ soft thresholding operators [30]. The diagram in Fig. 1(a) presents the iterative algorithm, which relies on knowledge of $\mathbf{H}_1, \mathbf{H}_2$ and selection of $\lambda_1$ and $\lambda_2$.

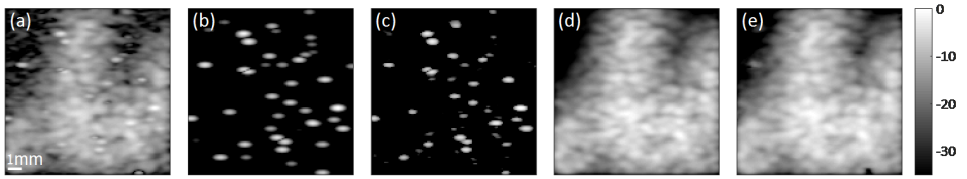



**Fig. 2**: Simulation results. (a) A single frame from an input movie of super-imposed UCA and tissue clutter. (b) Ground-truth UCA frame. (c) Recovered UCA frame via the unfolded network. (d) Ground-truth tissue signal. (d) Recovered tissue signal via the unfolded network. Color bar is in dB.

## 4. CONVOLUTION BASED DEEP LEARNING RPCA

### 4.1. Unfolding the iterative algorithm

As previously mentioned, an iterative algorithm can be considered as a recurrent neural network, in which the $k$th iteration can be regarded as the $k$th layer in a feedforward network [10]. On the other hand, convolutions are linear operators, thus considering (6), for each layer in the network we swap each of the matrices dependent on $\mathbf{H}_1$ and $\mathbf{H}_2$ with convolution layers (kernels) $\mathbf{P}_1^k, \ldots, \mathbf{P}_6^k$ of appropriate sizes. The kernels as well as the regularization parameters $\lambda_1^k$ and $\lambda_2^k$, are learned in training. By doing so, the following equations for the $k$th layer are obtained

$$\mathbf{L}^{k+1} = \text{SVT}_{\lambda_1^k}\left\{\mathbf{P}_5^k * \mathbf{L}^k + \mathbf{P}_3^k * \mathbf{S}^k + \mathbf{P}_1^k * \mathbf{D}\right\},$$
$$\mathbf{S}^{k+1} = \mathcal{T}_{\lambda_2^k}\left\{\mathbf{P}_6^k * \mathbf{L}^k + \mathbf{P}_4^k * \mathbf{S}^k + \mathbf{P}_2^k * \mathbf{D}\right\},$$

with $*$ being a convolution operator. The latter can be considered as a single layer in a multi-layer feedforward network. A diagram of a single layer from the unfolded architecture is given in Fig. 1(b), where the supposedly known model matrices were replaced by the 2D convolution kernels $\mathbf{P}_1^k, \ldots, \mathbf{P}_6^k$, which are learned as part of the training process of the overall network.

Contrary to previous works in unfolding RPCA which considered training fully connected (FC) layers [10], we employ convolution kernels in our implementation, which allows us to achieve spatial invariance while reducing the number of learned parameters considerably. In many applications, the recovered matrices $\mathbf{S}$ and $\mathbf{L}$ represent a 3D volume, or movie, of dynamic objects imposed on a (quasi) static background. Each column in $\mathbf{S}$ and $\mathbf{L}$ is a vectorized frame from the recovered sparse and low-rank movies. Thus, we consider in practice our data as a 3D volume and apply 2D convolutions. The SVT operation at the $k$th layer is performed after reshaping the input 3D volume into a 2D matrix, by vectorizing and column-wise stacking each frame.

The thresholding coefficients are learned independently for each layer. Given the $k$th layer, the actual thresholding values for both the SVT and soft-thresholding operations are given by $\text{thr}_L^k = \sigma(\lambda_L^k) \cdot a_L \cdot \max(L^k)$, and $\text{thr}_S^k = \sigma(\lambda_S^k) \cdot a_S \cdot \text{mean}(S^k)$, where $\sigma(\cdot)$ is a sigmoid function, $a_L$ and $a_S$ are fixed scalars (in our application we chose $a_L = 0.4$ and $a_S = 1.8$) and $\lambda_L^k$ and $\lambda_S^k$ are learned in each layer by the network.

### 4.2. Training the unfolded network

The unfolded network presented in the previous section is trained using back-propagation in a supervised manner. Generally speaking, we obtain training examples $\mathbf{D}_i$ and corresponding sparse $\hat{\mathbf{S}}_i$ and low-rank $\hat{\mathbf{L}}_i$ decompositions. In practice, $\hat{\mathbf{S}}_i$ and $\hat{\mathbf{L}}_i$ can either be obtained from simulations or by decomposing $\mathbf{D}_i$ using iterative algorithms such as FISTA. The loss function is chosen as the sum of the mean squared errors (MSE) between the predicted $\mathbf{S}$ and $\mathbf{L}$ values of the network and $\hat{\mathbf{S}}_i$, $\hat{\mathbf{L}}_i$, respectively,

$$\mathcal{L}(\boldsymbol{\theta}) = \frac{1}{2N}\sum_{i=1}^{N} ||f_S(\mathbf{D}_i, \boldsymbol{\theta}) - \hat{\mathbf{S}}_i||_F^2 + \frac{1}{2N}\sum_{i=1}^{N} ||f_L(\mathbf{D}_i, \boldsymbol{\theta}) - \hat{\mathbf{L}}_i||_F^2.$$

In the latter equation, $f_{S/L}$ is the sparse/low-rank output of the unfolded network with learnable parameters $\boldsymbol{\theta} = \{\mathbf{P}_1^k, \ldots, \mathbf{P}_6^k, \lambda_1^k, \lambda_2^k\}$, $k = 1, \ldots, K$, where $K$ is the number of chosen layers.

### 4.3. Application to ultrasound imaging

In this section we describe how the unfolded network can be applied to improve vascular visualization in contrast enhanced ultrasound (US) imaging. In US imaging, typically a series of pulses are transmitted to the imaged medium. The resulting echoes from the sonicated medium are received in each transducer element and then combined in a process called beamforming to produce a focused image. As presented in [31], after demodulation the complex analytical (IQ) signal can be represented as

$$D(x, z, t) = I(x, z, t) + jQ(x, z, t),$$

where $I(x, z, t), Q(x, z, t)$ are the in-phase and quadrature components of the demodulated signal, $x, z$ are the vertical and axial coordinates, and $t$ indicates frame number. The signal $D(x, z, t)$ is a sum of IQ demodulated echoes returned from the UCA ($S(x, z, t)$) as well as from the tissue ($L(x, z, t)$), contaminated by additive noise ($N(x, z, t)$):

$$D(x, z, t) = L(x, z, t) + S(x, z, t) + N(x, z, t).$$

Acquiring a series of $t = 1, \ldots, T$ frames and stacking them as vectors in a matrix leads to the following model

$$\mathbf{D} = \mathbf{L} + \mathbf{S} + \mathbf{N}. \tag{7}$$

Assuming that each movie frame is of size $M \times M$ pixels, each of the matrices of (7) are of size $M^2 \times T$. Model (7) corresponds to the low-rank and sparse model with $\mathbf{H}_1 = \mathbf{H}_2 = \mathbf{I}$.

In this application, UCA flow within the vasculature and allow their clear visualization. However, the tissue clutter signal first needs to be removed. Given a set of $T$ scans of flowing UCAs, we apply FISTA to the data and achieve approximations for the low-rank tissue and sparse UCA signals, with fixed thresholding parameters ($\lambda_1 = 0.02$, $\lambda_2 = 0.001$) and for 30000 iterations.

Our unfolded network consists of 10 layers. The first two layers use convolution kernels of size $5 \times 5 \times 1$, while the rest of the layers have $3 \times 3 \times 1$ kernels. As the demodulated input signal $\mathbf{D}$ is complex valued, complex convolutions were implemented [17].

Training a deep network with millions of parameters typically requires a vast amount of training examples, and in practice, US scans of specific organs are not available in abundance. Thus, convolutions instead of FC layers help in reducing the total number of trained parameters. Furthermore, to produce a large number of training examples, we rely on two strategies: patch-based analysis and simulations.

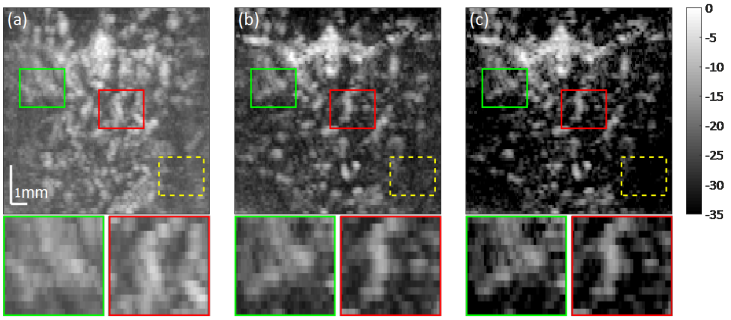



**Fig. 3**: Extraction of UCA from tissue signal. (a) SVD based separation. (b) L+S FISTA separation. (c) Deep network separation, with the unfolded architecture of the FISTA algorithm. Color bar is in dB.

First, we divide the US movie into 3D patches (axial coordinate, lateral coordinate and frame number) and apply FISTA on each of these 3D patches to obtain the desired decomposition. In practice each such patch is of size $32 \times 32$ pixels ($1.64 \times 1\text{mm}^2$), over 20 consecutive frames and between consecutive patches there exists an overlap of $50\%$, which enables us to increase the number of training data. Second, we generate additional data by simulating an acquired movie of randomly flowing UCAs, cluttered by tissue, in a realistic simulation. The simulation provides a perfect ground truth, which helps improve the reconstruction result of the network over its iterative counterpart.

Implementation was done in Python 3.5, using the PyTorch 0.4.0 package. Training was performed using the ADAM optimizer with a constant learning rate of 0.01. For the *in-vivo* results, we first train the network on a simulated movie of UCAs and tissue signal for 10 epochs. We then refine the network by training on an *in-vivo* scan of a rat for an additional 10 epochs. *In-vivo* results were obtained after applying the network to a scan of a different rat, on which the network did not train.

## 5. RESULTS

In this section, we present results of improved vascular visualization with a fixed-complexity deep unfolded architecture for both simulations and *in-vivo* scans.

Figure 2 shows the recovery results of the network on a single frame from simulated data (panel (a) shows an example input frame). The network was trained only on simulations for 10 epochs. Panels (c) and (e) show the UCA and tissue recoveries of the network, respectively. Visual inspection reveals these recoveries to be very similar to the ground truth images (panels (b) and (d), respectively).

Figure 3 presents *in-vivo* results. Panel (a) depicts the resulting SVD based extraction of the UCA signal, while panel (b) shows the FISTA based separation and panel (c) shows the result of the unfolded network. In all three cases, once the CEUS movie is extracted from the input movie, in order to present a single representative image, for each pixel, we take its maximum absolute value over the movie frames. This technique is commonly referred to as maximum intensity projection (MIP).

In each panel, the green and red boxes indicate selected areas, whose enlarged views are presented in the corresponding green and red boxes below each panel. Visual inspection of the entire panels (a)-(c) shows that both FISTA and unfolded network achieve CEUS signal separation with lower residual clutter/tissue background than the naive SVD approach. Moreover, as the unfolded network was also trained on simulated data, for which we have perfect tissue suppression, its performance in terms of contrast seems to surpass FISTA. Considering the enlarged regions below panels (a)-(c) further supports this conclusion, showing better contrast of the FISTA and unfolded network outputs.

To further quantify the performance of each method, we provide a metric to assess the contrast ratio of the output of each method. For each panel we calculate the contrast ratio (CR), defined as $\text{CR} = \mu_s/\mu_b$, and assess the performance of each method. CR is calculated between a selected patch, e.g. the red or green boxes in panels (a)-(c) and a reference patch, marked by the dashed yellow patches, which represents the background, for the same image. That is, for each panel we estimate the CR values of the red - yellow and green - yellow boxes, where $\mu_s$ is the mean of the red/green box and $\mu_b$ is the mean of the dashed yellow patch. Table 1 provides the estimated CR values of each method, respectively.

**Table 1**: CR values for the selected green and red rectangles of Fig. 3, as compared with the dashed yellow background rectangle in each corresponding panel. All values are in dB.

| | **SVD** | **FISTA** | **Unfolded** |
|---|---|---|---|
| Green box | 4.68 | 5.52 | 15.24 |
| Red box | 4.56 | 5.24 | 14.88 |

In CR, higher values imply higher contrast ratios, which suggest better noise suppression and better signal depiction. Considering both tables, it is evident that the FISTA and unfolded network approaches achieve better CR values.

## 6. CONCLUSION

In this work, we unfold a first order iterative algorithm for RPCA decomposition as a deep network, and replace the FC layers with learnable convolution kernels. In contrast to previous results, our network does not need to know the rank of the low-rank part a-priori and is suitable for complex valued data. We then applied our methodology to the problem of tissue clutter suppression in CEUS and showed that the acquired movie can be modeled as a composition of a low-rank part for the tissue signal and a sparse matrix for the UCA signal. We then train an unfolded network of 10 layers on this task and apply the network to simulated and *in-vivo* data, showing better visual quality (higher contrast) than the commonly used SVD filter and a FISTA based iterative algorithm with 30000 iterations.

## 7. ACKNOWLEDGMENT

We would like to thank Prof. Luo Jianwen and Yi Yang from the center for biomedical imaging research in Tsinghua university for providing the *in-vivo* data.

# References


[1] T. Bouwmans and E. H. Zahzah, "Robust PCA via principal component pursuit: A review for a comparative evaluation in video surveillance," *Computer Vision and Image Understanding*, vol. 122, pp. 22–34, 2014.

[2] Z. Lin, "A review on low-rank models in data analysis," *Big Data & Information Analytics*, vol. 1, no. 2/3, pp. 139–161, 2016.

[3] E. J. Candès, X. Li, Y. Ma, and J. Wright, "Robust principal component analysis?" *Journal of the ACM (JACM)*, vol. 58, no. 3, p. 11, 2011.

[4] A. Georghiades and P. Belhumeur, "Illumination cone models for faces recognition under variable lighting," in *Proceedings of CVPR98*, 1998.

[5] R. P. T. H. Papadimitriou, C.H. and S. Vempala, "Latent semantic indexing: A probabilistic analysis," *Journal of Computer and System Sciences*, vol. 61, no. 2, pp. 217–235, 2000.

[6] C. C. Xu, H. and S. Sanghavi, "Robust pca via outlier pursuit," in *Advances in Neural Information Processing Systems*, 2010, pp. 2496–2504.

[7] B. T. Guyon, C. and E. Zahzah, "Foreground detection via robust low rank matrix factorization including spatial constraint with iterative reweighted regression," in *Pattern Recognition (ICPR), 2012 21st International Conference on*. IEEE, 2012, pp. 2805–2808.

[8] ——, "Moving object detection via robust low rank matrix decomposition with irls scheme," in *International Symposium on Visual Computing*. Springer, 2012, pp. 665–674.

[9] A. Beck and M. Teboulle, "Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems," *IEEE Transactions on Image Processing*, vol. 18, no. 11, pp. 2419–2434, 2009.

[10] P. Sprechmann, A. M. Bronstein, and G. Sapiro, "Learning efficient sparse and low rank models," *IEEE transactions on pattern analysis and machine intelligence*, vol. 37, no. 9, pp. 1821–1833, 2015.

[11] K. Gregor and Y. LeCun, "Learning fast approximations of sparse coding," in *Proceedings of the 27th International Conference on International Conference on Machine Learning*. Omnipress, 2010, pp. 399–406.

[12] H. Sreter and R. Giryes, "Learned convolutional sparse coding," in *2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 2018, pp. 2191–2195.

[13] N. Samuel, T. Diskin, and A. Wiesel, "Deep mimo detection," *arXiv preprint arXiv:1706.01151*, 2017.

[14] Q. Qiu and G. Sapiro, "Learning transformations for clustering and classification," *The Journal of Machine Learning Research*, vol. 16, no. 1, pp. 187–225, 2015.

[15] M. G. Mardani, M. and G. Giannakis, "Dynamic anomalography: Tracking network anomalies via sparsity and low rank," *IEEE Journal of Selected Topics in Signal Processing*, vol. 7, no. 1, pp. 50–66, 2013.

[16] X. H. Feng, J. and S. Yan, "Online robust PCA via stochastic optimization," in *Advances in Neural Information Processing Systems*, 2013, pp. 404–412.

[17] C. Trabelsi, O. Bilaniuk, Y. Zhang, D. Serdyuk, S. Subramanian, J. F. Santos, S. Mehri, N. Rostamzadeh, Y. Bengio, and C. J. Pal, "Deep complex networks," *arXiv preprint arXiv:1705.09792*, 2017.

[18] B. Furlow, "Contrast-enhanced ultrasound," *Radiologic technology*, vol. 80, no. 6, pp. 547S–561S, 2009.

[19] T. Opacic, S. Dencks, B. Theek, M. Piepenbrock, D. Ackermann, A. Rix, T. Lammers, E. Stickeler, S. Delorme, G. Schmitz *et al.*, "Motion model ultrasound localization microscopy for preclinical and clinical multiparametric tumor characterization," *Nature communications*, vol. 9, no. 1, p. 1527, 2018.

[20] N. De Jong, F. Ten Cate, C. Lancee, J. Roelandt, and N. Bom, "Principles and recent developments in ultrasound contrast agents," *Ultrasonics*, vol. 29, no. 4, pp. 324–330, 1991.

[21] S. Bjaerum, H. Torp, and K. Kristoffersen, "Clutter filter design for ultrasound color flow imaging," *IEEE transactions on ultrasonics, ferroelectrics, and frequency control*, vol. 49, no. 2, pp. 204–216, 2002.

[22] F. W. Mauldin, D. Lin, and J. A. Hossack, "The singular value filter: a general filter design strategy for pca-based signal separation in medical ultrasound imaging," *IEEE transactions on medical imaging*, vol. 30, no. 11, pp. 1951–1964, 2011.

[23] 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.

[24] A. Beck and M. Teboulle, "A Fast Iterative Shrinkage-Thresholding Algorithm," *SIAM Journal on Imaging Sciences*, vol. 2, no. 1, pp. 183–202, 2009.

[25] R. Otazo, E. Candès, and D. K. Sodickson, "Low-rank plus sparse matrix decomposition for accelerated dynamic mri with separation of background and dynamic components," *Magnetic Resonance in Medicine*, vol. 73, no. 3, pp. 1125–1136, 2015.

[26] E. J. Candès and B. Recht, "Exact matrix completion via convex optimization," *Foundations of Computational mathematics*, vol. 9, no. 6, p. 717, 2009.

[27] J.-J. Moreau, "Proximité et dualité dans un espace hilbertien," *Bulletin de la Société mathématique de France*, vol. 93, pp. 273–299, 1965.

[28] C. Mu, Y. Zhang, J. Wright, and D. Goldfarb, "Scalable robust matrix recovery: Frank–wolfe meets proximal methods," *SIAM Journal on Scientific Computing*, vol. 38, no. 5, pp. A3291–A3317, 2016.

[29] J.-F. Cai, E. J. Candès, and Z. Shen, "A singular value thresholding algorithm for matrix completion," *SIAM Journal on Optimization*, vol. 20, no. 4, pp. 1956–1982, 2010.

[30] Y. C. Eldar and G. Kutyniok, *Compressed sensing: theory and applications*. Cambridge University Press, 2012.

[31] A. Bar-Zion, C. Tremblay-Darveau, O. Solomon, D. Adam, and Y. C. Eldar, "Fast vascular ultrasound imaging with enhanced spatial resolution and background rejection," *IEEE transactions on medical imaging*, vol. 36, no. 1, pp. 169–180, 2017.