

TENDER: TENSOR Non-local Deconvolution Enabled Radiation Reduction in CT Perfusion

Ruogu Fang^{a,*}, Ajay Gupta^b, Junzhou Huang^c, Pina Sanelli^d

^a*School of Computing and Information Sciences, Florida International University, USA*

^b*Department of Radiology, Weill Cornell Medical College, USA*

^c*Department of Computer Science, University of Texas at Arlington, USA*

^d*Department of Radiology, Hofstra Northwell School of Medicine, USA*

Abstract

Stroke is the leading cause of long-term disability and the second leading cause of mortality in the world, and exerts an enormous burden on the public health. CT remains one of the most widely used imaging modality for stroke diagnosis. However when coupled with CT perfusion, the excessive radiation exposure in repetitive imaging to assess treatment response and prognosis has raised significant public concerns regarding its potential hazards to both short- and long-term health outcomes. Tensor total variation has been proposed to reduce the necessary radiation dose in CT perfusion without comprising the image quality by fusing the information of the local anatomical structure with the temporal blood flow model. However the local search in the framework fails to leverage the non-local information in the spatio-temporal data. In this paper, we propose TENDER, an efficient framework of non-local tensor deconvolution to maintain the accuracy of the hemodynamic parameters and the diagnostic reliability in low radiation dose CT perfusion. The tensor total variation is extended using non-local spatio-temporal cubics for regularization to integrate contextual and non-local information. We also propose an efficient framework consisting of fast nearest neighbor search, accelerated optimization and parallel computing to im-

*Corresponding author

Email addresses: rfang@cs.fiu.edu (Ruogu Fang), ajg9004@med.cornell.edu (Ajay Gupta), jzhuang@exchange.uta.edu (Junzhou Huang), psanelli@northwell.edu (Pina Sanelli)

prove the efficiency and scalability of the non-local spatio-temporal algorithm. Evaluations on clinical data of subjects with cerebrovascular disease and normal subjects demonstrate the advantage of non-local tensor deconvolution for reducing radiation dose in CT perfusion.

Keywords: low radiation dose, CT perfusion, tensor non-local deconvolution, total variation, stroke

1. Introduction

Stroke remains the leading cause of disability and second leading cause of mortality worldwide, thus exerting an enormous public health burden for long-term disability care costing \$34 billion annually. Timely diagnosis is crucial to inform treatment decisions in acute stroke as “time is brain”. Therefore, it is critical to develop reliable in-vivo quantitative imaging markers for stroke for diagnosis and guiding treatment decisions. Computed tomography perfusion (CTP) is an ideal imaging modality to assess hemodynamic changes in the brain given its widespread availability, speed, affordability and high spatial resolution. Despite these benefits, CTP delivers one of the highest radiation doses of noninvasive imaging studies, and as a result, has raised significant public concerns because of its potential short- and long-term biological effects including cancer induction, skin damage and early cataract formation [1, 2, 3, 4]. This is especially concerning in stroke patients whose disease requires serial monitoring of hemodynamic status with perfusion imaging to assess treatment response and prognosis [5]. Lowering the radiation exposure would reduce the potential health hazard, improve healthcare quality and safety, and make CTP modality fully utilized for a wider population. However, a major challenge in dose reduction strategies in CTP is that such approaches inevitably lead to an increase in noise and therefore less accurate hemodynamic parameter quantification given the methods used today.

In this paper, we propose an efficient framework TENDER (TEnsor Non-local Deconvolution Enabled Radiation reduction) to minimize the radiation

exposure without compromising the image quality, especially the perfusion parameter accuracy. This fast deconvolution method extends the prior work of tensor total variation (TTV) [6, 7] with non-local regularization to improve the image quality and interpretation of low-dose CTP. Instead of restricting the regularization of residue functions to the adjoining voxels in the spatial domain and neighboring frames in the temporal domain, we consider both long-range dependency and the global connections in the spatial and temporal dimensions. We propose to integrate the state-of-art low-dose deconvolution method together with the non-local regularization to increase the robustness of the method with efficient algorithm. To the best of our knowledge, this integration of two approaches in a spatio-temporal framework to regularize the flow-scaled residue impulse functions has never been proposed, and it can make significant improvement in the perfusion parameter estimation. Furthermore, the efficient algorithm to accelerate TENDER computation and optimization would make the proposed algorithm clinically valuable.

2. Related Work

There are two major steps in hemodynamic parameter estimation for perfusion imaging: reconstruction and deconvolution. Therefore there are mainly two types of approaches to reduce the radiation exposure.

2.1. Reconstruction

Current state-of-the-art methods to reduce radiation exposure from CTP are problematic because they mainly focus on the reconstruction step which does not improve the parameter estimation process [8, 9, 10, 11, 12]. However, the stability and accuracy of the parameter estimation step is critical for the precise quantitative estimation of hemodynamic parameters. Among the few methods to improve the deconvolution step in low-dose CTP, the improvement in performance been limited without leveraging the existing high-dose data [13]. With the advent of big medical data, a data-driven mathematical approach employing a spatio-temporal model has the potential to significantly reduce the radiation exposure in CTP.

2.2. Deconvolution

While the first class of approaches does not solve the inherent instability problem in the quantification (deconvolution) process of CTP, the second class of approaches directly addresses this instability issue. Among these methods, the information redundancy and sparsity is a property that has shed light into medical image analysis, ranging from deformable models [14, 15, 16], segmentation [17], fast MRI quantification [18], and low-dose quantification problems [19, 20, 21], but the sparsity frameworks need training data for dictionary learning. In another line of work, tensor total variation (TTV) deconvolution [6, 7] has been recently proposed to significantly reduce the radiation dosage in CTP with improved robustness and quantitative accuracy by integrating the anatomical structure correlation and the temporal blood flow model. The anatomical structure of the brain encompasses long-range similarity of the same tissue classes, as shown in Fig. 1(a). However the locality property of the current TTV algorithm limits the search for similar patterns in the 4-connected adjacent neighborhood, neglecting the long-range or global correlations of the entire brain structure. This locality limitation has led to noticeable absence or artifact of the delicate structures, such as the capillary, the insula and the parietal lobe, which are critical indicators for the clinical diagnosis of cerebrovascular diseases. Fig. 1(b) shows the importance of accurate depiction of hemodynamic parameters. The delicate vascular and cerebral structures are critical biomarkers of the existence and severity of the cerebrovascular diseases. Naturally, integrating non-local correlation into the estimation process of the hemodynamic parameters would yield more precise depiction of the pathological regions in the brain [22, 23]. As an extension of [22], this work significantly extend introduction, related work and experimental parts.

2.3. Contributions

The contribution of this work is three-fold: First, the non-local connections are explored to leverage the anatomical and structural similarity of the same tissue classes in both the spatial and the temporal dimensions. Second, effi-

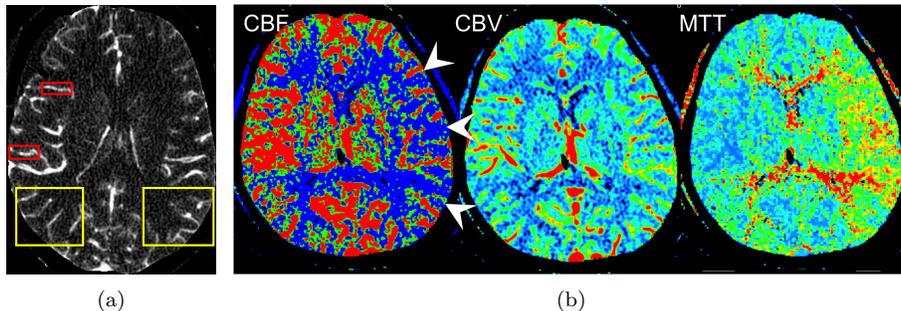


Figure 1: (a) The illustration of long-range similarity in the brain. The red and yellow boxes show the non-local regions which have similar patterns. (b) Perfusion parameter maps (CBF - cerebral blood flow, CBV - cerebral blood volume, and MTT - mean transit time) of a 22-year old with severe left middle cerebral artery (MCA) stenosis. Arrows indicate the regions with altered hemodynamic function represented as abnormally decreased CBF and prolonged MTT. This pattern is indicative of ischemic tissue at risk of stroke in the left hemisphere (right side of the image).

cient parallel implementation and similarity computation using window offsets reduce the computational time of the non-local algorithm. Third, extensive experiments on low-dose CTP clinical data of subjects with cerebrovascular diseases and normal subjects are performed. The experiments demonstrate the superiority of the non-local framework, compared with the local TTV method. The advantages include more accurate preservation of the fine structures and higher spatial resolution for the low-dose data.

This paper is organized as follows. Section 3 presents the current tensor total variation model for low-dose CTP deconvolution with its limitations. Section 4 introduces our proposed TENDER model for tensor non-local deconvolution to overcome the locality constraints in the TTV model, followed by Section 5 on the efficiency method for this spatio-temporal optimization problem, including fast nearest neighbor search in the non-local neighborhood, efficiency optimization, and parallel computing. Section 6 and 7 describes the experiment setup and shows the experimental results on clinical data. The paper is concluded with Section 9.

3. Tensor Total Variation Model

In this section, we will first briefly review the tensor total variation (TTV) model for the low-dose CTP and discuss its deficiency in accurate estimation of delicate structure and examing pattern complexities.

To reduce the radiation dose in CT perfusion imaging, tensor total variation (TTV) [6] is recently proposed to efficiently and robustly estimate the hemodynamic parameters. It integrates the anatomical structure correlation and the temporal continuation of the blood flow signal. The TTV algorithm optimizes a cost function with one linear system for the deconvolution and one smoothness regularization term, as below:

$$K_{TTV} = \arg \min_{K \in \mathbb{R}^{T \times N}} \left(\frac{1}{2} \|AK - C\|_2^2 + \|K\|_{TTV}^\gamma \right). \quad (1)$$

The first term is the temporal convolution model. In this term, $A \in \mathbb{R}^{T \times T}$ is a block-circulant matrix representing the arterial input function (AIF), which is the input signal to the linear time-invariant system of the capillary bed. The block-circulant format makes the deconvolution insensitive to delays in the AIF. $C \in \mathbb{R}^{T \times N}$ is the contrast agent concentration (CAC) curve of all the voxels in the volume of interest (VOI). Both A and C are extracted from the CTP data. $K \in \mathbb{R}^{T \times N}$ is the unknown of this optimization problem - the flow-scaled residue functions of the VOI. Here T is the duration of the signal, and $N = N_1 \times N_2 \times N_3$ is the total number of voxels in the sagittal, coronal and axial directions.

The second term is the tensor total variation regularizer. The TTV regularization is defined as

$$\|K\|_{TTV}^\gamma = \sum_{d=t,x,y,z} \sqrt{\sum_{d=1}^4 (\gamma_d \nabla_d \tilde{K}_{t,x,y,z})^2}, \quad (2)$$

where ∇_d is the forward finite difference operator in the d^{th} dimension, and $\tilde{K} \in \mathbb{R}^{T \times N_1 \times N_2 \times N_3}$ is the 4-dimensional volume reshaped from matrix K with temporal signal for one dimension and spatial signal for three dimensions. t, x, y, z are the indices for the temporal and spatial dimensions. The outside summation means that the square root of the sum of the first order derivative is

summed over all the temporal points t and spatial voxels x, y, z of \tilde{K} . L_1 norm is used in the forward finite difference operator ∇_d to preserve the edges, and the regularization parameters γ_d designates the regularization strength for each dimension. In [6], the temporal regularization parameter $\gamma_t = 10^{-8}$, while the spatial regularization parameter $\gamma_{x,y,z} = 10^{-4}$. Larger γ_d with $d = t, x, y, z$ will lead to higher smoothing regularization for the flow-scaled residue functions \tilde{K} in the d^{th} dimension. Cerebral blood flow (CBF) maps can be computed from K as the maximum value at each voxel over time. More details about the TTV framework can be found in [6].

TTV has achieved significant performance improvement on the digital brain phantom and low- and ultra-low dose clinical CTP data at 30, 15 and 10 mAs [6]. However the locality property of the tensor total variation regularization limits the capability of preserving the small and fine anatomical structures, details and texture in the brain. These structures include the vessels, gray-white matter junction, insular and basal ganglia regions. They are essential indicators of the location and severity of the ischemia or acute stroke. It may also create new distortions, such as blurring, staircase effect and wavelet outliers due to the regularization on the adjacent voxels, as shown in Fig. 2.

4. TENDER Model

Due to the locality limit of the TTV model, we introduce the TENDER model, along with an efficient algorithm to compute the spatio-temporal cubic similarity. We will also discuss the time complexity of the TENDER model.

First introduced by [24], non-local total variation has been studied to address the limitations of conventional total variation model, including the blocky effect, the missing of the small edges and the lack of long-range information sharing [25, 26, 27]. It has also been applied to 4D computed tomography [28] and magnetic resonance imaging reconstruction [29]. This work is the first attempt to integrate non-local tensor total variation with the spatio-temporal deconvolution problem in 4D CTP.

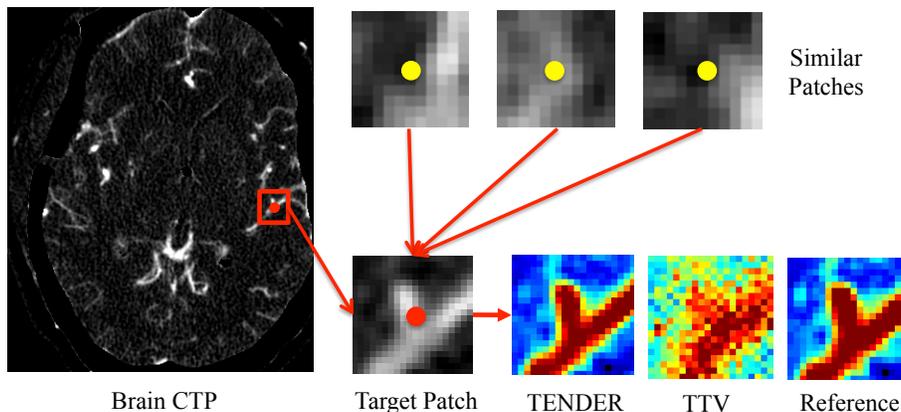


Figure 2: Illustration of TENDER model in a 2D image. TENDER regularization term for voxel i (red dot) is a weighted summation of the difference between voxel i and the most similar voxels (yellow dots) in the search window with width W (red box). The weight $w(i, j)$ depends on the patches around the voxels. Compared to local-TTV, which only considers the 4-connected local neighborhood, TENDER preserves the accuracy and contrast of the vascular structure with higher fidelity of the reference patch. The actual TENDER regularization is imposed on 4D spatio-temporal flow-scaled residue impulse functions across different slices and time points.

The TENDER model links each voxel in the volume with the long-range voxels using a weighted function. For every voxel i , instead of computing the forward finite difference on the 4-connected neighbors, we search in a neighborhood window $N(i)$ with window size W , and minimize the weighted differences between the target voxel and all voxels in the window. Specifically, the non-local tensor total variation can be formulated as:

$$\|K\|_{TENDER} = \sum_i \sqrt{\sum_j (K(i) - K(j))^2 w(i, j)}. \quad (3)$$

Here $K(i)$ denotes the value of flow-scaled residue impulse function K at spatio-temporal voxel i , and $w(i, j)$ is a similarity function between the voxel i and j . The more similar the voxels i and j , the higher the weight function $w(i, j)$ is. We use an exponential function of the patches surrounding the two voxels to model their similarity

$$w(i, j) = \frac{1}{Z(i)} e^{-\frac{\|K(P_i) - K(P_j)\|_2^2}{\sigma^2}}, \quad (4)$$

where Z is a normalization factor, with $Z(i) = \sum_j w(i, j)$ and σ is a filter

parameter that controls the shape of the similarity function. P_i is a small patch around voxel i with radius d . In this way, when two patches are identical or similar, the weight w will be close to 1; when the two patches are very different, the weight w will approach 0. Non-local total variation has shown superior performance signal reconstruction and denoising [25, 26], and by fusing it with the temporal convolution model with a weight α for the non-local spatial term, we get

$$K_{TENDER} = \arg \min_{K \in \mathbb{R}^{T \times N}} \left(\frac{1}{2} \|AK - C\|_2^2 + \alpha \|K\|_{TENDER} \right). \quad (5)$$

The TENDER model searches for the similar patches in a larger window instead of the adjacent 4-connected neighbors in the local TTV. In this way, the similar tissue patterns of the same tissue types in the long-range regions of the brain can assist to reduce the artifact and noise in the deconvolution process. This allows the TENDER model to deconvolve the low-dose CTP volume using non-local dependency by removing the noise without distorting the salient structures, as shown in Fig. 2.

For the weighting parameter α , we find that in the experimental results, $\alpha = 1$ give descent results for the TENDER model, while the optimization result is robust to the parameter α and does not vary significantly at different values of α .

It is worthy to note that because the voxel i is any voxel in the spatio-temporal domain of the flow-scaled residue impulse function $K \in \mathbb{R}^{T \times N}$, TENDER is searching the similar patches in the spatio-temporal domain, which includes the multiple slices in the axial direction and the various time points in the temporal sequences.

5. Efficient Optimization and Time Complexity

We implement this algorithm by MATLAB and C++ using mex in MATLAB 2013a environment (MathWorks Inc, Natick, MA) and Windows 8 operating system with an Intel Core i5 with 4 cores and 32GB RAM.

5.1. Notations

Let's define some parameters first. Let N be the total number of voxels in the entire volume. W be the search window size for the similar voxels around voxel i . d is the radius of the patch around the voxel. N_b is the number of similar voxels chosen to regularize the voxel i in order to speed up the computation. m is the dimension of the spatio-temporal tensor. σ is the Gaussian parameter to control the shape of the similarity function.

In this work, for a 2D slice in the brain CTP data of 512×512 voxels, 120 seconds of scanning duration, $W = 5$ voxels, $d = 4$ voxels, $N_b = 15$, $\sigma = 0.5$. $m = 4$ because the flow-scaled residue impulse functions are spatio-temporal tensor with 4 dimensions.

5.2. Brute-Force Search

The TENDER model has a higher time complexity compared to the local TTV. For each voxel i in the volume, we need to calculate the patch difference between the target voxel and every other voxel in the search window. Then we rank all the patch differences in voxel i 's search window in an ascending order, and pick up the first N_b patches for optimizing the value of i .

The time complexity of the brutal force non-local TTV is $O(N \cdot ((2W + 1)(2d + 1))^m + N \cdot (2W + 1)^m \log(N_b))$. For the parameters above, the computational time reaches up to nearly *10 hours*, which is unrealistic in clinical applications.

5.3. Fast Nearest Neighbor Search

An efficient method to compute the intensity difference between two patches is used to accelerate the non-local TTV is needed. Specifically, at each offset $\vec{w} = (w_x, w_y, w_z, w_t)$ in the search window W , a new matrix D of the same size to the brain volume is created to precompute the patch differences, with $D_{\vec{w}} = \sum_i (K(i + \vec{w}) - K(i))^2$. This matrix keeps the sum of the squared differences from the upper left corner to the current voxel. When computing the differences between the two patches at location j and offset w , we only need to compute the value $D(j_x + d, j_y + d) - D(j_x + d, j_y) - D(j_x, j_y + d) + D(j_x, j_y)$. This

accelerating method to find the nearest neighbors reduced the time complexity to $O(N \cdot (2W + 1)^m + \log(N_b))$. The space complexity is $N \cdot (2W + 1)^m$.

5.4. Efficient Optimization Algorithm

Due to the relatively slow update in the non-local TTV term, we propose a fast TENDER algorithm to optimize the objective function in Eq. (5), as outlined in Algorithm 1. The efficient algorithm consists of three major steps: (1) Gradient descent to update the temporal convolution term; (2) Proximal map to update the spatial non-local total variation term; (3) Parameter update to control the learning rate.

Here we illustrate each step in the proposed efficient algorithm in detail:

(1) **Steepest gradient descent:** Let's denote $f(r^n) = \frac{1}{2} \|Ar^n - C\|^2$, then the gradient of function f at point r^n is $\nabla f(r^n) = A^T(Ar^n - C)$. This step only costs time complexity of $\mathcal{O}(p \log(p))$, where p is the total number of spatio-temporal pixels of the residue impulse functions. Since the Teoplitz matrix A is ill-posed, we design a steepest gradient descent method that updates the step size s adaptively to avoid the issue with a constant step size.

(2) **Proximal map:** The proximal map associated with a continuous convex function $g(x)$ and any scalar $\rho > 0$ is defined as [30]

$$prox_{\rho}(g)(x) := \arg \min_u \{g(u) + \frac{1}{2\rho} \|u - x\|^2\} \quad (6)$$

In this step, $K^n = prox_{\rho}(2\alpha \|K\|_{TENDER})(fold(K_g))$ has a closed form solution and can be computed in time $\mathcal{O}(p \log(p))$, where $fold(K_g)$ folds the matrix K_g into a tensor $\tilde{K} \in \mathbb{R}^{T \times N_1 \times N_2 \times N_3}$. We set $\rho = 1$ in our implementation.

(3) **Parameter update.** Parameters r and t can be updated by adding up scalars or vectors, costing $\mathcal{O}(p)$ and $\mathcal{O}(1)$ respectively.

Therefore, the total cost of each iteration in the proposed TENDER algorithm is approximately $\mathcal{O}(p \log(p))$. Compared to the local total variation method proposed in [6, 7], the proposed TENDER model can effectively leverage the contextual information to avoid the blocky artifacts and preserve the fine structures.

Algorithm 1 The framework of TENDER algorithm.

Input: $K^0 = r^1 = 0, t^1 = C = 0, \tau$
Output: Flow-scaled residue functions $K \in \mathbb{R}^{T \times N_1 \times N_2 \times N_3}$.
for $n = 1, 2, \dots, N$ **do**
 $C = C + 1$
 (1) Steepest gradient descent $K_g = r^n + s^{n+1}A^T(C - Ar^n)$
 where $s^{n+1} = \frac{vec(Q)^T vec(Q)}{vec(AQ)^T vec(AQ)}$, $Q \equiv A^T(Ar^n - C)$, $vec(\cdot)$ vectorizes a matrix
 (2) Proximal map:
 if $C = \tau$ (*Acceleration Step*) **then**
 $K^n = \text{prox}_\gamma(2\alpha\|K\|_{TENDER})(fold(K_g)), C = 0$
 end if
 (3) Update t, r $t^{n+1} = (1 + \sqrt{1 + 4(t^n)^2})/2, r^{n+1} = K^n + ((t^n - 1)/t^{n+1})(K^n - K^{n-1})$
end for

Another key feature of the proposed algorithm is the acceleration step. In the accelerated step, instead of alternating between the non-local TTV term and the temporal convolution term once each iteration, we update the non-local TTV term *fewer* times than updating the temporal convolution term. Specifically, a counter C controls the frequency of non-local TTV term update. Only when the temporal convolution term is updated τ times in the steepest gradient descent step (1), the non-local TTV term is updated once in the proximal map step (2) in Algorithm 1. This acceleration scheme saves the computation time by avoiding frequently updating the non-local TTV term, and has shown sufficient accuracy in the experimental results.

5.5. Parallel Computing

The intrinsic nature of non-local TTV algorithm allows for multi-threading and parallel computing on the multi-core clusters or grids. We divide the entire brain volume into sub-volumes, with each of them processed by one processor. The patch difference computation for every voxel i and the weight calculation for all the voxels after selecting the top N_b neighbors can be paralleled.

6. Experiments

The goal of our proposed method is to accurately estimate the hemodynamic parameters in low-dose CTP by robust deconvolution. To reach this goal, we first simulate the low-dose CTP data from the high-dose data by adding correlated Gaussian noise, and then apply the compared deconvolution methods on the high- and low-dose CTP data. Finally we compare the quantitative and visual results of the estimated perfusion parameter maps. The procedure is illustrated in Fig. 3.

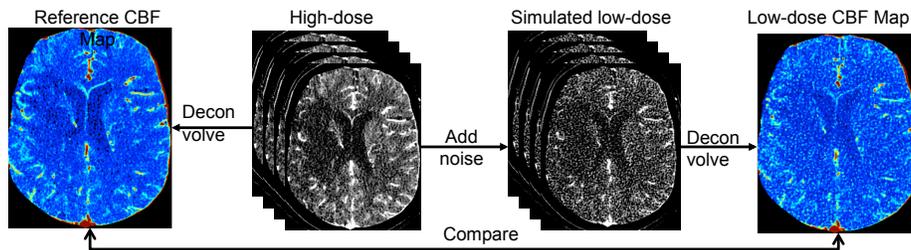


Figure 3: Simulation of low-dose CTP data from high-dose CTP data and the evaluation framework

6.1. Low-Dose Simulation

Due to the ethical issues and potential health risk associated with scanning the same subject twice using different radiation doses, we follow the experimental setting in [6] to simulate low-dose CTP data at 15 mAs by adding correlated Gaussian noise with standard deviation of $\sigma = 25$ [31]. Note that this low-dose simulation technique is a widely method for CT algorithm evaluation in the medical field [32, 33]. The deconvolution methods are evaluated on the simulated low-dose CTP data. The quality of the CBF maps of all methods are evaluated by comparing with the reference maps (high-dose CTP) using peak signal-to-noise ratio (PSNR). While PSNR may not be the best evaluation metric for the clinical dataset, it is an objective quantitative reflection of the fidelity between the perfusion maps of the low-dose and the high-dose CTP data.

6.2. Clinical Dataset

Our method is evaluated on a clinical dataset of 10 subjects admitted to the NewYork-Presbyterian Hospital/Weill Cornell Medical College with mean age (range) of 53 (42-63) years and four of them had brain deficits due to aneurysmal subarachnoid hemorrhage (aSAH) or ischemic stroke, and the rest were normal. CTP images were collected with a standard protocol using GE Lightspeed Pro-16 scanners (General Electric Medical Systems, Milwaukee, WI) with cine 4i scanning mode and 60 second acquisition at 1 rotation per second, 0.5 sec per sample, using 80 kVp and 190 mA. Four 5-mm-thick sections with pixel spacing of 0.43 mm between centers of columns and rows were assessed at the level of the third ventricle and the basal ganglia, yielding a spatio-temporal tensor of $512 \times 512 \times 4 \times 118$ where there are 4 slices and 119 temporal samples. Approximately 45 mL of nonionic iodinated contrast was administered intravenously at 5 mL/s using a power injector with a 5 second delay.

6.3. Competing Methods

To evaluate the efficacy of the proposed TENDER model in perfusion map quantification of low-dose CTP, we evaluate the performance on multiple subjects using a number of state-of-art quantification methods, including:

1. sSVD [34]: Standard Singular Value Decomposition is the foundation of non-parametric estimation methods for perfusion imaging, which applies singular value decomposition (SVD) to the matrix remove the noise in small singular values.
2. bSVD [35]: Block-circulant SVD decomposes the block-circulant matrix of arterial input function to overcome the error caused by bolus delay. It has been one of the most effective methods for CTP quantification.
3. Tikhonov [36]: Tikhonov regularization approach uses a smooth weighting function for the singular values to minimize the impact of noise.
4. TTV [6]: Tensor total variation exploits the spatio-temporal correlation of the tissue in a small neighborhood region, which has already shown excellent performance in low-dose CTP quantification.

7. Experimental Results

In this section, we demonstrate the visual and quantitative results of CBF map estimation using our proposed TENDER model and the compared methods.

7.1. The Normal Case

We first assess the compared methods on a normal subject (Fig. 4). For each subject, from left to right it shows the reference map, the low-dose maps of standard singular value decomposition (sSVD), block-circulant singular value decomposition (bSVD), Tikhonov, TTV, and our proposed TENDER. Obviously, the singular value-based methods (sSVD, bSVD, Tikhonov) tend to over-estimate the CBF map and lead to large variance in the homogeneous areas in the white matter and cerebrospinal fluid (CSF). In comparison, spatio-temporal model-based methods (TTV and TENDER) maintains the fidelity of the CBF map, with clear-cut boundary between the vessel and the non-vessel tissues (shown in the ROI images), without over-estimation of the perfusion parameters. However, TTV still over-estimates the CBF value, and the vessels in the close-up view are dilated due to the local smoothing using the tensor total variation regularization. TENDER overcomes both issues. The quantitative accuracy of the perfusion maps improves significantly, and more noticeably, the small vessels in the brain are precisely preserved without dilation or rupture, as we can observe in the local TTV results.

7.2. The Ischemic Case

We then apply the compared methods on a subject with right frontoparietal craniotomy and ischemia in the right anterior cerebral artery (RACA) and right middle cerebral artery (RMCA) territories (Fig. 5). The entire brain image and the close-up views demonstrate significant improvement in the overall accuracy and preservation of the delicate anatomical structures using the non-local TTV method for both the deficit and the normal subjects. sSVD tends to severely over-estimate CBF, while SVD-based methods also over-estimate perfusion parameters. TTV performs better than the SVD-based methods in preserving

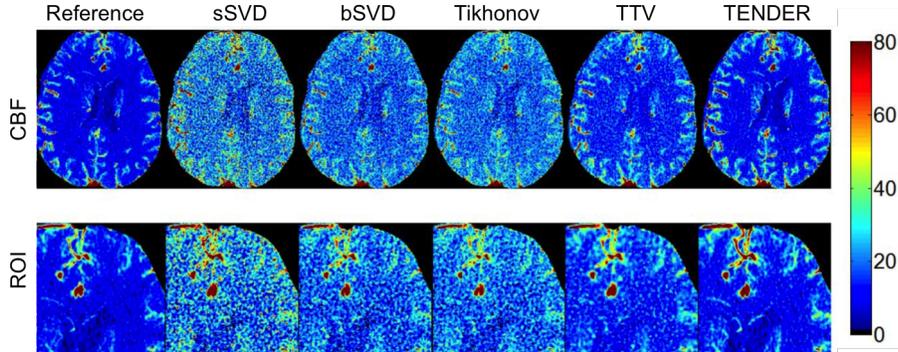


Figure 4: Results from a normal subject. From left to right shows the reference map (high-dose CTP), the low-dose maps of standard singular value decomposition (sSVD), block-circulant singular value decomposition (bSVD), Tikhonov, TTV, and our proposed TENDER. The first row is the whole brain CBF map and the second row is the closeup view.

the quantitative accuracy and the contrast resolution between different tissue classes, but still with severe noise and artifacts compared to the reference. Due to the non-local search to fully utilize the far-reaching neighborhood information, TENDER model is not confined to the 4-connected neighborhood for the spatio-temporal regularization and achieves the best performance in quantifying the CBF map and preserving image quality in the ischemic subject.

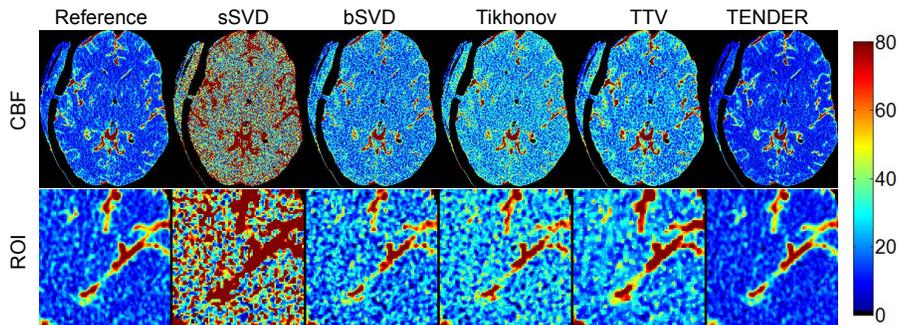


Figure 5: Results from a subject with right frontoparietal craniotomy and ischemia in the right anterior cerebral artery (RACA) and right middle cerebral artery (RMCA) territories. From left to right shows the reference map (high-dose CTP), the low-dose maps of standard singular value decomposition (sSVD), block-circulant singular value decomposition (bSVD), Tikhonov, TTV, and our proposed TENDER. The first row is the whole brain CBF map and the second row is the closeup view of a selected region.

7.3. Quantitative Evaluation

Finally we demonstrate the quantitative comparison in terms of PSNR for all the 10 subjects, as shown in Table 1. As there are both normal and abnormal subjects with aSAH or ischemic stroke, the large variation of the subject population help to demonstrate the robustness of the proposed TENDER model. In every case, TENDER model outperforms the compared methods, and overall TENDER gains over 6 dB in PSNR compared to SVD-based methods, and 3 dB over TTV. Box-plot revealing the central tendency of the compared methods is shown in Fig. 6. Our proposed method significantly outperforms all other comparison methods ($p < 0.05$) using one-tail Student’s t test.

Subject	sSVD	bSVD	Tikh	TTV	TENDER
1	10.83	19.66	19.32	24.04	26.55
2	8.75	17.14	17.67	21.25	22.38
3	6.65	15.85	15.74	19.91	21.89
4	2.16	19.25	19.24	15.55	22.17
5	15.26	27.21	26.66	22.96	37.70
6	10.59	18.46	18.79	23.31	27.22
7	9.81	18.32	19.69	24.32	25.45
8	9.45	19.09	18.90	23.09	24.72
9	13.69	22.57	22.79	27.54	28.81
10	10.99	19.29	19.61	20.04	21.12
Average	9.82	19.68	19.84	22.20	25.79

Table 1: Peak signal-to-noise ratio (PSNR) in dB of cerebral blood flow compared with the ground truth at high radiation dose of 10 subjects using the baseline methods and the proposed TENDER model. The best performance in terms of average PSNR is highlighted with bold font.

7.4. Running Time

Using the accelerated optimization algorithm, efficient search in the non-local spatio-temporal neighborhood and the parallel computing technique, our proposed TENDER model is nearly 20 times faster than the non-local regularization method without acceleration.

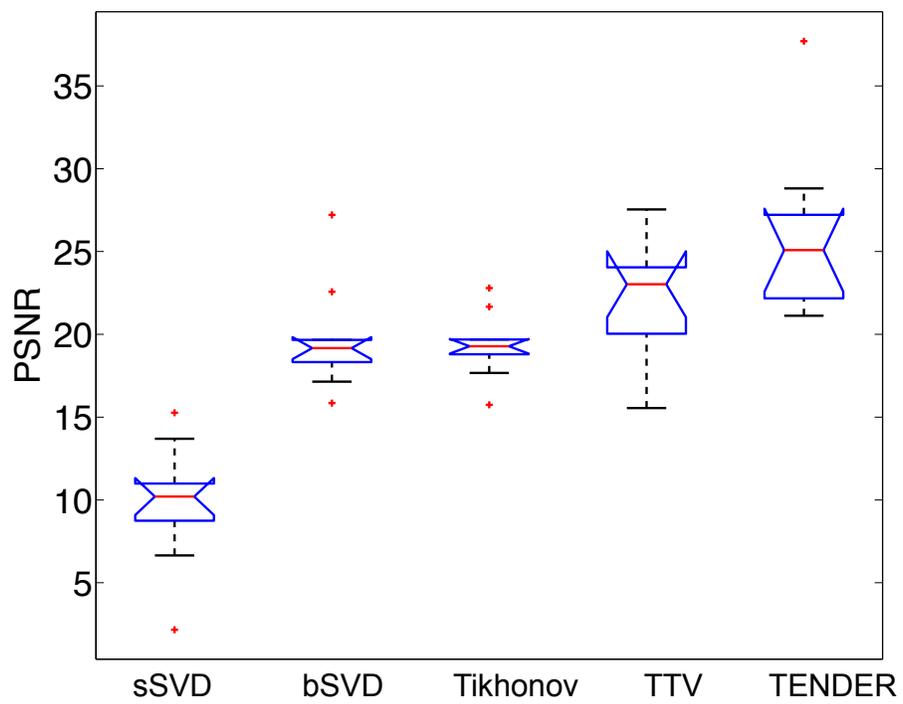


Figure 6: Boxplot of PSNR and SSIM for the 10 clinical subjects. The proposed TENDER method significantly outperforms all other comparison methods ($p < 0.05$).

8. Discussion

The history of perfusion imaging dates back to 1980 when Leon Axel proposed the original framework and principles of CT perfusion [37]. Perfusion imaging started with neuroimaging of a pre-selected region of the brain using a dynamic sequential scanning after the injection of contrast agent as it travels through the capillary bed of the brain [38, 39]. Recently, whole brain imaging [40] and perfusion imaging of other parts of the body such as heart [41], kidney [42], liver [43], pancreas [44] and prostate [45] have played a central role in vascular disease diagnosis. CT perfusion imaging has led to numerous quantification methods to model the perfusion flow and estimate the dynamic perfusion parameters such as cerebral blood flow (CBF), cerebral blood volume (CBV), mean transit time (MTT) and time-to-peak (TTP) [46, 46]. The basic model builds on indicator-dilution theory [47], and the first models quantify the hemodynamic parameters by evaluating the contrast concentration curves of each voxel individually [34, 48]. We call this type of model “voxel independent model”. This classic model has been widely adopted in clinical practice to estimate the perfusion parameters for vascular diseases such as stroke and subarachnoid hemorrhage [49, 50].

However the independent assumption in this classic model entirely ignores the fact that the “brain is connect”, in both spatial domain and the temporal passage. The neglect of this spatio-temporal connection may not be an enormous issue at sufficient radiation dose when the signal-to-noise ratio is high enough for accurate computation of the perfusion parameters, but can render the quantification accuracy a severe problem when the radiation dose is reduced for safer medical imaging [32, 51].

While a number of approaches have been attempted to improve the perfusion map estimation in low-dose or ultra-low dose CTP using spatial filtering or noise removal in the reconstructed serial CT images [8, 52, 11, 9], the oscillatory nature and unstable deconvolution CTP perfusion map estimation remain the bottleneck for accurate quantification. Spatio-temporal regulariza-

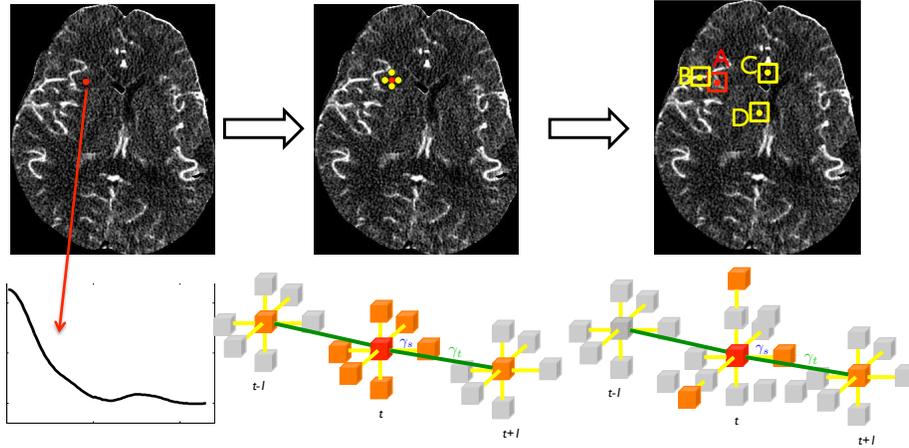


Figure 7: Illustration of the evolution of deconvolution algorithms for CTP, from independent voxel to local neighborhood, and to non-local regions. In the left independent model, every voxel is computed without consideration of spatial and temporal context. In the middle local TTV model, only the local neighborhoods (6 connected in 3D space and 2 connected in temporal domain) are used for robust deconvolution. In the right non-local TENDER model, non-local search of the broader neighborhood regions enable more robust estimation of the perfusion parameters.

tion has been developed to stabilize the deconvolution process, via patch-based dictionaries learned from high-dose perfusion maps [20, 19], tissue-specific [53], spatio-temporal deconvolution [13, 54, 6, 18]. This is an important step towards leveraging the contextual information of the connected brain tissue, yet the limitation of these methods is that the connection is confined to local neighborhood, usually the directly connected voxels, which confines the contextual information in limited region. Since the voxels dimensions in a typical CTP spatio-temporal data are much smaller compared to the tissue anatomy, the variation of blood perfusion should have regional effects rather than single or directly-connected voxel effects. Therefore, within the extended voxel neighborhood the perfusion parameters are expected to be of low variation or even constant. Based on these observations, the non-local TENDER model leverages the extended neighborhood of the voxel to find similar spatio-temporal patterns for regularization in low-dose data. The evolution of the deconvolution algorithms described above is illustrated in Fig. 7.

Furthermore, while the non-local search and similarity computation leads

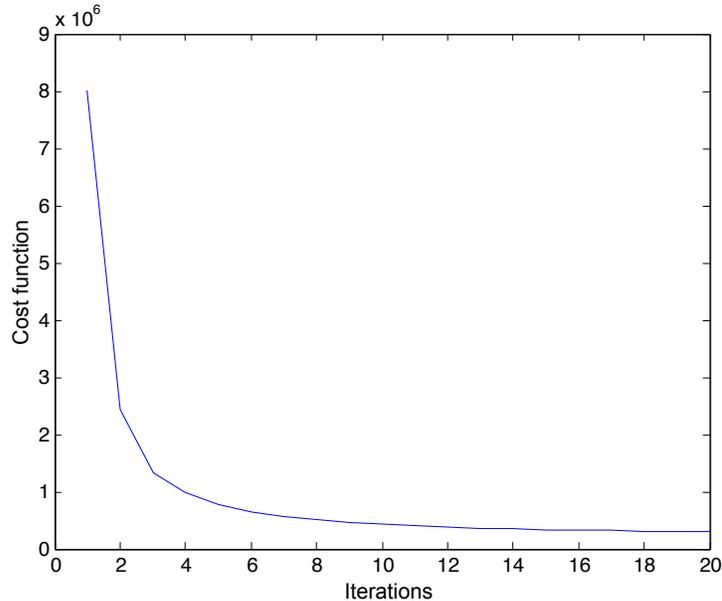


Figure 8: Convergence curve of the cost function for TENDER algorithm.

to higher computational cost, our efficient optimization and fast nearest neighbor search schema demonstrates fast convergence and reasonable computational time. Specifically, the algorithm converges within 10 iterations (Fig. 8) after our accelerated optimization. Since the algorithm is implemented in MATLAB platform and run on a single PC desktop, grid or cluster computing and implementation in C++ are expected to speed up the experiments.

9. Conclusion

Reducing the radiation exposure in ionizing radiation based imaging without compromising the image quality remains a critical challenge in medical imaging, especially for perfusion imaging with repetitive high-dose scans. In this paper, we proposed an efficient TENDER model to reduce the radiation dose in CT perfusion via non-local tensor total variation regularization. The non-local similarities of the same tissue classes in the brain structure are leveraged to stabilize the spatio-temporal residue functions. The overall quantitative accuracy is sig-

nificantly improved with the delicate anatomical structures such as vessels are well preserved to assist clinical diagnosis. Fast optimization and implementation schemes are presented to reduce the time complexity and computational cost. Extensive evaluations with comparison to the state-of-art algorithms, including sSVD, bSVD, Tikhonov and TTV, demonstrate the superior performance of the non-local TTV method in low-dose deconvolution and perfusion parameter estimation.

10. Acknowledgements

Research reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health under Award Number UL1TR000457, NSF IIS-1423056, CMMI-1434401, CNS-1405985. The content is solely the responsibility of the authors and does not necessarily represent the official views of the National Institutes of Health.

References

- [1] M. Wintermark, M. Lev, Fda investigates the safety of brain perfusion CT, *American Journal of Neuroradiology* 31 (1) (2010) 2–3. [2](#)
- [2] Y. Imanishi, A. Fukui, H. Niimi, D. Itoh, K. Nozaki, S. Nakaji, K. Ishizuka, H. Tabata, Y. Furuya, M. Uzura, et al., Radiation-induced temporary hair loss as a radiation damage only occurring in patients who had the combination of MDCT and DSA, *European radiology* 15 (1) (2005) 41–46. [2](#)
- [3] I. Shuryak, R. K. Sachs, D. J. Brenner, Cancer risks after radiation exposure in middle age, *Journal of the National Cancer Institute* 102 (21) (2010) 1628–1636. [2](#)
- [4] A. B. de González, K. P. Kim, J. M. Samet, Radiation-induced cancer risk from annual computed tomography for patients with cystic fibrosis, *American journal of respiratory and critical care medicine* 176 (10) (2007) 970–973. [2](#)

- [5] M. Cohnen, H.-J. Wittsack, S. Assadi, K. Muskalla, A. Ringelstein, L. Poll, A. Saleh, U. Müdder, Radiation exposure of patients in comprehensive computed tomography of the head in acute stroke, *American Journal of Neuroradiology* 27 (8) (2006) 1741–1745. [2](#)
- [6] R. Fang, S. Zhang, T. Chen, P. Sanelli, Robust low-dose CT perfusion deconvolution via tensor total-variation regularization, *Medical Imaging, IEEE Transactions on* 34 (7) (2015) 1533–1548. [doi:10.1109/TMI.2015.2405015](#). [3](#), [4](#), [6](#), [7](#), [11](#), [13](#), [14](#), [20](#)
- [7] R. Fang, P. C. Sanelli, S. Zhang, T. Chen, Tensor total-variation regularized deconvolution for efficient low-dose CT perfusion, in: *Medical Image Computing and Computer-Assisted Intervention–MICCAI 2014*, Springer, 2014, pp. 154–161. [3](#), [4](#), [11](#)
- [8] N. Saito, K. Kudo, T. Sasaki, M. Uesugi, K. Koshino, M. Miyamoto, S. Suzuki, Realization of reliable cerebral-blood-flow maps from low-dose CT perfusion images by statistical noise reduction using nonlinear diffusion filtering, *Radiological physics and technology* 1 (1) (2008) 62–74. [3](#), [19](#)
- [9] A. M. Mendrik, E.-j. Vonken, B. van Ginneken, H. W. de Jong, A. Riordan, T. van Seeters, E. J. Smit, M. A. Viergever, M. Prokop, Tips bilateral noise reduction in 4d CT perfusion scans produces high-quality cerebral blood flow maps, *Physics in Medicine and Biology* 56 (13) (2011) 3857. [3](#), [19](#)
- [10] Z. Tian, X. Jia, K. Yuan, T. Pan, S. B. Jiang, Low-dose CT reconstruction via edge-preserving total variation regularization, *Physics in medicine and biology* 56 (18) (2011) 5949. [3](#)
- [11] J. Ma, J. Huang, Q. Feng, H. Zhang, H. Lu, Z. Liang, W. Chen, Low-dose computed tomography image restoration using previous normal-dose scan, *Medical physics* 38 (2011) 5713. [3](#), [19](#)
- [12] M. Supanich, Y. Tao, B. Nett, K. Pulfer, J. Hsieh, P. Turski, C. Mistretta, H. Rowley, G.-H. Chen, Radiation dose reduction in time-resolved CT an-

- giography using highly constrained back projection reconstruction, *Physics in medicine and biology* 54 (14) (2009) 4575. [3](#)
- [13] L. He, B. Orten, S. Do, W. C. Karl, A. Kambadakone, D. V. Sahani, H. Pien, A spatio-temporal deconvolution method to improve perfusion CT quantification, *Medical Imaging, IEEE Transactions on* 29 (5) (2010) 1182–1191. [3](#), [20](#)
- [14] Y. Yu, S. Zhang, K. Li, D. Metaxas, L. Axel, Deformable models with sparsity constraints for cardiac motion analysis, *Medical image analysis* 18 (6) (2014) 927–937. [4](#)
- [15] S. Zhang, Y. Zhan, D. N. Metaxas, Deformable segmentation via sparse representation and dictionary learning, *Medical Image Analysis* 16 (7) (2012) 1385–1396. [4](#)
- [16] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. Metaxas, X. Zhou, Deformable segmentation via sparse shape representation, in: G. Fichtinger, A. Martel, T. Peters (Eds.), *Medical Image Computing and Computer-Assisted Intervention*, Vol. 6892 of *Lecture Notes in Computer Science*, Springer Berlin Heidelberg, 2011, pp. 451–458. [4](#)
- [17] S. Zhang, Y. Zhan, M. Dewan, J. Huang, D. N. Metaxas, X. S. Zhou, Towards robust and effective shape modeling: Sparse shape composition, *Medical image analysis* 16 (1) (2012) 265–277. [4](#)
- [18] R. Fang, J. Huang, W.-M. Luh, A spatio-temporal low-rank total variation approach for denoising arterial spin labeling MRI data, in: *Biomedical Imaging (ISBI), 2015 IEEE 12th International Symposium on*, IEEE, 2015, pp. 498–502. [4](#), [20](#)
- [19] R. Fang, K. Karlsson, T. Chen, P. C. Sanelli, Improving low-dose blood–brain barrier permeability quantification using sparse high-dose induced prior for patlak model, *Medical image analysis* 18 (6) (2014) 866–880. [4](#), [20](#)

- [20] R. Fang, T. Chen, P. C. Sanelli, Towards robust deconvolution of low-dose perfusion CT: Sparse perfusion deconvolution using online dictionary learning, *Medical image analysis* 17 (4) (2013) 417–428. [4](#), [20](#)
- [21] R. Fang, T. Chen, P. Sanelli, Sparsity-based deconvolution of low-dose brain perfusion CT in subarachnoid hemorrhage patients, in: *Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on*, 2012, pp. 872–875. [doi:10.1109/ISBI.2012.6235687](https://doi.org/10.1109/ISBI.2012.6235687). [4](#)
- [22] R. Fang, M. Ni, J. Huang, Q. Li, T. Li, Efficient 4D non-local tensor total-variation for low-dose CT perfusion deconvolution, in: *18th Annual International Conference on Medical Image Computing and Computer Assisted Intervention, Workshop on Medical Computer Vision: Algorithms for Big Data*, Elsevier, 2015. [4](#)
- [23] R. Fang, M. Ni, J. Huang, Q. Li, T. Li, Robust low-dose CT perfusion deconvolution via non-local tensor total variation, in: *Biomedical Engineering Society Annual Meeting, BMES*, 2015. [4](#)
- [24] A. Sawatzky, (Nonlocal) Total Variation in Medical Imaging, PhD Thesis. [7](#)
- [25] X. Zhang, M. Burger, X. Bresson, S. Osher, Bregmanized nonlocal regularization for deconvolution and sparse reconstruction, *SIAM Journal on Imaging Sciences* 3 (3) (2010) 253–276. [doi:10.1137/090746379](https://doi.org/10.1137/090746379). [7](#), [9](#)
- [26] M. Mignotte, A non-local regularization strategy for image deconvolution, *Pattern recognition letters* 29 (16) (2008) 2206–2212. [7](#), [9](#)
- [27] A. Elmoataz, O. Lezoray, S. Boughleux, Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing, *Image Processing, IEEE Transactions on* 17 (7) (2008) 1047–1060. [7](#)
- [28] X. Jia, Y. Lou, B. Dong, Z. Tian, S. Jiang, 4D computed tomography reconstruction from few-projection data via temporal non-local regularization,

Medical Image Computing and Computer-Assisted Intervention–MICCAI 2010 (2010) 143–150. [7](#)

- [29] J. Huang, F. Yang, Compressed magnetic resonance imaging based on wavelet sparsity and nonlocal total variation, in: Biomedical Imaging (ISBI), 2012 9th IEEE International Symposium on, IEEE, 2012, pp. 968–971. [7](#)
- [30] A. Beck, M. Teboulle, Fast gradient-based algorithms for constrained total variation image denoising and deblurring problems, *Image Processing, IEEE Transactions on* 18 (11) (2009) 2419–2434. [11](#)
- [31] A. J. Britten, M. Crotty, H. Kiremidjian, A. Grundy, E. J. Adam, The addition of computer simulated noise to investigate radiation dose and image quality in images with spatial correlation of statistical noise: an example application to X-ray CT of the brain, *The British Journal of Radiology* 77 (916) (2004) 323–328. [doi:10.1259/bjr/78576048](https://doi.org/10.1259/bjr/78576048). [13](#)
- [32] K. Juluru, J. Shih, A. Raj, J. Comunale, H. Delaney, E. Greenberg, C. Hermann, Y. Liu, A. Hoelscher, N. Al-Khori, et al., Effects of increased image noise on image quality and quantitative interpretation in brain CT perfusion, *American Journal of Neuroradiology* 34 (8) (2013) 1506–1512. [13](#), [19](#)
- [33] D. P. Frush, C. C. Slack, C. L. Hollingsworth, G. S. Bisset, L. F. Donnelly, J. Hsieh, T. Lavin-Wensell, J. R. Mayo, Computer-simulated radiation dose reduction for abdominal multidetector CT of pediatric patients, *American Journal of Roentgenology* 179 (5) (2002) 1107–1113. [13](#)
- [34] L. Østergaard, R. M. Weisskoff, D. A. Chesler, C. Gyldensted, B. R. Rosen, High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis, *Magnetic resonance in medicine* 36 (5) (1996) 715–725. [14](#), [19](#)

- [35] H.-J. Wittsack, A. M. Wohlschläger, E. Ritzl, R. Kleiser, M. Cohnen, R. Seitz, U. Mödder, CT-perfusion imaging of the human brain: advanced deconvolution analysis using circulant singular value decomposition, *Computerized Medical Imaging and Graphics* 32 (1) (2008) 67–77. [14](#)
- [36] A. Fieselmann, M. Kowarschik, A. Ganguly, J. Hornegger, R. Fahrig, Deconvolution-based CT and MR brain perfusion measurement: theoretical model revisited and practical implementation details, *Journal of Biomedical Imaging* 2011 (2011) 14. [14](#)
- [37] L. Axel, Cerebral blood flow determination by rapid-sequence computed tomography: theoretical analysis., *Radiology* 137 (3) (1980) 679–686. [19](#)
- [38] K. Miles, M. Hayball, A. Dixon, Colour perfusion imaging: a new application of computed tomography, *The Lancet* 337 (8742) (1991) 643–645. [19](#)
- [39] M. Koenig, E. Klotz, B. Luka, D. J. Venderink, J. F. Spittler, L. Heuser, Perfusion CT of the brain: diagnostic approach for early detection of ischemic stroke, *Radiology* 209 (1) (1998) 85–93. [19](#)
- [40] E. Siebert, G. Bohner, M. Dewey, F. Masuhr, K. Hoffmann, J. Mews, F. Engelken, H. Bauknecht, S. Diekmann, R. Klingebiel, 320-slice CT neuroimaging: initial clinical experience and image quality evaluation, *The British Journal of Radiology* 82 (979) (2009) 561–570, PMID: 19221186. [doi:10.1259/bjr/27721218](https://doi.org/10.1259/bjr/27721218). [19](#)
- [41] H. Machida, I. Tanaka, R. Fukui, Y. Shen, T. Ishikawa, E. Tate, E. Ueno, Current and novel imaging techniques in coronary CT, *RadioGraphics* 35 (4) (2015) 991–1010, PMID: 26046942. [doi:10.1148/rg.2015140181](https://doi.org/10.1148/rg.2015140181). [19](#)
- [42] X.-R. Cai, Q.-C. Zhou, J. Yu, Y.-Z. Feng, Z.-H. Xian, W.-C. Yang, X.-K. Mo, Assessment of renal function in patients with unilateral ureteral obstruction using whole-organ perfusion imaging with 320-detector row com-

- puted tomography, PLoS ONE 10 (4) (2015) e0122454. [doi:10.1371/journal.pone.0122454](https://doi.org/10.1371/journal.pone.0122454). 19
- [43] P. V. Pandharipande, G. A. Krinsky, H. Rusinek, V. S. Lee, Perfusion imaging of the liver: current challenges and future goals 1, *Radiology* 234 (3) (2005) 661–673. 19
- [44] L. Delrue, P. Blanckaert, D. Mertens, E. Cesmeli, W. P. Ceelen, P. Duyck, Assessment of tumor vascularization in pancreatic adenocarcinoma using 128-slice perfusion computed tomography imaging, *Journal of computer assisted tomography* 35 (4) (2011) 434–438. 19
- [45] N. Cullu, M. Kantarci, H. Ogul, S. Adanur, A. Okur, E. Koc, B. Pirimoglu, L. Karaca, Y. Kizrak, O. Polat, Feasibility study of CT perfusion imaging for prostate carcinoma, *European radiology* 24 (9) (2014) 2236–2244. 19
- [46] A. Konstas, G. Goldmakher, T.-Y. Lee, M. Lev, Theoretic basis and technical implementations of CT perfusion in acute ischemic stroke, part 1: theoretic basis, *American Journal of Neuroradiology* 30 (4) (2009) 662–668. 19
- [47] P. Meier, K. L. Zierler, On the theory of the indicator-dilution method for measurement of blood flow and volume, *Journal of applied physiology* 6 (12) (1954) 731–744. 19
- [48] L. Østergaard, A. G. Sorensen, K. K. Kwong, R. M. Weisskoff, C. Gyldensted, B. R. Rosen, High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. part ii: Experimental comparison and preliminary results, *Magnetic Resonance in Medicine* 36 (5) (1996) 726–736. 19
- [49] T. E. Mayer, G. F. Hamann, J. Baranczyk, B. Rosengarten, E. Klotz, M. Wiesmann, U. Missler, G. Schulte-Altdorneburg, H. J. Brueckmann, Dynamic CT perfusion imaging of acute stroke, *American journal of neuroradiology* 21 (8) (2000) 1441–1449. 19

- [50] M. Wintermark, N. Ko, W. Smith, S. Liu, R. Higashida, W. Dillon, Vasospasm after subarachnoid hemorrhage: utility of perfusion CT and CT angiography on diagnosis and management, *American journal of neuroradiology* 27 (1) (2006) 26–34. [19](#)
- [51] A. Murphy, A. So, T.-Y. Lee, S. Symons, R. Jakubovic, L. Zhang, R. I. Aviv, Low dose CT perfusion in acute ischemic stroke, *Neuroradiology* 56 (12) (2014) 1055–1062. [19](#)
- [52] E. Tong, M. Wintermark, CTA-enhanced perfusion CT: an original method to perform ultra-low-dose CTA-enhanced perfusion CT, *Neuroradiology* 56 (11) (2014) 955–964. [19](#)
- [53] R. Fang, H. Jiang, J. Huang, Tissue-specific sparse deconvolution for brain CT perfusion, *Computerized Medical Imaging and Graphics* 46 (2015) 64–72. [20](#)
- [54] C. Frindel, M. C. Robini, D. Rousseau, A 3-d spatio-temporal deconvolution approach for mr perfusion in the brain, *Medical image analysis* 18 (1) (2014) 144–160. [20](#)