

# SINGLE IMAGE SUPER-RESOLUTION OF NOISY 3D DENTAL CT IMAGES USING TUCKER DECOMPOSITION

*J. Hatvani*<sup>†\*</sup>   *A. Basarab*<sup>†</sup>   *J. Michetti*<sup>†</sup>   *M. Gyöngy*<sup>\*</sup>   *D. Kouamé*<sup>†</sup>

<sup>†</sup>Université Paul Sabatier Toulouse 3  
Institut de Recherche en Informatique de Toulouse  
118 r. de Narbonne, F-31062 Toulouse

<sup>\*</sup> Pázmány Péter Catholic University  
Faculty of Information Technology and Bionics  
50/a Práter utca, H-1083 Budapest

## ABSTRACT

Tensor decomposition has proven to be a strong tool in various 3D image processing tasks such as denoising and super-resolution. In this context, we recently proposed a canonical polyadic decomposition (CPD) based algorithm for single image super-resolution (SISR). The algorithm has shown to be an order of magnitude faster than popular optimization-based techniques. In this work, we investigated the added value brought by Tucker decomposition. While CPD allows a joint implementation of the denoising and deconvolution steps of the SISR model, with Tucker decomposition the denoising is realized first, followed by deconvolution. This way the ill-posedness of the deconvolution caused by noise is partially mitigated. The results achieved using the two different tensor decomposition techniques were compared, and the robustness against noise was investigated. For validation, we used dental images. The superiority of the proposed method is shown in terms of peak signal-to-ratio, structural similarity index, the canal segmentation accuracy, and runtime.

**Index Terms**— 3D single image super-resolution, tensor factorization, Tucker decomposition, dental CT

## 1. INTRODUCTION

Single image super-resolution (SISR) techniques aim to improve the observed image without further measurements, offering a safe and cost-effective reconstruction method. This is especially true in medical imaging, where additional radiation is to be avoided. For instance in dentistry, the position and structure of the tooth canal is determined using cone-beam computed tomography (CBCT), and is outstandingly important in case of routine root canal treatments, where the current success rate of 60-85% [1, 2] could be improved with SISR.

The super-resolution problem assumes an image degradation model where the high-resolution (HR) image is convolved by a blurring kernel, downsampled, and corrupted by additive noise, resulting in the low-resolution (LR) image.

The problem is further complicated by its dimensionality, as the processing of 3D images is computationally more challenging. This ill-posed inverse problem was addressed in the literature by minimization techniques using different well-known regularizers [3–6], or deep learning techniques [7–9]. While the first group of algorithms is computationally inefficient as a result of large matrix operations, deep learning techniques require large training sets, which are usually difficult to obtain for medical applications.

In our previous work [10] an SISR technique (TF-SISR) using canonical polyadic decomposition (CPD) was proposed, adapting the idea of [11] used in multispectral-hyperspectral image fusion. In that work promising results were obtained for high SNR. In this paper an SISR method is proposed based on a different tensor factorization, the Tucker decomposition (TD), already used in multi- and hyper-spectral imaging [12]. The proposed algorithm, denoted by TD-SISR, is shown to be more accurate and more robust to noise than the previous TF-SISR method.

In the next section, the two tensor decomposition techniques are presented. Section 3 summarizes the TF-SISR algorithm and presents the TD-based SISR. Section 4 compares the results of the proposed technique against TF-SISR for simulated and real CBCT data, and shows its robustness to noise. Sections 5 and 6 provide a discussion about the results, and give conclusive remarks and perspectives of this work.

## 2. TENSOR OPERATIONS

A 3D image can be defined as a third-order tensor  $\mathbf{X} \in \mathbb{R}^{I \times J \times K}$ . Its mode- $n$  fibers are the analogues of columns and rows ( $X(:, j, k)$ ,  $X(i, :, k)$  and  $X(i, j, :)$ ).

$\mathbf{X}$  can be multiplied by a 2D matrix along each of its  $n$  dimensions ( $P_1 \in \mathbb{R}^{I^* \times I}$ ,  $P_2 \in \mathbb{R}^{J^* \times J}$ ,  $P_3 \in \mathbb{R}^{K^* \times K}$  |  $I^*, J^*, K^* \in \mathbb{Z}$  respectively), called the mode- $n$  products ( $\times_n$ ) [10]

$$\mathbf{T} = \mathbf{X} \times_1 P_1 \times_2 P_2 \times_3 P_3 =$$

$$\sum_{i=1}^I \sum_{j=1}^J \sum_{k=1}^K X(i, j, k) P_1(:, i) \circ P_2(:, j) \circ P_3(:, k), \quad (1)$$

This research has been partially supported by Pázmány University KAP19-20.

where  $\mathbf{T} \in \mathbb{R}^{I^* \times J^* \times K^*}$ , and  $\circ$  is the outer product.

The tensor unfolding flattens a tensor into a 2D matrix. It can happen along any of the modes; the mode- $n$  fibers are ordered alphabetically as columns of the matrix.

## 2.1. Canonical Polyadic Decomposition

The CPD [13] used in our previous work [10] factorizes the tensor  $\mathbf{X}$  as a sum of  $R$  rank-1 tensors (such tensors can be written as the outer product of three 1D arrays, here  $U^n(:, r)|n = 1, 2, 3$ )

$$\mathbf{X} \approx \sum_{r=1}^R U^1(:, r) \circ U^2(:, r) \circ U^3(:, r), \quad (2)$$

where  $\bar{U} = \{U^1 \in \mathbb{R}^{I \times R}, U^2 \in \mathbb{R}^{J \times R}, U^3 \in \mathbb{R}^{K \times R}\}$  is a set of three 2D matrices, known as the CPD of the tensor  $\mathbf{X}$ . The minimal number of such rank-1 tensors that can exactly compose the original tensor is called the rank of the tensor,  $F$ . The tensor rank is NP-hard to find, but using it a unique decomposition can be achieved under mild conditions [14]. The shorthand of this decomposition is  $\mathbf{X} = \llbracket U^1, U^2, U^3 \rrbracket$ .

## 2.2. Tucker Decomposition

The  $n$ -rank of a tensor is the number of its independent mode- $n$  fibers. For a third-order tensor three such values,  $\{1\text{-rank}, 2\text{-rank}, 3\text{-rank}\}$  can be defined. While the CPD defines the rank of a tensor,  $F$ , for the TD the  $n$ -rank is a more practical tool - but the two terms should not be confused. TD (also called higher order SVD, or multi-linear SVD) is the multidimensional generalization of the 2D SVD, written as

$$\mathbf{X} \approx \Sigma \times_1 V_1 \times_2 V_2 \times_3 V_3, \quad (3)$$

where  $\Sigma$  is a core tensor, and  $V_1 \in \mathbb{R}^{I \times R_1}, V_2 \in \mathbb{R}^{J \times R_2}, V_3 \in \mathbb{R}^{K \times R_3}$  are the orthonormal bases of the subspaces spanned by the mode- $n$  fibers [13]. If  $R_1, R_2, R_3$  equal the  $n$ -ranks of  $\mathbf{X}$ , the decomposition is unique, but for lower numbers and in the presence of noise it becomes inexact. As opposed to the 2D SVD,  $\Sigma$  is not diagonal, but shows the level of interaction between the different modes. Due to orthonormality, it can be obtained as

$$\Sigma = \mathbf{X} \times_1 V_1^T \times_2 V_2^T \times_3 V_3^T, \quad (4)$$

The shorthand of TD is  $\mathbf{X} = \llbracket \Sigma; V_1, V_2, V_3 \rrbracket$ . Singular values similar to the 2D case ( $SV_n$  for  $n = 1, 2, 3$ ) can be defined. These are calculated as the Frobenius norms of 2D slices taken from the core tensor  $\Sigma$ , along mode- $n$ , fixing the index of the component in question.

# 3. METHODS

## 3.1. Problem formulation

The SISR image degradation model considered herein assumes that the HR image  $\mathbf{X}$  is blurred by a separable (usually

Gaussian) kernel  $h$ , is downsampled by rate  $r$ , and corrupted by additive white Gaussian noise  $\mathbf{N}$ . The output LR image  $\mathbf{Y} \in \mathbb{R}^{I/r \times J/r \times K/r}$  is given by the vector-matrix equation

$$\text{vec}(\mathbf{Y}) = D H \text{vec}(\mathbf{X}) + \text{vec}(\mathbf{N}), \quad (5)$$

where  $H$  is a block-circulant matrix with circulant blocks defined from  $h$ ,  $D$  is an  $r$ -fold downsampling operator, and  $\text{vec}(\cdot)$  vectorizes a tensor alphabetically. In this work the blurring kernel is assumed to be known, and is estimated for the current application as explained in [9].

## 3.2. Previous TF-SISR using CPD

In our previous work [10] the denoising, deconvolution and upsampling of  $\mathbf{Y}$  were implemented jointly, separated for the 3 dimensions. As both the blurring and downsampling operations are separable to the 3 dimensions ( $D_1, D_2, D_3$  and  $H_1, H_2, H_3$ ), the alternating least squares algorithm finding the optimal CPD can be realized with

$$\begin{aligned} U^1 &= (D_1 H_1)^\dagger \mathbf{Y}^{(1)} (D_3 H_3 U^3 \odot D_2 H_2 U^2)^\dagger{}^T \\ U^2 &= (D_2 H_2)^\dagger \mathbf{Y}^{(2)} (D_3 H_3 U^3 \odot D_1 H_1 U^1)^\dagger{}^T \\ U^3 &= (D_3 H_3)^\dagger \mathbf{Y}^{(3)} (D_2 H_2 U^2 \odot D_1 H_1 U^1)^\dagger{}^T, \end{aligned} \quad (6)$$

as derived in [10]. These steps are repeated until  $\bar{U}$  converges.  $\mathbf{Y}^{(n)}$  is the  $n$ -unfolded tensor, and  $\odot$  also implements unfolding, only in the CPD-domain. Denoising is realized by choosing a small enough  $R$ . The symbol  $\dagger$  denotes the Moore-Penrose pseudoinverse with Tikhonov regularization

$$A^\dagger = (A^T A + \epsilon I)^{-1} A^T. \quad (7)$$

## 3.3. Proposed TD-SISR

The idea of the method proposed herein is to denoise the image before deconvolution, in order to stabilize the ill-posed operation, as earlier suggested in the literature [15].

As explained in Section 2.2, the singular values of each mode ( $SV_1, SV_2, SV_3$ ) can be calculated from  $\Sigma$ . Similarly to the 2D case, by picking the relevant components having a singular value higher than a threshold  $R_n$ , a denoised version of  $\mathbf{Y}$ ,  $\hat{\mathbf{Y}}$  may be achieved [16].

$$\hat{\mathbf{Y}} = \bar{\Sigma} \times_1 \bar{V}_1 \times_2 \bar{V}_2 \times_3 \bar{V}_3$$

, where

$$\begin{aligned} \bar{V}_n &= V_n(:, i) \quad | \quad SV_n(i) \geq R_n \\ \bar{\Sigma} &= \mathbf{Y} \times_1 \bar{V}_1^T \times_2 \bar{V}_2^T \times_3 \bar{V}_3^T \end{aligned} \quad (8)$$

Unlike in the 2D case, this truncated approximation might not be optimal in the least squares sense, but gives a reasonable estimate [13].

After obtaining the denoised image, the deconvolution is realized using a Tikhonov-regularized deconvolution separated for the three modes.

$$\hat{\mathbf{X}} = \hat{\mathbf{Y}} \times_1 (D_1 H_1)^\dagger \times_2 (D_2 H_2)^\dagger \times_3 (D_3 H_3)^\dagger. \quad (9)$$

## 4. RESULTS

### 4.1. Datasets and metrics

HR dental images were acquired using a QuantumFX micro-CT system (Perkin Elmer, resolution 1 LP/mm at 50% MTF, meaning that spatial frequencies of 1 line pair per mm are depicted at 50% of the modulation transfer function). In simulation these images were blurred (with a Gaussian kernel, standard deviation  $\sigma_1=\sigma_2=\sigma_3=8$ ), downsampled (at  $r=2$ ) and Gaussian noise was added (white noise at different SNR levels). For simulations, a lower premolar was chosen (280268492 pixels). For real data experiments the LR images were CBCT images acquired with a Carestream 81003D system (resolution 10 LP/mm at 50% MTF). Following [9], a Gaussian blurring kernel was estimated from the CBCT data, and served as an input for the SISR algorithms. For this application, the estimated standard deviations of the Matlab Gaussian functions were  $\sigma_1=8.2$ ,  $\sigma_2=7.5$ ,  $\sigma_3=1.3$ . The volumes naturally contain noise because of the measurements and of the reconstruction algorithm. However, the level of noise is low for extracted teeth due to the absence of surrounding structures. To mitigate this and to further explore the robustness of the SISR algorithms to noise, Gaussian noise corresponding to different SNRs was artificially added to the experimental data. The source of the real data was an upper molar (324248442 pixels).

The improvement of the LR images was measured through their peak signal-to-noise ration (PSNR) and structural similarity index (SSI) [17], calculated against the micro-CT HR images (which are considered as the ground truth due to their high spatial resolution and SNR, but unavailable in clinics because of the excessive radiation dose). The metrics were calculated for the teeth only, excluding the background.

The efficiency of the algorithms was also measured on the segmented root canals following a dedicated method developed in [18], calculating their volume and Dice coefficient [19] against the segmented HR micro-CT volumes.

The algorithms were implemented in Matlab 2017 [20], and for basic tensor operations Tensorlab [21] was used.

### 4.2. Simulation results

For TF-SISR  $R = 500$  was chosen, and in both methods  $\epsilon = 1$  was set for (7) following [10]. As it can be seen in Fig. 2, the singular values decay rapidly (mind the logarithmic scale). For TD-SISR  $R_1=R_2=R_3=40$  (SV values under 1) were chosen as they were generally sufficient for all noise levels.

Table 1 shows the quantitative results of the SISR methods. The PSNR is improved for each case and both methods, compared to the simulated LR image. TD-SISR gave better results, except for the extremely noisy, 20 dB case. The SSI gave similar results, with no improvement in the 20 dB case. After these results the segmentation was carried out

at 25 dB. The improvement is confirmed by the Dice coefficients, showing the superiority of the TD-SISR method.

**Table 1. Metrics in simulation**

	Simulated LR	TF-SISR	TD-SISR
runtime	-	17.96 s	<b>1.86 s</b>
PSNR (dB)			
no noise	28.56	31.48	<b>34.99</b>
30 dB	28.45	31.17	<b>34.39</b>
25 dB	28.36	31.08	<b>31.40</b>
20 dB	27.98	<b>30.01</b>	29.33
SSI [0, 1]			
no noise	0.9623	0.9680	<b>0.9823</b>
30 dB	0.9612	0.9650	<b>0.9763</b>
25 dB	0.9572	0.9595	<b>0.9653</b>
20 dB	<b>0.9463</b>	0.9453	0.9417
Segmentation at 25 dB			
Dice	0.8976	0.9242	<b>0.9425</b>

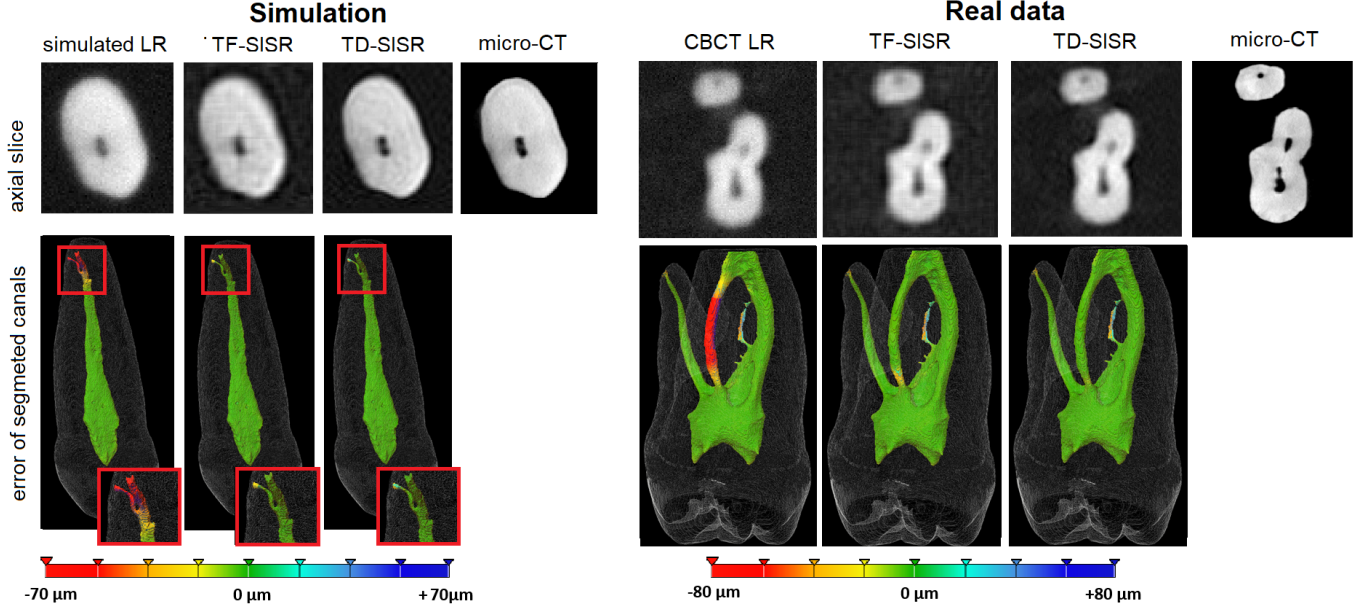
### 4.3. Real data results

For the real data in TF-SISR the same settings were used as in simulation, and for TD-SISR  $R_1, R_2, R_3 = 50$  was set after plotting the SVs (Fig. 3).

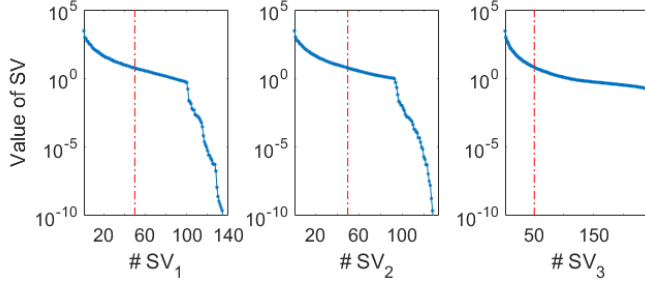
The metrics have shown milder improvements compared to the simulation, but both PSNR and SSI improved in all cases with both methods, and TD-SISR gave superior results. The volume of the segmented 25 dB images also improved regarding the Dice coefficient.

**Table 2. Metrics in simulation**

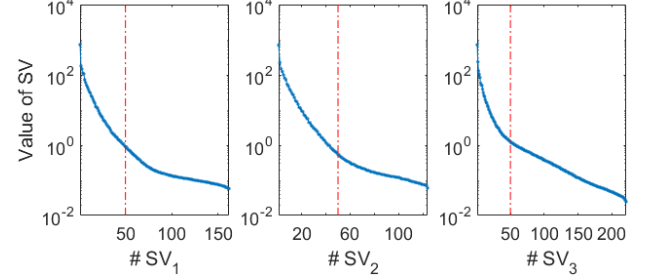
	Simulated LR	TF-SISR	TD-SISR
runtime	-	17.71 s	<b>1.46 s</b>
PSNR (dB)			
no noise	19.55	21.25	<b>21.61</b>
30 dB	19.30	20.84	<b>21.57</b>
25 dB	19.10	20.13	<b>21.09</b>
20 dB	18.91	20.21	<b>20.29</b>
SSI [0, 1]			
no noise	0.8647	0.8907	<b>0.8935</b>
30 dB	0.8610	0.8870	<b>0.8929</b>
25 dB	0.8478	0.8784	<b>0.8908</b>
20 dB	0.8173	0.8555	<b>0.8814</b>
Segmentation at 25 dB			
Dice	0.8939	0.9189	<b>0.9304</b>



**Fig. 1.** Results of SISR methods under 25 dB noise, both in simulation and in real data. The first row shows a single axial slice taken from the volumes. The second row shows the distance between the segmented HR and LR, enhanced LR volumes.



**Fig. 2.** Simulation - singular values without added noise in all three modes, on a logarithmic scale. The vertical lines represent the chosen truncation thresholds



**Fig. 3.** Real data - singular values without added noise in all three modes, on logarithmic scale. The vertical lines represent the chosen truncation thresholds

## 5. DISCUSSION

In contrast to the earlier TF-SISR method no iterations are applicable in TD-SISR. In TD-SISR 3 thresholds  $R_1, R_2, R_3$  have to be defined for the three modes, while in TF-SISR only one parameter,  $R$  influences the denoising step. However, the singular values of TD-SISR correspond to the importance of the components, while  $R$  in TF-SISR bears no such meaning. This makes the setting of TD-SISR parameters easier, and its efficiency is validated by the qualitative and quantitative results. The runtime of TD-SISR is lower because of the lack of iterations, but calculating the SVD for even larger volumes might be a bottleneck [22].

## 6. CONCLUSION

In this paper a new SISR technique was proposed, using Tucker decomposition for the denoising, and a Tikhonov-regularized deconvolution. Even though 2 additional parameters have to be set, it gave faster and quantitatively better results in noisy images compared to our previous method, TF-SISR. Images of 280268492 and 324248442 pixels were super-resolved under 2 s with standard Matlab implementation. In future work, the connection between the TF-SISR and TD-SISR parameters along with their robustness, and the general inverse problem including thresholding constraint will be investigated.

## 7. REFERENCES

- [1] Y-L Ng, V Mann, S Rahbaran, J Lewsey, and K Gulabivala, "Outcome of primary root canal treatment: systematic review of the literature—part 1. effects of study characteristics on probability of success," *International endodontic journal*, vol. 40, no. 12, pp. 921–939, 2007.
- [2] H. M Eriksen, LL. Kirkevang, and K. Petersson, "Endodontic epidemiology and treatment outcome: general considerations," *Endodontic Topics*, vol. 2, no. 1, pp. 1–9, 2002.
- [3] A. Toma, L. Denis, B. Sixou, JB. Pialat, and F. Peyrin, "Total variation super-resolution for 3D trabecular bone micro-structure segmentation," in *Proc. 22nd European Signal Processing Conference (EUSIPCO)*, Lisbon, Portugal, Sept. 2014, pp. 2220–2224.
- [4] F. Shi, J. Cheng, L. Wang, PT. Yap, and D. Shen, "LRTV: MR image super-resolution with low-rank and total variation regularizations," vol. 34, no. 12, pp. 2459–2466, 2015.
- [5] W. Zhang, H. Zhang, L. Wang, A. Cai, L. Li, and B. Yan, "Limited angle CT reconstruction by simultaneous spatial and radon domain regularization based on TV and data-driven tight frame," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 880, pp. 107–117, 2018.
- [6] J. V Manjón, P. Coupé, A. Buades, V. Fonov, D L. Collins, and M. Robles, "Non-local MRI upsampling," *Medical image analysis*, vol. 14, no. 6, pp. 784–792, 2010.
- [7] S. Cengiz, M. d C Valdes-Hernandez, and E. Ozturk-Isik, "Super resolution convolutional neural networks for increasing spatial resolution of <sup>1</sup>h magnetic resonance spectroscopic imaging," in *Proc. 21st MIUA Annu. Conf.*, Edinburgh, UK, July11-13 2017, Springer, pp. 641–650.
- [8] Y. Zhang and M. An, "Deep learning- and transfer learning-based super resolution reconstruction from single medical image," *J. Healthc. Eng.*, vol. 2017, Art. ID 5859727, 20 pages, 2017.
- [9] Janka Hatvani, András Horváth, Jérôme Michetti, Adrian Basarab, Denis Kouamé, and Miklós Gyöngy, "Deep learning-based super-resolution applied to dental computed tomography," *IEEE Transactions on Radiation and Plasma Medical Sciences*, vol. 3, no. 2, pp. 120–128, 2018.
- [10] Janka Hatvani, Adrian Basarab, Jean-Yves Tournet, Miklós Gyöngy, and Denis Kouamé, "A tensor factorization method for 3-d super resolution with application to dental ct," *IEEE transactions on medical imaging*, vol. 38, no. 6, pp. 1524–1531, 2018.
- [11] C. I. Kanatsoulis, X. Fu, N. D. Sidiropoulos, and W. Ma, "Hyperspectral super-resolution: A coupled tensor factorization approach," vol. 66, no. 24, pp. 6503–6517, Dec 2018.
- [12] C. Prévost, K. Usevich, P. Comon, and D. Brie, "Hyperspectral super-resolution with coupled tucker approximation: Recoverability and svd-based algorithms," 2019.
- [13] T. G Kolda and B. W. Bader, "Tensor decompositions and applications," *SIAM review*, vol. 51, no. 3, pp. 455–500, 2009.
- [14] L. Chiantini and G. Ottaviani, "On generic identifiability of 3-tensors of small rank," *SIAM Journal on Matrix Analysis and Applications*, vol. 33, no. 3, pp. 1018–1037, 2012.
- [15] C. Tsutake and T. Yoshida, "Vaguelette-wavelet deconvolution via compressive sampling," *IEEE Access*, vol. 7, pp. 54533–54541, 2019.
- [16] R. James, A. M. Jolly, A. C., and D. Michael, "Image denoising using adaptive pca and svd," in *2015 Fifth International Conference on Advances in Computing and Communications (ICACC)*, Sep. 2015, pp. 383–386.
- [17] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," vol. 13, no. 4, pp. 600–612, Apr. 2004.
- [18] J. Michetti, A. Basarab, F. Diemer, and D. Kouame, "Comparison of an adaptive local thresholding method on CBCT and  $\mu$ CT endodontic images," *Physics in Medicine & Biology*, vol. 63, no. 1, pp. 015020, 2017.
- [19] L. R. Dice, "Measures of the amount of ecologic association between species," *Ecology*, vol. 26, no. 3, pp. 297–302, 1945.
- [20] MATLAB, *version 9.3.0 (R2017a)*, The MathWorks Inc., Natick, Massachusetts, 2017.
- [21] N. Vervliet, O. Debals, L. Sorber, M. Van Barel, and L. De Lathauwer, "Tensorlab 3.0," Mar. 2016, Available online.
- [22] A. K. Menon and C. Elkan, "Fast algorithms for approximating the singular value decomposition," *ACM Trans. Knowl. Discov. Data*, vol. 5, no. 2, Feb. 2011.