2021 43rd Annual International Conference of the
IEEE Engineering in Medicine & Biology Society (EMBC)
Oct 31 - Nov 4, 2021. Virtual Conference

Forward Model and Deep Learning Based Iterative Deconvolution for Robust Dynamic CT Perfusion

Viswanath P. Sudarshan*, Pavan Kumar Reddy*, Jayavardhana Gubbi, Balamuralidhar Purushothaman
Embedded Devices and Intelligent Systems, TCS Research, Bangalore, India

***Abstract*— Perfusion maps obtained from low-dose computed tomography (CT) data suffer from poor signal to noise ratio. To enhance the quality of the perfusion maps, several works rely on denoising the low-dose CT (LD-CT) images followed by conventional regularized deconvolution. Recent works employ deep neural networks (DNN) for learning a direct mapping between the noisy and the clean perfusion maps ignoring the convolution-based forward model. DNN-based methods are not robust to practical variations in the data that are seen in real-world applications such as stroke. In this work, we propose an iterative framework that combines the perfusion forward model with a DNN-based regularizer to obtain perfusion maps directly from the LD-CT dynamic data. To improve the robustness of the DNN, we leverage the anatomical information from the contrast-enhanced LD-CT images to learn the mapping between low-dose and standard-dose perfusion maps. Through empirical experiments, we show that our model is robust both qualitatively and quantitatively to practical perturbations in the data.**

I. INTRODUCTION

Computed tomography perfusion (CTP) imaging enables fast quantitative diagnosis of the hemodynamic parameters and has recently been widely established for clinical assessment of acute ischemic stroke. Typical CTP scans involve a dynamic scanning session with a bolus injection of the contrast agent during which several consecutive CT scans are performed. The amount of radiation exposure from a dynamic session of CTP imaging is a cause of concern and compromises patient safety[14]. With the clinics adopting the *as low as reasonably achievable* (ALARA) [16] principle, low-dose CTP imaging can lead to better diagnosis and prognosis of several cerebrovascular pathologies. However, lowering the dose results in degradation of the signal-to-noise ratio (SNR) in the reconstructed CT images and consequently, the estimated perfusion maps. Recently, there has been an increased emphasis on reducing the ionizing radiations without compromising the image quality [15].

Early influential works focused on estimating the perfusion maps directly from the 4-D spatiotemporal contrast-enhanced CT data using model-based approaches. For standard-dose CT imaging, typically, truncated singular value decomposition (TSVD) method is used to obtain the quantitative cerebral blood flow (CBF) maps [10]. Prior works on estimation of improved CTP images from LD-CT data can be categorized as: (i) regularized deconvolution to obtain CBF maps from noisy LD-CT data; (ii) denoising/enhancing the the LD-CT images to obtain an estimate of the SD-CT images followed by application of TSVD algorithm to obtain improved CBF maps; and (iii) post-deconvolution denoising/enhancement of the CBF maps obtained from LD-CT images to estimate perfusion maps at standard dose.

Deep-learning methods have shown substantially improved performance for estimating the quantitative perfusion maps compared to other learning-based methods [8], [19]. While such works focus on improving the quantitative accuracy of the estimated perfusion maps from low-dose CT images or CT perfusion maps, they do not model the convolution-based forward process [5] and hence, are susceptible to degradation in performance when presented with input data that deviates from the data used for training the DNN. Robustness to practical perturbations is critical in applications like medical imaging where the perturbations in the data might be related to the scanner or patient's physiology.

Deep neural network (DNN) based methods that map from a given source image (input) to target image (reference) often evaluate the performance of the models on test data that are quite similar to training data. Several works have studied the degradation in performance of DNN models when presented with testing data that deviates from training data [7], [11]. To improve the robustness of the DNN-based models to solve *ill-posed* inverse problems, hybrid models are proposed that leverage (i) the forward model for data consistency (measurement space) and (ii) DNN for regularization (image space). Such hybrid models have been studied for image reconstruction of modalities like MRI *e.g.*, [20], PET [6], and have been shown to be robust to different acquisition strategies, *e.g.*, variations in k-space trajectory for MRI [1]. This paper proposes a novel framework that leverages the convolution model for perfusion maps generation from low-dose dynamic CT data. Specifically, we propose an iterative framework that leverages the denoising capability of the DNNs trained at a specific noise level. Additionally, we observe that our reconstructions improve over the state-of-the-art typically in a few (≤ 10) iterations.

II. METHOD

A. Forward Model for Perfusion Maps Generation

Let $A(t)$ represent the arterial input function (AIF), determined from the set of N dynamic LD-CT data, $\{X^n_{LD}\}^N_{n=1}$. Let $\{X^n_{SD}\}^N_{n=1}$ represent the corresponding set of dynamic standard-dose CT (SD-CT) images. Let $R_v(t)$ denote the unknown tissue residue function within a small region v. $R_v(t)$ denotes the amount of residual contrast in the vascular structures. Then, the time-concentration curve $C_v(t)$ is given by the convolution-based forward model [5] $C_v(t) =$

*Equal contribution

978-1-7281-1178-0/21/$31.00 ©2021 IEEE

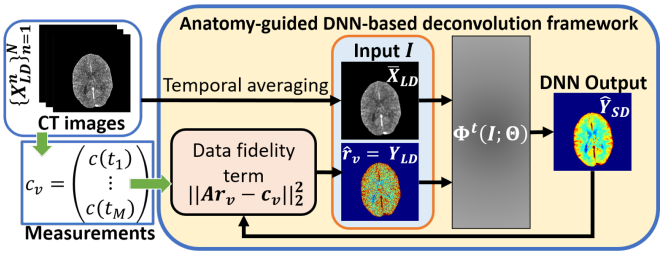

Fig. 1. Proposed iterative framework for generating perfusion maps. The LD-CT images are averaged along the temporal dimension (to remove noise) and utilized as a multi-channel input for the DNN to leverage the structural information. $\Phi^t(\cdot;\cdot)$ denotes the DNN used at the t-th iteration.

$\int_0^t A(\tau)R_v(t-\tau)\,d\tau$. In the discrete setting, let the column vectors c_v and r_v, each containing M elements obtained at equally spaced intervals, denote the measured contrast and the unknown residue function, respectively. Thus, the forward model becomes $c_v = \mathbf{A}r_v$, where matrix $\mathbf{A}$ is block-circulant with columns representing $A(t)$ as described in [18]. In the iterative setting, estimate r_v is obtained based on the convolution model as data likelihood and a suitable regularization function $\mathcal{D}(\cdot)$ yielding the objective function

$$\widehat{r_v} = \arg\min_{r_v} \|\mathbf{A}r_v - c_v\|^2 + \lambda\mathcal{D}(r_v), \tag{1}$$

where $\lambda > 0$ is the regularization parameter. The CBF maps are obtained by suitably scaling the image r_v as in [4].

We use the alternate direction method of multipliers (ADMM) algorithm with iteration-dependent DNNs acting as regularizers to solve the above-mentioned optimization problem. To solve Eq. 1 the ADMM update equations can be written as follows [2]:

$$r_v^{k+1} = \arg\min_{r_v} \|\mathbf{A}r_v - c_v\|^2 + \frac{\rho}{2}\|r_v - (s^k - \bar{u}^k)\|^2, \tag{2}$$

$$s^{k+1} = \arg\min_{s} \lambda\mathcal{D}(s) + \frac{\rho}{2}\|s - ({r_v}^{k+1} + \bar{u}^k)\|^2, \tag{3}$$

$$\bar{u}^{k+1} = \bar{u}^k + (r_v^{k+1} - s^{k+1}), \tag{4}$$

where, $\bar{u}^k = u^k/\rho$ is the scaled Lagrangian multiplier, s is the variable splitting term used to convert the unconstrained problem in Eq. 1 into a constrained problem. Eq. 3 corresponds to proximal map of the regularizer function $\mathcal{D}(\cdot)$ and acts as a denoiser. Several works have shown that replacing this update equation with a well-known denoiser improves the image quality during reconstruction [2]. Since the noise affecting the estimated CBF maps is different (due to the underlying forward process) than typical noise distributions (*e.g.*, Gaussian or Poisson) that affect natural images, we train a set of CNN denoisers to update s in Eq. 3. Further, to overcome reducing levels of noise across the iterative process, at the k-th iteration, we use the DNN $\mathcal{D}_\sigma^k$ which is trained using CBF maps generated from CT images affected by noise level characterized by σ. The proposed framework is illustrated in Figure 1.

Low-dose simulation. We follow the low-dose CT data-generation process as explained in several prior works [3], [8], [4]. Let I_{SD} and I_{LD} represent the tube currents employed for standard-dose and low-dose CT data acquisitions, respectively. Similarly, let σ_{SD} and σ_{LD} represent the noise standard deviation corresponding to I_{SD} and I_{LD}, respectively. Then, to generate the low-dose CT data, $\{X_{LD}^n\}_{n=1}^N$ given the standard-dose images $\{X_{SD}^n\}_{n=1}^N$, the noise standard deviation σ to be added is given by the relation $\sigma_{LD} = \sigma_{SD} + \sigma$. We add Gaussian noise with suitable standard deviation to the SD-CT images as in [3] to generate the set of LD-CT images.

B. DNN Model

Let Y_{LD} represent the CBF map obtained using TSVD-based deconvolution with $\{X_{LD}^n\}_{n=1}^N$. Similarly, let Y_{SD} represent the CBF map obtained from $\{X_{SD}^n\}$. Let $\bar{X}_{LD}$ denote the mean of the spatio-temporal data $\{X_{LD}^n\}_{n=1}^N$ along the temporal direction. We use a DNN $\Phi(I;\Theta)$, for the regression task, where Θ represents the network parameters and I represents the pair of input images $X := \{Y_{LD}, \bar{X}_{LD}\}$. The regressor $\Phi(I;\Theta)$ learns a mapping between the input pair of images to the output image Y_{SD}.

We use a DNN that follows a residual learning approach and the architecture is similar to that of the DNN used in [8] that is modified to accept a multi-channel (two-channel) input. We employ a patch-based training, with patch-size 40 $\times$ 40, such that the noise characteristics are captured within a patch, similar to [8]. The input and output patch sizes are the same (by padding in the convolutional layers), and since the network is a fully convolutional network, the DNN is capable of handling inputs of arbitrary size during inference. The DNN consists of 17 layers with convolutional layers, rectified linear units as activation function, and batch-normalization for stabilizing the training process. At each iteration of the proposed algorithm, as the noise in the reconstructed image (r_v) keeps reducing, we employ a separate denoiser $\mathcal{D}_{\sigma^q}^i := \Phi^q(I^q;\Theta_q,\sigma^q)$ where I^q denotes the input at the q-th iteration and σ^q the corresponding noise standard deviation.

III. Data and Experiments

We use the contrast-enhanced dynamic CT images available as part of the ischemic stroke lesion (ISLES) challenge 2018 [12]. The dataset consists of 94 volumes of dynamic CT images each containing varying number of axial slices (two to eight) and time-frames (40 to 50). We consider the provided CT data as standard-dose data, *i.e.*, SD-CT images. We simulated the LD-CT images by adding noise to the set of SD-CT images as mentioned earlier. To obtain the perfusion maps at SD and LD, first, we remove (replace with zero) the bone regions with a threshold of 120 HU. Secondly, we computed the AIF and VOF as detailed in [13], [9]. We add spectral noise (to generate LD-CT) within the brain region, and obtain corresponding perfusion (CBF) maps at SD and LD using the TSVD-based deconvolution algorithm [18]. The SSIM between the set of SD-CBF and the set of LD-CBF images averaged over the entire dataset was around 0.80. We denote the above-mentioned lower dosage level as LD-1. Assuming that the CT images for perfusion imaging are obtained with a tube current of 150 mAs (typically clinics employ $\geq$ 150 mAs [14]), the resultant noise corresponds to a tube current of around 15 mAs. Thus, the data at LD-1

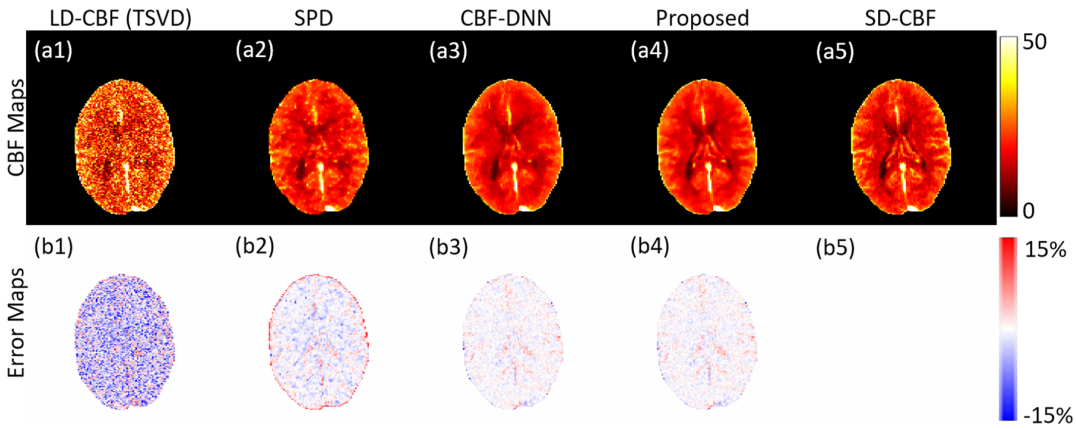

Fig. 2. Validation on test data LD-1. Estimated CBF maps: (a1) TSVD: truncated SVD method applied to LD-CT images; (a2) SPD: sparse dictionary prior method; (a3) CBF-DNN: residual-learning-based DNN with LD-CBF as input; (a4) Proposed: forward-model + DNN-regularizer-based deconvolution of LD-CT images. (a5) SD-CBF: reference CBF map at standard dose.

corresponds to a dose reduction factor (DRF) of at least 10×. We randomly select 59 subjects for training, 5 for validation, and 30 for testing.

A. Additional data with SNR degradation

To validate the robustness of the model to data that deviates from the training set, we generated two additional datasets with further reduced dose compared to LD-1. We focus on out-of-distribution data arising from two different acquisition methods, yielding LD-CT images at lower SNR levels LD-2 and LD-3, generated as follows.

Further reduced tube current (LD-2). As in LD-1, we adjusted the added noise to the SD-CT images such that the average SSIM between the SD-CBF and the LD-CBF images averaged over the test set was around 0.78 (LD-2). This test set corresponds to a tube current of 10 mAs which represents a DRF of at least 15×.

Reduced number of frames (LD-3). In this case, we downsample the LD-CT time-frames (select alternate frames) at LD-1 by a factor of two resulting in a further 2× reduction in dose. Thus, effective DRF is approximately 20×. The average SSIM between the SD-CBF and the LD-CBF images averaged over the test set was around 0.77 (LD-3).

IV. Results and Discussion

We compare our proposed method with three other methods spanning across regularized model-based and standalone DNN methods: (i) deconvolution using TSVD in which regularization is achieved by eliminating the singular values (of matrix $\mathbf{A}$) with low-values (based on a heuristic threshold) [10]; (ii) sparse deconvolution using a dictionary prior, with the dictionary trained on SD-CBF images, as described in [3] (SPD); (iii) DNN-based CBF map denoiser (CBF-DNN), with a similar architecture as described in [8]. We employ the same train-validation-test split for all the methods and also tune the underlying hyperparameters for all the methods using the validation set such that we obtained the least relative root mean squared error (RRMSE) defined between two images A and B as $RRMSE(A,B) = \|A - B\|_F / \|A\|_F$, where $\|\cdot\|_F$ denotes the Frobenius norm. For quantitative comparison across methods on all the three test sets, we use the structural similarity index (SSIM) [17].

Figure 2 shows qualitative results of the predicted CBF maps using LD-CT data corresponding to LD-1, with LD-CBF (Figure 2 (a1)) as input and SD-CBF (Figure 2 (a5)) as reference. The SPD method (Figure 2 (a2)) removes substantial amount of noise compared to the TSVD method (Figure 2 (a2)). While the CBF-DNN method (Figure 2 (a3)) improves substantially over the SPD method by further removing noise, the predicted image from the proposed method (Figure 2 (a4)), shows a superior recovery of structure and contrast (especially in the subcortical regions) in addition to further reduced noise. Most importantly, our method that includes anatomical information from the averaged low-dose CT images, yields images with the least residual magnitudes (Figure 2 (b4)).

Figure 3 shows predicted CBF maps obtained using data corresponding to LD-2 and LD-3, as detailed in Section III-A. Both CBF-DNN and our proposed methods are evaluated using models trained on noise-level corresponding to LD-1. In general, the intensities in the CBF maps are lower compared to corresponding CBF maps in Figure 2 due to further reduced dose arising from reduced tube current (LD-2: rows (a) and (b)) and reduced number of CT frames acquired (LD-3: rows (c) and (d)). Similar to LD-1 (Figure 2), both SPD (Figure 3 (a2)) and CBF-DNN (Figure 3 (a3)) remove substantial amount of noise compared to the TSVD method (Figure 3 (a1)). On the other hand, our proposed method (Figure 3 (a4)) shows a superior recovery of structure and contrast (especially in the subcortical regions) in addition to removal of noise, as highlighted in the zoomed regions (Figure 3 (b4)). In the case of data from reduced frames, the CBF-DNN method (Figure 3 (a3)) improves substantially over the SPD method by producing a smoother perfusion image. However, compared to the proposed method (Figure 3 (c4)), the CBF-DNN fails to capture certain anatomical structures *e.g.*, the frontal horn of the lateral ventricles (see Figures 3 (d3)–(d5)).

Figure 4 shows the quantitative performance of all the methods for the three different inputs discussed in this paper:

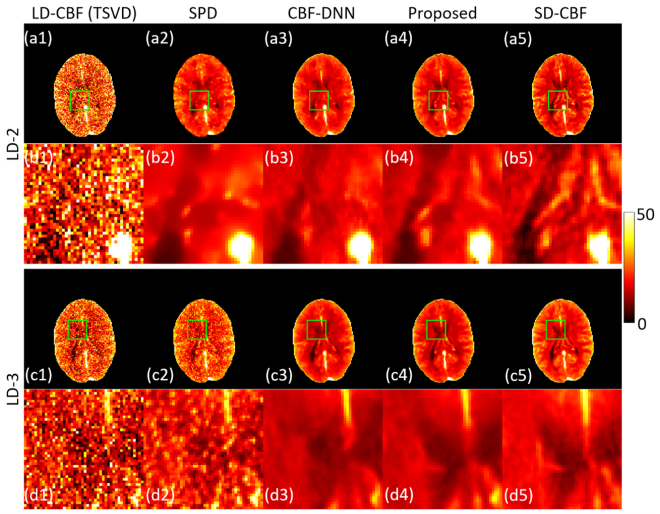

Fig. 3. Validation on two different datasets with SNR degradation, LD-2 (rows (a) and (b) and LD-3 (rows (c) and (d)). Rows (b) and (d) show the zoomed regions of interest indicated by the green boxes in the panels in rows (a) and (c).

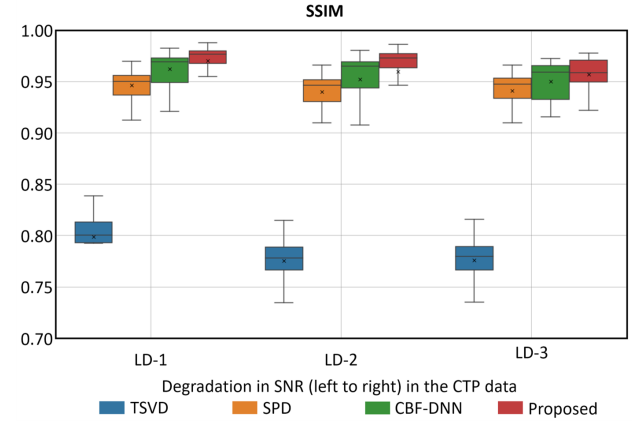

Fig. 4. Quantitative evaluation for all methods on three different datasets with decreasing SNR: LD-1, LD-2, and LD-3. The box plots represent the SSIM evaluated for all the slices for each test subject.

LD-1, LD-2, and LD-3, achieving a DRF upto 20×. The box plots are obtained on all of the axial slices of the entire test set. Although all the methods show a drop in performance as we move from LD-1 to LD-3, our method offers the least degradation across all noise-levels. For all the cases, our method improves over CBF-DNN, and both these methods improve significantly over SPD. The improvement of DNN-based methods over SPD is consistent with [8].

V. Conclusion

This paper proposes a iterative framework that combines the model-based deconvolution with a DNN-based regularizer generating CBF perfusion maps from LD-CT data. Our novel contributions include employing (ii) a DNN-based regularizer in an iterative framework along with the forward model and (ii) the anatomical information from the CTP images as a multi-channel input to the DNN within the iterative framework. Through empirical analysis, we observed that our method leverages the LD-CT image patches along with the LD-CBF images to provide improved stability (during training) and robustness (during inference). The proposed framework provides improved CBF maps for severe reduction in the tube current (upto 20×) both qualitatively and quantitatively compared to the state of the art. Our method demonstrated improved robustness to practical perturbations in the data from two additional acquisition scenarios, *i.e.*, reduced tube current and reduced number of CT frames. Future works call for a detailed clinical validation, ablation studies, and evaluation on other out of distribution data.

References

[1] HK Aggarwal, MP Mani, and M Jacob. MoDL: Model-based deep learning architecture for inverse problems. *IEEE Trans Med Imag.*, 38(2), 2018.

[2] SH Chan, X Wang, and O Elgendy. Plug-and-play admm for image restoration: Fixed-point convergence and applications. *IEEE Trans Comput Imag.*, 3(1), 2016.

[3] R Fang, T Chen, and P Sanelli. Towards robust deconvolution of low-dose perfusion CT: Sparse perfusion deconvolution using online dictionary learning. *Med Imag Anal.*, 17(4):417–428, 2013.

[4] R Fang, S Zhang, T Chen, and P Sanelli. Robust low-dose CT perfusion deconvolution via tensor total-variation regularization. *IEEE Trans Med Imag.*, 34(7), 2015.

[5] A Fieselmann, M Kowarschik, A Ganguly, J Hornegger, and R Fahrig. Deconvolution-based CT and MR brain perfusion measurement: theoretical model revisited and practical implementation details. *Int J Biomed Imag.*, 2011.

[6] K Gong, J Guan, K Kim, X Zhang, J Yang, Y Seo, G El Fakhri, J Qi, and Q Li. Iterative PET image reconstruction using convolutional neural network representation. *IEEE Trans Med Imag.*, 38(3), 2018.

[7] YC Hsu, Y Shen, H Jin, and Z Kira. Generalized ODIN: Detecting out-of-distribution image without learning from out-of-distribution data. In *CVPR*, 2020.

[8] VS Kadimesetty, S Gutta, S Ganapathy, and PK Yalavarthy. Convolutional neural network-based robust denoising of low-dose computed tomography perfusion maps. *IEEE Trans Radiat Plasma Med Sci.*, 3(2), 2018.

[9] Y Kao, M Teng, Y Kao, Y Chen, C Wu, W Chen, F Chiu, and F Chang. Automatic measurements of arterial input and venous output functions on cerebral computed tomography perfusion images: A preliminary study. *Comp Biol Med.*, 51, 2014.

[10] TS Koh, XY Wu, LH Cheong, and CCT Lim. Assessment of perfusion by dynamic contrast-enhanced imaging using a deconvolution approach based on regression and singular value decomposition. *IEEE Trans Med Imag.*, 23(12), 2004.

[11] K Lee, K Lee, H Lee, and J Shin. A simple unified framework for detecting out-of-distribution samples and adversarial attacks. In *Adv Neural Info Proc Sys.*, pages 7167–7177, 2018.

[12] O Maier, B Menze, J von der Gablentz, et al. ISLES 2015-A public evaluation benchmark for ischemic stroke lesion segmentation from multispectral MRI. *Med Imag Anal.*, 35, 2017.

[13] Kim Mouridsen, Søren Christensen, Louise Gyldensted, and Leif Østergaard. Automatic selection of arterial input function using cluster analysis. *Mag Reson Med.*, 55(3), 2006.

[14] JM Niesten, IC van der Schaaf, AJ Riordan, H, et al. Radiation dose reduction in cerebral ct perfusion imaging using iterative reconstruction. *Eur Radiol.*, 24(2), 2014.

[15] A Othman, S Afat, M Brockmann, O Nikoubashman, C Brockmann, K Nikolaou, and M Wiesmann. Radiation dose reduction in perfusion CT imaging of the brain: A review of the literature. *J Neuroradiology*, 43(1), 2016.

[16] SD Voss, GH Reaman, SC Kaste, and TL Slovis. The ALARA concept in pediatric oncology. *Pediat Radiol.*, 39, 2009.

[17] Z Wang, A Bovik, H Sheikh, and E Simoncelli. Image quality assessment: from error visibility to structural similarity. *IEEE Trans Image Proc.*, 13(4), 2004.

[18] HJ Wittsack, A Wohlschläger, EK Ritzl, R Kleiser, M Cohnen, R Seitz, and U Mödder. CT-perfusion imaging of the human brain: advanced deconvolution analysis using circulant singular value decomposition. *Comp Med Imag Graphics*, 32(1), 2008.

[19] Dufan Wu, Hui Ren, and Quanzheng Li. Self-supervised dynamic CT perfusion image denoising with deep neural networks. *IEEE Trans Radiat Plasma Med Sci*, 2020.

[20] Y Yang, J Sun, H Li, and Z Xu. ADMM-CSNet: A deep learning approach for image compressive sensing. *IEEE Trans Patt Anal Mach Intell.*, 42(3), 2018.