Supplementary Material of "Tensor Ring Decomposition Guided Dictionary Learning for OCT Images Denoising"

Parisa Ghaderi Daneshmand[1], Hossein Rabbani[*,1], Senior Member, IEEE

[1]Medical Image & Signal Processing Research Center, School of Advanced Technologies in Medicine, Isfahan University of Medical Sciences, Isfahan, Iran

I. Overview

This supplementary material includes the following sections. In Section II, we investigate the computational complexity of our proposed TRGDL model. In Section III, we have provided an empirical examination to depict the convergence behavior of the suggested TRGDL denoising method. To show the effect of the number of the B-scans in our algorithm, we try to use our algorithm on 3, 4, and 5 B-scans, in Section IV. In Section V, we provided the quantitative and qualitative results for denoising OCT images of dataset-3. Section VI presents a comparison of the results obtained by the proposed TRGDL method with those obtained using the averaging method. Finally, we provide a simplified toy example to illustrate the dimensions of all variables in our TRGDL algorithm in Section VII.

II. Computational Complexity

In each iteration, the computational complexity of the suggested TRGDL model mainly contains two main parts. 1) Considering fixed dictionaries and solving four sub-problems including $\mathcal{L}^k$, $\mathcal{A}^k$, $\mathcal{Q}^k$, and $\mathcal{M}^k$ subproblems for each group tensor (k=1, 2… , K) . 2) Dictionary updating. In the first part, the main computational complexity lies in the solving the $\mathcal{L}^k$ sub-problem and updating $\mathcal{Z}^k$. Assuming that we have a tensor $\mathcal{Y}^k \in R^{I_1 \times I_2 \times I_3}$ where $I_1=I_2=I_3=I$, TR-rank is considered as $R_1=R_2=R_N=R$, the computational cost of updating the $\mathcal{L}^k$ is $O\left(R^6+R^4I^2+R^2I^3\right)$, and therefore the overall complexity for K group tensors is about $O\left(K(R^6+R^4I^2+R^2I^3)\right)$. In the second part, the complexity of updating $\mathbf{D}_S$ and $\mathbf{D}_T$ is $O\left(P_{total}(\omega_S pp)^2+P_{total}(\omega_T V)^2\right)$.

III. Convergence Analysis

In this section, we have provided an empirical examination to depict the convergence behavior of the suggested TRGDL denoising method. For this purpose, we conducted our experiments on three datasets and recorded the value of objective function in Eq. (6) (in the paper) at each iteration for all subjects. Fig. S1 illustrates the evaluations of objective function values versus the iterations for a sample subject in each dataset. In the three curves, we can clearly observe that with iteration numbers increasing, the functional energy value monotonically decreases, which proves a good convergence property in our suggested TRGDL model. We conduct this examination on all subjects in three datasets and the same behavior is observed.

[*]Corresponding author

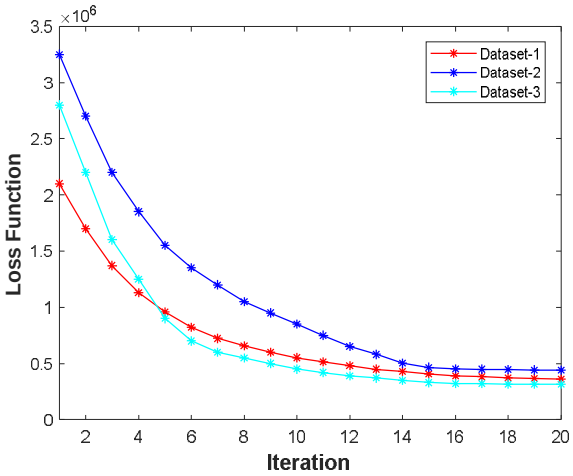


Fig.S1 Empirical convergence curve of objective function against iterations for a sample subject in three OCT datasets i.e., Dataset-1, Dataset-2, Dataset-3

IV. Analysis of the number of OCT slices of Dataset-2 and Dataset-3

To show the effect of the number of the B-scans in our algorithm, we try to use our algorithm on 3, 4, and 5 B-scans. We run these experiments for dataset-2 and dataset-3. Tables S1 and S2 present the average quantitative results on dataset-2 and dataset-3 for CNR and MSR, obtained from our proposed TRGDL model using 3, 4, and 5 nearby B-scans.

Table.S1 Mean of the CNR, MSR results for denoising retinal images in dataset-2 using 3, 4, and 5 B-scans in our proposed method

	3-Bscans as input	**4-Bscans as input**	**5-Bscans as input**
CNR	5.18	5.71	**5.96**
MSR	10.99	12.90	**14.06**

Table.S2 Mean of the CNR, MSR results for denoising retinal images in dataset-3 using 3, 4, and 5 B-scans in our proposed method

	3-Bscans as input	**4-Bscans as input**	**5-Bscans as input**
CNR	4.67	5.02	**5.13**
MSR	11.37	13.26	**14.29**

As Tables S1 & S2 show, the quantitative results of our proposed model using 5 nearby B-scans is better than the TRGTDL model with 3- and 4-Bscans in terms of CNR, and MSR.

V. Results for OCT Images Denoising of Dataset-3

This section provides qualitative and quantitative results for the noise suppression experiment of dataset-3. Fig.S2 and Table. S3 show results of all compared techniques. As Fig.S2 shows, the proposed TRGDL has the most pleasing performance in maintaining layer structures and reducing the noise.

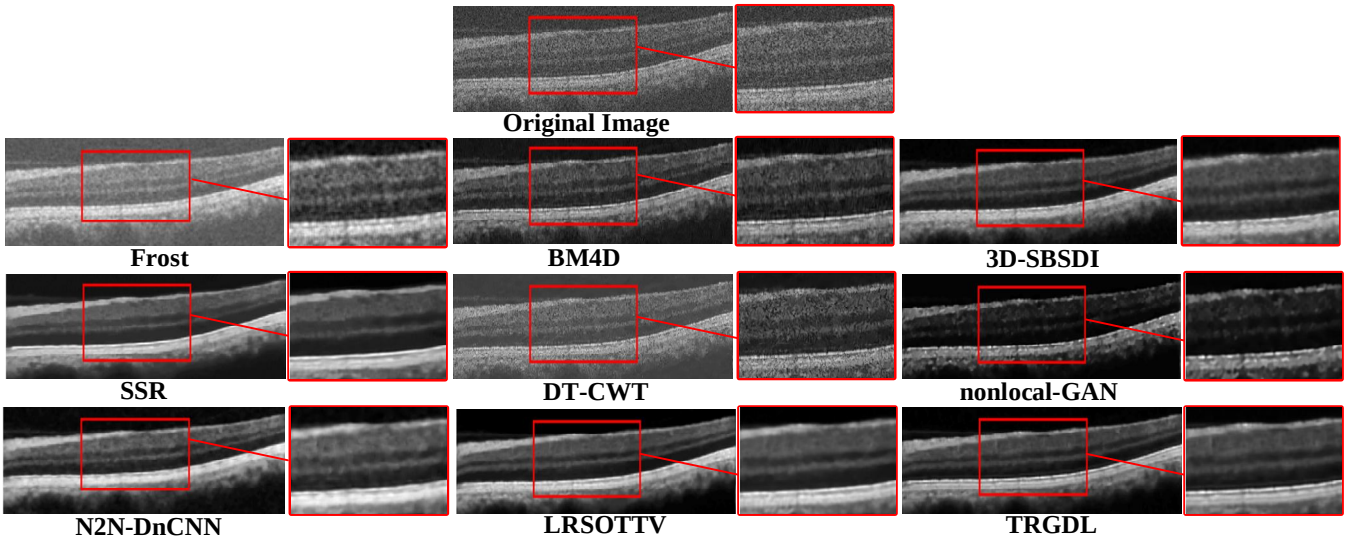


Fig.S2 One retinal OCT image in the dataset-3 and their denoising results utilizing the Frost [1], BM4D [2], 3D-SBSDI [3], SSR [4], DT-CWT [5], nonlocal-GAN [6], N2N-DnCNN [7], LRSOTTV [8], and the suggested TRGDL approach.

Moreover, this method outperforms other rival methods in terms of CNR, and MSR. In addition, the result of all examined approaches is close to each other in EP, and our method's EP value holds the third position among all approaches, closely approaching the top score attained by BM4D. This underscores the effectiveness of the suggested TRGDL in preserving layer structures in OCT images. Table. S3 further illustrates a significant improvement in all metrics when comparing the suggested TRGDL model to other techniques.

Table.S3 Mean & Standard deviation of the MSR, CNR, and EP results for denoising dataset-3 by various approaches. Best results are bolded. Where $p < 0.05$, the measures for each test approach are statistically significant and characterized by a* symbol.

	CNR		MSR		EP	
	Mean± SD	**P-value**	**Mean± SD**	**P-value**	**Mean± SD**	**P-value**
Noisy	2.24±0.29	4.00E-05*	4.66±0.31	4.00E-05*	1.0000± 0.0000	4.00E-05*
Frost	4.12±0.84	4.00E-05*	9.34±1.24	4.00E-05*	0.9298±0.0394	4.00E-05*
BM4D	3.93±0.70	4.00E-05*	8.59± 0.96	4.00E-05*	**0.9673**±0.0220	4.00E-05*
3D-SBSDI	5.01±1.26	1.93E-03*	12.82±1.99	2.01E-04*	0.9426±0.0340	8.31E-04*
SSR	4.71±1.27	9.83E-04*	13.23±2.13	6.08E-03*	0.9433±0.0345	8.94E-03*
DT-CWT	3.47±0.55	4.00E-05*	6.99±0.68	4.00E-05*	0.9641±0.0142	4.00E-05*
nonlocal-GAN	2.91±0.64	4.00E-05*	13.37± 2.23	4.95E-02*	0.9425±0.0347	9.27E-04*
LRSOTTV	5.04±1.27	4.06E-03*	13.30± 2.23	1.73E-03*	0.9359±0.0396	3.13E-04*
N2N-DnCNN	4.96±1.24	1.54 E-03*	13.09± 2.58	3.30E-03*	0.9345±0.0366	4.00E-05*
TRGDL	**5.13**±1.26	-	**14.29**±1.99	-	0.9460±0.0320	-

VI. Comparison of the proposed TRGDL method with the averaging method

To compare the result of our TRGDL method with the averaging method, we utilized one additional dataset, namely, dataset-4. This dataset is a large public dataset which is introduced in [9]. This dataset contains several volumes of size 640×400×400 voxels that were acquired using the RTVueXR SD-OCT system.

First, we randomly chose 30 normal volumes from the large dataset. Then, to implement the averaging strategy, in each volume, we selected one B-scan and 19 noisy adjacent B-scans around each chosen B-scan. Next, we obtained the averaged image from these 20 overlapping B-scans using the averaging method.

To implement our TRGDL method, we considered each chosen B-scan in above and considered 4 noisy adjacent B-scans around each chosen B-scan. Accordingly, 30 normal volumes of size 640×400×5 voxels are considered as input of our algorithm.

Fig. S3 & S4 display the visual denoising results of the proposed TRGDL method and the averaging method on two representative OCT images of dataset-4.

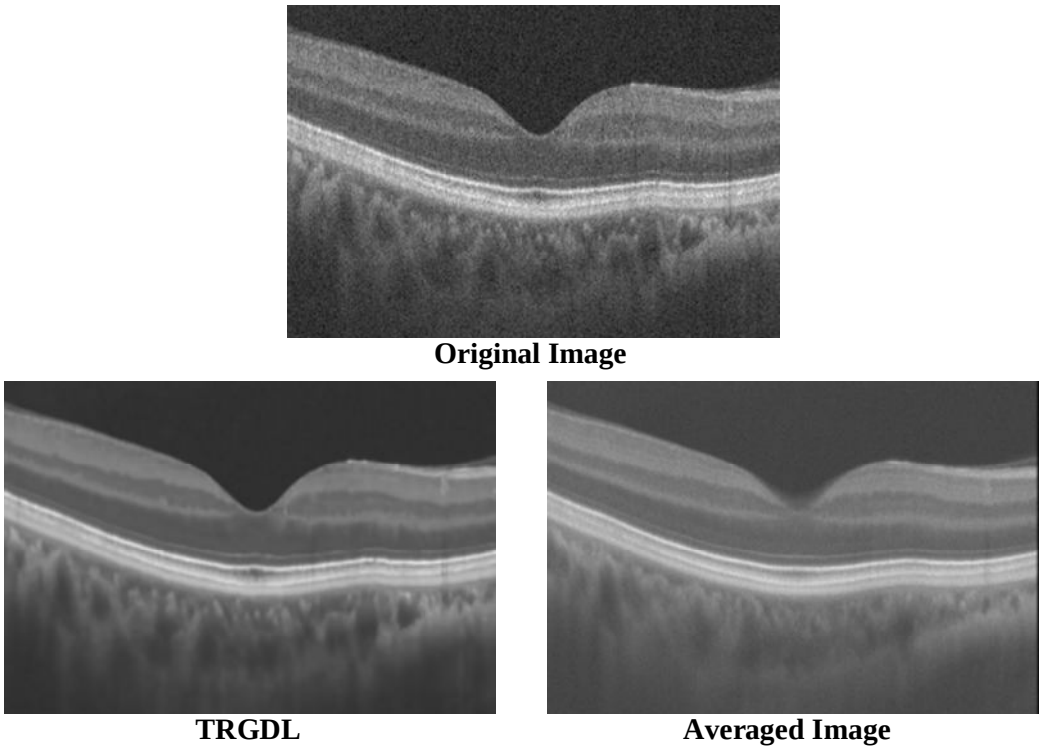
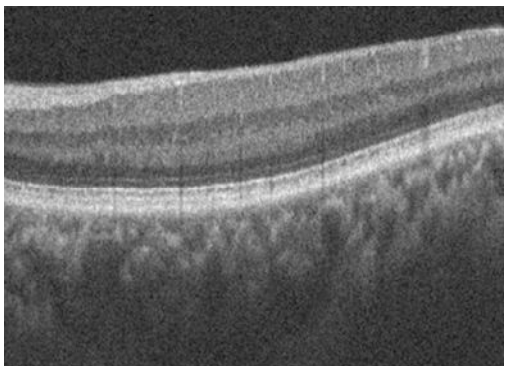
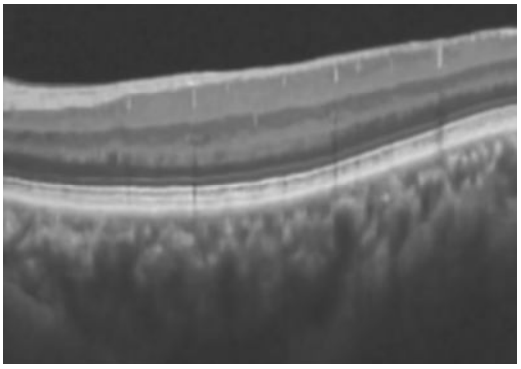


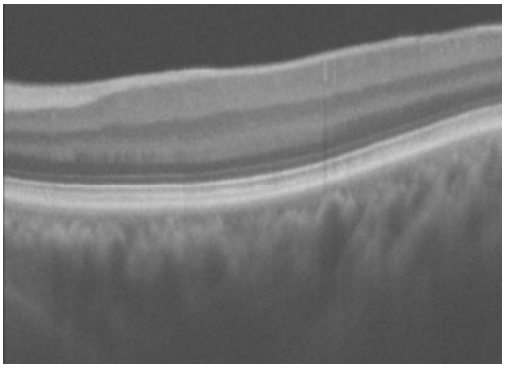
Fig.S3 One retinal OCT image in dataset-4 and their denoising results utilizing the suggested TRGDL approach and the averaging method



Original Image



TRGDL



Averaged Image

Fig.S4 One retinal OCT image in dataset-4 and their denoising results utilizing the suggested TRGDL approach and the averaging method

A visual comparison of the results of our suggested TRGDL model and the average method reveals that both approaches remove the speckle noise but the average method vanishes the details and meaningful structures. Moreover, Table S4 shows the quantitative results. As Table S4 shows, our TRGDL method is better than the average method in terms of CNR. And the average method has a higher MSR value than our method.

Table S4 Mean & Standard deviation of the MSR, CNR results for denoising 30 retinal images in the dataset-4 by TRGDL method and the averaging method

	CNR	**MSR**
	Mean± SD	**Mean± SD**
Noisy	4.25± 0.47	6.60±0.60
Average Image	5.72±0.95	**10.10**±1.41
TRGTDL	**6.13**±0.97	9.55±1.28

VII. Dimensional Insight: A Toy Example Demonstrating the Variables in Our Algorithm

In this section, we provide a simplified toy example to illustrate the dimensions of all variables in our algorithm. Suppose we have a dataset with a size of 21×21×5 (the spatial height ×spatial width ×the number of slices). We extract the cubic patches with the 5-pixel step and size 7×7×5 from it. Therefore, we will have 64 cubic patches. For patch clustering, we first unfolded each cubic patch as a matrix of size 49×5 and then we clustered the similar patches into K groups. Consider we have K=3 groups (or clusters), and each group, denoted as $\mathcal{Y}^{k} \in R^{49\times5\times m^{(k)}}$, includes $m^{(k)}$ similar patches. For k=1, 2, and 3, $m^{(k)}$ is equal to 20, 20, and 24, respectively. Suppose we want to learn spatial and temporal dictionaries with redundancy ratios of 1.2 and 1.4, respectively (i.e., $\mathbf{D}_S \in R^{49\times58}$ and $\mathbf{D}_T \in R^{5\times7}$). Additionally, we set the TR rank as [6, 4, 3].

In our algorithm, the TR decomposition model is applied to each group. Let's focus on one group, for example first group (i.e., $\mathcal{Y}^{1} \in R^{49\times5\times20}$). Table S5 summarizes the dimensions of key variables in Eqs. (21)- (31) for updating first group.

Table S5 Dimensions of key variables in Eqs. (21)-(31) for the Toy Example

Variable	Dimension	Title
$\mathcal{Y}^{(1)}$	49×5×20	First noisy group
$[R_1, R_2, R_3]$	[6, 4, 3]	TR rank
$\mathbf{D}_S$	49×58	Spatial dictionary
$\mathbf{D}_T$	5×7	Temporal dictionary
$\mathcal{A}^{(1)}$	58×7×20	Sparse coefficients tensor
$\left(\mathcal{Z}^{1}\right)^{(1)}$	6×49×4	First TR factor
$\left(\mathcal{Z}^{1}\right)^{(2)}$	4×5×3	Second TR factor
$\left(\mathcal{Z}^{1}\right)^{(3)}$	3×20×6	Third TR factor
$\left[\mathcal{Z}^{1}\right]$	6×4900×6	multilinear product of three TR factors
$\hat{\mathcal{L}}^{1} = \Psi(\left[\mathcal{Z}^{1}\right])$	49×5×20	low-rank approximation of the first group tensor
$\mathbf{D} = \mathbf{D}_S \otimes \mathbf{D}_T$	245×406	Kronecker product of two dictionaries
$\mathbf{Y}^{1}_{(3)}$	20×245	Unfolding $\mathcal{Y}^{(1)}$ along the third mode
$\mathbf{A}^{1}_{(3)}$	20×406	Unfolding $\mathcal{A}^{(1)}$ along the third mode
$\mathcal{Q}^{(1)}$	58×7×20	auxiliary variable

The updating process continues and the variables corresponding to three groups are updated according to the above table. To update $\mathbf{D}_S$ and $\mathbf{D}_T$, we consider $\mathcal{T}^{k} = \frac{\mathcal{Y}^{k} + \alpha_r \mathcal{L}^{k}}{1+\alpha_r}$. Then, we stack three groups. Table S6 summarizes the dimensions of key variables in Eqs. (32)- (36).

Table S6 Dimensions of key variables in Eqs. (32)-(36) for the Toy Example

Variable	Dimensions	Title
$\mathcal{T} = \left(\mathcal{T}^{1}, \mathcal{T}^{2}, \mathcal{T}^{3}\right)$	49×5×64	stack $\mathcal{T}^{(1)}$, $\mathcal{T}^{(2)}$and $\mathcal{T}^{(3)}$ at the 3rd mode
$\mathcal{A} = \left(\mathcal{A}^{1}, \mathcal{A}^{2}, \mathcal{A}^{3}\right)$	58×7×64	stack $\mathcal{A}^{(1)}$, $\mathcal{A}^{(2)}$and $\mathcal{A}^{(3)}$ at the 3rd mode
$\mathcal{F}$	58×5×64	$\mathcal{F} = \mathcal{A} \times_2 \mathbf{D}_T$
$\mathbf{F}_{(1)}$	58×320	Unfolding $\mathcal{F}$ along the first mode
$\mathbf{T}_{(1)}$	49×320	Unfolding $\mathcal{T}$ along the first mode
$\mathcal{H}$	49×7×64	$\mathcal{H} = \mathcal{A} \times_1 \mathbf{D}_S$
$\mathbf{H}_{(2)}$	7×3136	Unfolding $\mathcal{H}$ along the second mode
$\mathbf{T}_{(2)}$	5×3136	Unfolding $\mathcal{T}$ along the second mode

Bibliography

[1] V. S. Frost, J. A. Stiles, K. S. Shanmugan, and J. C. Holtzman, "A model for radar images and its application to adaptive digital filtering of multiplicative noise," *IEEE Transactions on pattern analysis and machine intelligence*, no. 2, pp. 157-166, 1982.

[2] M. Maggioni, V. Katkovnik, K. Egiazarian, and A. Foi, "Nonlocal transform-domain filter for volumetric data denoising and reconstruction," *IEEE transactions on image processing*, vol. 22, no. 1, pp. 119-133, 2012.

[3] L. Fang *et al.*, "Fast acquisition and reconstruction of optical coherence tomography images via sparse representation," *IEEE transactions on medical imaging*, vol. 32, no. 11, pp. 2034-2049, 2013.

[4] L. Fang, S. Li, D. Cunefare, and S. Farsiu, "Segmentation based sparse reconstruction of optical coherence tomography images," *IEEE transactions on medical imaging*, vol. 36, no. 2, pp. 407-421, 2016.

[5] P. Arun, V. P. Gopi, and P. Palanisamy, "Despeckling of OCT images using DT-CWT based fusion technique," *Optik*, vol. 263, p. 169332, 2022.

[6] A. Guo, L. Fang, M. Qi, and S. Li, "Unsupervised denoising of optical coherence tomography images with nonlocal-generative adversarial network," *IEEE Transactions on Instrumentation and Measurement*, vol. 70, pp. 1-12, 2020.

[7] B. Qiu *et al.*, "Comparative study of deep neural networks with unsupervised Noise2Noise strategy for noise reduction of optical coherence tomography images," *Journal of Biophotonics*, vol. 14, no. 11, p. e202100151, 2021.

[8] P. G. Daneshmand, A. Mehridehnavi, and H. Rabbani, "Reconstruction of optical coherence tomography images using mixed low rank approximation and second order tensor based total variation method," *IEEE Transactions on Medical Imaging*, vol. 40, no. 3, pp. 865-878, 2020.

[9] M. Li *et al.*, "OCTA-500: A Retinal Dataset for Optical Coherence Tomography Angiography Study," *arXiv e-prints*, p. arXiv: 2012.07261, 2020.