Convolutional Neural Network-Based Robust Denoising of Low-Dose Computed Tomography Perfusion Maps

Venkata S. Kadimesetty, Sreedevi Gutta, Sriram Ganapathy, and Phaneendra K. Yalavarthy, *Senior Member, IEEE*

***Abstract*—The low-dose computed tomography (CT) perfusion data has low signal-to-noise ratio resulting in derived perfusion maps being noisy. These low-quality maps typically requires a denoising step to improve their utility in real-time. The existing methods, including state-of-the-art online sparse perfusion deconvolution (SPD), largely relies on the convolutional model that may not be applicable in all cases of brain perfusion. In this paper, a denoising convolutional neural network (DCNN) was proposed that relies only on computed perfusion maps for performing the denoising step. The network was trained with a large number of low-dose digital brain phantom perfusion maps to provide an approximation to the corresponding high-dose perfusion maps. The batch normalization coupled with residual learning makes the trained model invariant to the dynamic range of the input low-dose perfusion maps. The denoising of the raw-data using the convolutional neural network was also attempted here and shown to have limited applicability in the low-dose CT perfusion cases. The digital perfusion phantom as well as *in-vivo* results indicate that the proposed DCNN applied in the derived map domain provides superior improvement compared to the online SPD with an added advantage of being computationally efficient.**

***Index Terms*—Cerebral blood flow (CBF), computed tomography (CT) perfusion, convolutional neural network (CNN), denoising, low-dose.**

I. Introduction

COMPUTED tomography (CT) perfusion is a well established clinical diagnosis method for cerebral perfusion evaluation, which provides spatial maps of cerebral blood flow (CBF), cerebral blood volume (CBV), and mean transit time (MTT) [1]. CT perfusion is relatively a fast technique to treat patients with symptoms ranging from acute stroke to subarachnoid hemorrhage (SAH) [2]. CT perfusion involves dynamic imaging of contrast bolus after injection in the cerebral vasculature. The perfusion maps computed from the acquired temporal/dynamic data are post-processed to result in the quantitative perfusion maps (CBF, CBV, and MTT). For these perfusion maps to be quantitatively accurate, higher radiation dosage is required compared to standard CT of the head. Hence, CT perfusion scans are known to be high exposure scans that can make patients suffer with several biological effects, such as skin burn and increased risk of cancer [3]. Often the risk of CT perfusion scans for the patients tends to be undermined as stroke is the third-leading cause of death in the world. As CT perfusion is being widely used for diagnosis as well as prognosis of several other cerebrovascular diseases other than stroke, the high dosage risk cannot be ignored. The important challenge to be solved for making CT perfusion as the main stream cerebrovascular imaging method is to provide quantitatively accurate CT perfusion maps (such as CBF) using low dose CT perfusion data.

Several mathematical models were proposed earlier to obtain quantitative perfusion maps from the temporal data. The model involves deconvolution that is based on the central volume principle, with the assumption that the tissue time-enhancement curves is a convolution of arterial input function (AIF) and residue function, which describes the contrast remaining in the tissue at any instant of time. The most widely used deconvolution method is circulant truncated singular value decomposition (TSVD) [4]–[6]. This method was proven to provide meaningful results in the normal (high-dose) CT perfusion cases, the challenge remains for it to have utility in the low-dose cases. It is also known that the low-dose CT perfusion imaging data tends to be more noisy, thus resulting in highly distorted perfusion maps.

Several works had proposed to denoise the dynamic CT data, thereby improving the quantitative accuracy of CBF maps. These methods include nonlocal means, bilateral filtering, wavelet-based methods, and iterative reconstruction methods [7]. A data-driven approach that learns the dictionary from the high-dose perfusion maps termed as sparse perfusion deconvolution (SPD) was proposed earlier to enhance the accuracy of the estimated perfusion maps using low-dose perfusion data and provide better diagnostic accuracy [8]. This method computes the dictionary globally for performing denoising. As SPD uses the dynamic data in

Manuscript received March 14, 2018; revised May 30, 2018 and July 6, 2018; accepted July 17, 2018. Date of publication July 27, 2018; date of current version March 1, 2019. This work was supported in part by the Department of Biotechnology (DBT) Innovative Young Biotechnologist Award under Grant BT/07/IYBA/2013-13, and in part by DBT Bioengineering under Grant BT/PR7994/MED/32/284/2013. *(Venkata S. Kadimesetty and Sreedevi Gutta contributed equally to this work.) (Corresponding author: Phaneendra K. Yalavarthy.)*

V. S. Kadimesetty, S. Gutta, and P. K. Yalavarthy are with the Department of Computational and Data Sciences, Indian Institute of Science, Bengaluru 560 012, India (e-mail: narayana.kvs@gmail.com; gutta.sreedevi@gmail.com; yalavarthy@iisc.ac.in).

S. Ganapathy is with the Department of Electrical Engineering, Indian Institute of Science, Bengaluru 560 012, India (e-mail: sriramg@iisc.ac.in).

Color versions of one or more of the figures in this paper are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/TRPMS.2018.2860788

the deconvolution/denoising approach, the complexity of the problem become intractable with increase in the number of time points. With the advances in the CT detector technology, the temporal resolution is increasing in multifold with the introduction of multirow detectors, making CT perfusion number of time points to be close to 200 in less than a minute [9]. Also, the recently established tissue-specific deconvolution process has a requirement of obtaining accurate segmentation of tissue types, such as vessel, gray matter, and white matter [10]. Obtaining these tissue types accurately in real-time is highly challenging especially in the low-dose CT perfusion case and often requires manual-/semi-automatic methods to be deployed, making their utility limited in terms of applying them for low-dose CT perfusion denoising.

As these methods are data-driven, it is natural to apply deep learning methods for performing the denoising of low-dose CT perfusion maps. Deep learning has become the state-of-the-art denoising method in computer vision and there has been increased interest in applying similar strategies to medical imaging problems [11], [12]. Deep networks are popular due to their high computational efficiency leading to efficient algorithms and often provide superior performance compared to sparse coding as well as dictionary learning-based methods [13]–[15]. Most methods in the literature were geared toward improving the reconstructed low-dose CT image quality using deep neural networks, but did not directly address the problem of denoising of perfusion maps derived from the noisy low-dose CT perfusion data.

A. Prior Work

Denoising forms an important step in many post-processing applications. The aim of denoising methods is to provide an estimate of ground truth from the noisy observation. A simple approach for performing denoising is to project the data to transform domain, where noise can easily be separated assuming that the noise was additive white Gaussian [16]. Another way is estimating the statistics of the image directly [17]. Sparse coding techniques have received increased attention in recent years, which reconstruct images from a sparse linear combination of a learned dictionary [17]. A method that combines sparse coding and deep neural networks was presented to denoise the images with white Gaussian noise [18]. Later, a patch-based denoising algorithm with a simple network was proposed [19]. It was shown that this network was competitive to the state-of-the-art block-matching and 3-D (BM3D) filtering algorithm [19]. In another work, a convolutional neural network (CNN) was proposed for denoising of natural images and was shown to be superior compared to wavelet and Markov random field methods [20]. Existing denoising algorithms assume a specific noise model with a certain noise level for training. Recently, a denoising CNN (DCNN) was proposed that can handle denoising with unknown noise level [21].

For denoising of low-dose CT images, a Gaussian mixture Markov random field model was utilized as a prior in the image denoising problem [22]. This method introduced a statistical framework that is data-driven to arrive at denoised low-dose CT images [22]. Later, a deep CNN was proposed to remove noise in the low-dose CT images and was shown to have a great potential on structure preservation and artifact reduction [23]. A similar approach for improving low-dose CT images was developed by applying CNNs to a directional wavelet transform of the CT image [24]. These methods have shown quantitative improvements in reducing noise of low-dose CT images. Inspired by these works, a residual encoder–decoder CNN with a deconvolution network and shortcut connections was proposed to noise suppression and lesion detection in low-dose CT images [25]. Later, a deep convolutional network based on skip connections and residual learning was proposed for low-dose CT images and was shown improvement over total-variation-based reconstruction methods [26]. Recently, a CNN with an adversarial CNN was proposed to estimate high-dose CT images from low-dose CT images [27]. It was shown that the appearance of the generated images was similar to high-dose CT images due to utilization of adversarial network [27]. A residual CNN trained with noise being the output of the network was developed for post-processing of low-dose images and was shown to be effective in reducing artifacts while preserving tissue details [28]. In another work, an iterative image reconstruction algorithm for low-dose CT with priors learned from an artificial neural network was proposed [29]. Again, none of this prior work directly addressed the problem of low-dose CT perfusion maps denoising. Also, it is important to note that the low-dose CT perfusion scanning protocol is different compared to the typical low-dose CT imaging protocol, as the perfusion scanning involves collecting temporal dynamics and often the high-dose CT perfusion data it self have poor statistics [30].

The low-dose CT perfusion raw data needs to be deconvolved to provide perfusion maps, improvement in raw data is typically attributed to the improvement in the quality of perfusion maps. The application of denoising to the derived perfusion maps has not been popular as the inverse noise introduced by the deconvolution process is often difficult to estimate. The recent progress in developing generalized denoising models that are purely data driven based on deep CNNs provide an excellent opportunity to perform the denoising as the underlying noise model for the derived maps is unknown and these CNNs provides much needed generalizability across various imaging scenarios (for example varying X-ray tube currents).

The recent work of Zhang *et al.* [21] attempted to provide generalized denoising framework that is beyond standard additive Gaussian noise model. Inspired by this paper, a deep CNN was introduced in here for denoising the low-dose CBF (perfusion) maps and the network was trained with low-dose CBF map as the input and the high-dose CBF map as the output. At the same time, utilization of deep CNNs for performing raw data denoising was also explored to see whether denoising applied in either/both data and map space will provide a better alternative for improving the low-dose CBF maps. In this paper, the residual learning and batch normalization (BN) as proposed by Zhang *et al.* [21] was utilized to speed up the training and increase the denoising performance. The

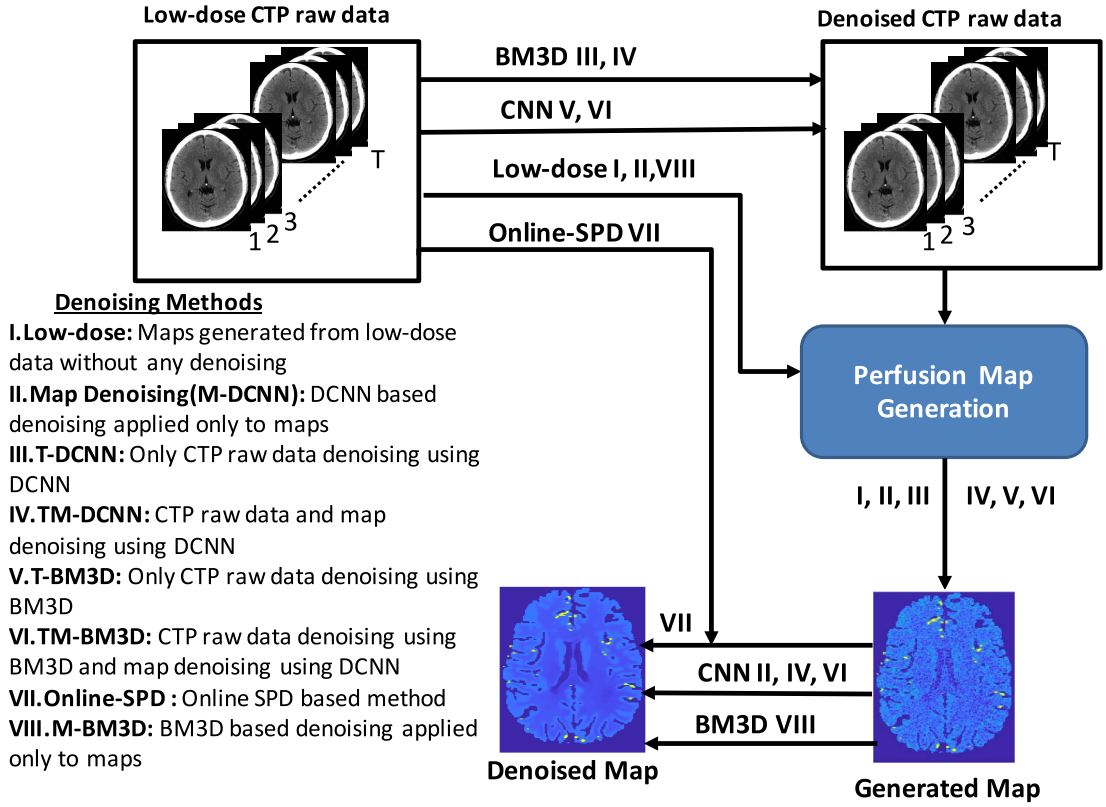

Fig. 1. Block diagram showing the denoising methods utilized in this paper. The roman numbers next to the arrows indicate the methods that utilize these steps. The full description of these methods is given in Section II-G. The training setup for CNN is given in Fig. 2.

standard denoising approaches, such as BM3D [19], [31], [32], was also applied in the raw data domain to highlight the generalizability of CNN-based methods. The proposed methods (summarized in Fig. 1) were compared to the state-of-the-art online dictionary-based SPD method both in digital brain perfusion phantom as well as *in-vivo* cases. Note that the high-dose CT perfusion map (CBF) was used as a reference to compare the output of the proposed deep learning-based denoising as well as online dictionary-based SPD methods.

II. Methodology

As stated earlier, the standard method for denoising of CBF map for comparison was online dictionary-based SPD method. Initially, the brain perfusion mathematical framework was presented, followed by a description of circulant TSVD-based deconvolution method. Later the online dictionary-based SPD method was described in detail. Finally, the proposed deep learning (CNN) denoising methods were presented.

A. Perfusion Parametric Model

The model to determine the amount of contrast in the region is given by a convolution defined as [4]

$$C_v(t) = \text{CBF} \int_0^t C_a(\tau) R(t-\tau) d\tau \tag{1}$$

where $C_v(t)$ is the tissue enhancement curve in the volume of interest (measured), CBF is the tissue blood flow (unknown), $C_a(t)$ is an AIF (typically determined using dynamic data), and $R(t)$ denotes the residue function of the tissue that quantifies the amount of contrast left in the vascular network at a certain time computed on a voxel by voxel basis (unknown). The value of one for R indicates that the entire contrast is within the vascular network and zero means the entire contrast has left.

The convolution operation represented in (1) can be discretized by measuring AIF and tissue enhancement curve at equally spaced points $t_1, t_2, \ldots t_T$, with a step size of Δt where T denotes the total number of time instants. The discretized equation can be represented as

$$C_v = \text{CBF}. * \Delta t. * \mathbf{C_a} * R \tag{2}$$

where C_v and R are column vectors containing concentrations and residues, respectively, at different time instances, and $\mathbf{C_a}$ represents a block-circulant matrix. Note that using circular deconvolution instead of linear deconvolution avoids the time shift problem and results in providing quantitatively accurate perfusion maps [33]. The deconvolution process involves estimation of (CBF and/or R) given C_v and $\mathbf{C_a}$, equivalently finding an inverse for $\mathbf{C_a}$. But $\mathbf{C_a}$ being ill-conditioned in nature, one can only determine pseudo-inverse, typically performed using TSVD [4], [5].

B. Circulant Truncated Singular Value Decomposition

Circulant TSVD is a popular method for performing the deconvolution in CT perfusion [34], [35]. The singular value decomposition (SVD) of the block circulant matrix $\mathbf{C_a}$ is

given as

$$\mathbf{C_a} = \mathbf{USV}^T \tag{3}$$

where **U** and **V** represents the left and right orthogonal matrices, and **S** is a diagonal matrix containing the singular values arranged in the descending order. Utilizing the SVD of the circulant matrix, (2) can be rewritten as

$$\text{CBF}.\Delta t.R = \mathbf{V}\Sigma\mathbf{U}^T.C_v \tag{4}$$

where Σ is a diagonal matrix having inverse of the diagonal values of **S**. The smaller singular values lead to large weighing coefficients resulting in inaccurate perfusion maps. TSVD regularizes the solution by eliminating the smaller singular values thereby removing the oscillatory terms from the solution generating more accurate maps. The truncation parameter plays an important role in regularizing the solution. In this paper, 30% of the maximum (first) singular value was used as a truncation parameter based on the experimental results presented in [36].

C. Online Sparse Perfusion Deconvolution

The perfusion maps computed from low-dose CT perfusion data are noisy, necessitating the need of a denoising step. Recent interest has been on sparse representations and dictionary learning methods [8]. Dictionary learning methods are typically data driven, where priors can be learned from maps to provide accurate results rather than using a prior that is based on simplifying assumptions (such as image being smooth). It is posed as an optimization problem with constrains. The least-squares minimization function based on (2) can be written as

$$\Gamma_{\text{LS}} = ||C_v - \mathbf{C_a}R||_2^2. \tag{5}$$

At $t = 0$, the $g = R(t = 0)$ becomes the CBF. The least-squares solution without regularization may not provide meaningful solutions due to the high level of noise in the low-dose CT perfusion data.

In SPD, a dictionary $\mathbf{D} \in \mathbf{R}^{P\times Q}$ was learned from the high-dose CBF maps, where P denotes the number of pixels in the input patch and Q indicates the number of patches in the dictionary [8], [10], [37]. This method assumes that the input patch is a sparse linear combination of the dictionary elements. The minimization function under this assumption becomes [8]

$$\Gamma = \mu_1||C_v - \mathbf{C_a}R||_2^2 + ||g - \mathbf{D}\alpha||_2^2 \ \ni \ ||\alpha||_0 \leq k \tag{6}$$

where μ_1 is the weight of the temporal term, $||.||_0$ indicates ℓ_0-norm that counts the number of nonzero entries, and $\alpha = [\alpha_1, \alpha_2, ..., \alpha_Q]$ are weights or coefficients for the linear combination of the dictionary elements (which are to be determined). Depending on the problem at hand, the constraint in (6) assures that the number of nonzero entries in α are less than k. Equation (6) is intractable due to the nonconvex property of ℓ_0-norm. The constraint in (6) can be relaxed to ℓ_1-norm, which makes it convex while still promoting sparsity [38]. Equation (6) in this case (unconstrained optimization) becomes [8]

$$\Gamma = \mu_1||C_v - \mathbf{C_a}R||_2^2 + ||g - \mathbf{D}\alpha||_2^2 + \mu_2||\alpha||_1 \tag{7}$$

where μ_2 represents the sparsity of α. The minimization function given in (7) can be solved iteratively using SPD method. The first step is to minimize (7) with respect to α with a fixed g, which can be written as [8]

$$\Gamma_1 = ||g - \mathbf{D}\alpha||_2^2 + \mu_2||\alpha||_1. \tag{8}$$

Equation (8) is minimized with respect to α using LARS-Lasso algorithm [39]. The second step is to find/update g with computed α leading to the minimization problem given as

$$\Gamma_2 = \mu_1||C_v - C_aR||_2^2 + ||g - \mathbf{D}\alpha||_2^2. \tag{9}$$

Equation (9) is minimized with respect to g. These minimizations (8), (9) are performed iteratively to compute α and g. Note that this minimization requires the initialization with randomly sampled patches of CBF map of high-dose CT perfusion and these patches are used to seek and update the dictionary **D** (which is initialized to over complete DCT dictionary). The parameter μ_1 balances the temporal model and dictionary and μ_2 controls the sparsity of the dictionary coefficients. In this paper, the parameters μ_1 and μ_2 were chosen heuristically for each case to obtain best results with lowest error norm between high-dose and low-dose CBF maps. The dictionary size of 256 was utilized to denoise a noisy image with patch size being 8×8 with an overlap of 7 pixels.

D. Block Matching Three-Dimensional Filtering

BM3D filtering is a nonlocal method to denoise images and to enhance the sparsity in the transform domain [31], [32]. It is a standard method of denoising in medical imaging and has been applied widely to both CT as well as magnetic resonance imaging data [31], [32]. The BM3D denoising algorithm consists of three steps. Initially, similar block images are stacked into 3-D data arrays and are decorrelated using the invertible 3-D transform. The spectra obtained in 3-D are filtered using hard thresholding. Lastly, the filtered spectra are inverted and placed back in their original position. The final denoised image is obtained from the weighted average of all clockwise estimates. The detailed description of this method can be found in [40]. In this paper, each volume of temporal CT perfusion data and maps were denoised using BM3D to compare with the CNN-based methods. The open-source code available at [41] was utilized here for implementing BM3D filtering. The parameters of the algorithm were chosen optimally based on the corresponding noise in the CTP data. For the map denoising, the parameters were chosen heuristically to result in best possible denoised estimate.

E. Proposed Map Denoising Convolutional Neural Network

Recently the CNNs-based denoising has become the main stream of denoising method especially in computer vision and allied areas. The success in using CNN for a wide range of tasks can be attributed to the availability of large amounts of training data and computational power with powerful graphics processing units (GPUs), which offer massive parallelization. In this paper, a DCNN was utilized to denoise the CBF maps obtained from low-dose perfusion data. The architecture of the network was similar to the network proposed for denoising

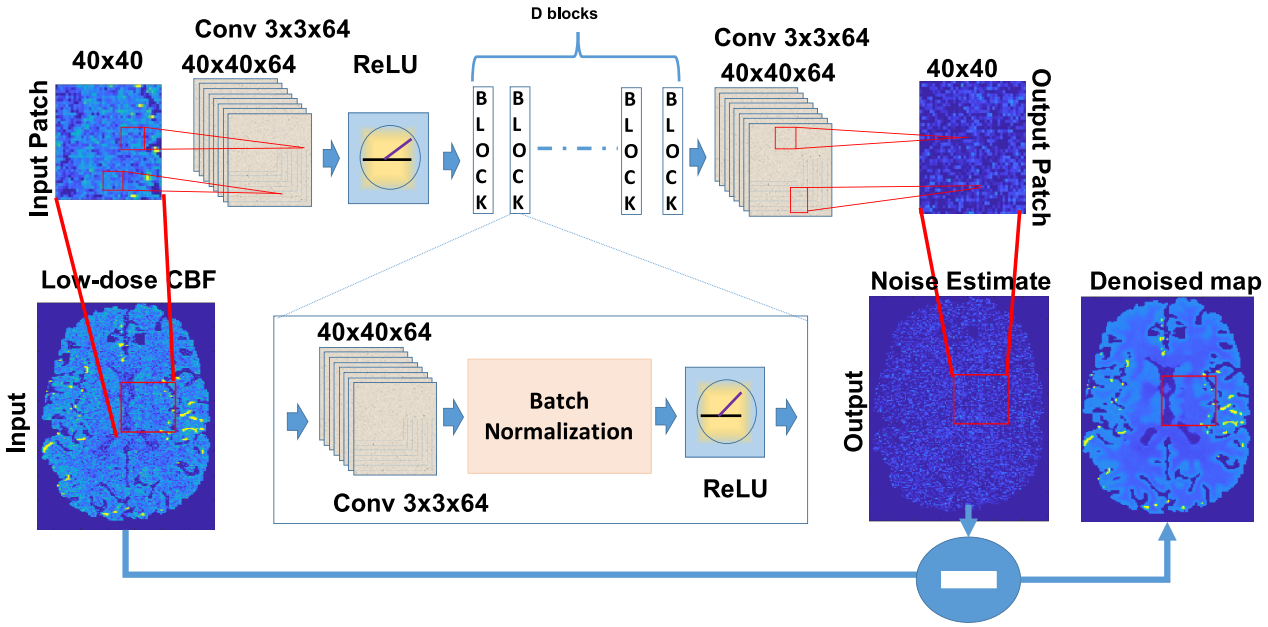

Fig. 2. Training setup of deep CNN architecture for denoising the low-dose CT perfusion map (methods-II, IV, and VI). Input to the network is a low-dose CT perfusion map (CBF) and output of the network is the estimation of the noise present in the map. The denoised CBF is obtained by subtracting the estimated noise from the low-dose CBF map. The basic block that is repeated 12 (indicated as D) times consisting of: a convolution layer consisting of 64 3×3 kernels, a BN layer, followed by an ReLU for nonlinearity. Testing is performed on the whole image without the requirement of patch extraction. Single precision was used in the network architecture and the maps are pesudo-colored for better representation. Note that the same architecture was also used for low-dose CT perfusion raw data denoising (utilized in methods-III and IV).

natural images corrupted with unknown Gaussian noise [21]. This network performs the training of denoising patch-wise. The depth of the network (number of layers) depends on the patch-size, which in turn depends on the noise level. Larger patch size should be considered for high noise levels to capture more context. A residual learning method combined with BN was utilized for the purpose of faster training and better denoising performance.

1) Convolutional Neural Network Architecture: The proposed CNN architecture was shown in Fig. 2. The input to the network is a low-dose noisy CBF map represented as $x \in \mathbf{R}^{p\times q}$, where p and q corresponds to the width and height of the input, respectively. The output of the network is a residual image ($s \in \mathbf{R}^{p\times q}$) determined from the noisy low-dose input. The denoised perfusion map is obtained by subtracting the residue from the noisy input, which can be expressed as $y = x - s$. The aim of the network is to determine the function F that maps noisy low-dose input to the predicted residue (based on high-dose CBF map), given as

$$s = F(x). \tag{10}$$

The loss function to determine the optimal weights for the network can be obtained by considering mean squared error, which can be written as

$$L = \frac{1}{2N}\sum_{i=1}^{N} ||F(x_i, \theta) - (x_i - y_i)||^2 \tag{11}$$

where the suffix i corresponds to the patch in the input image and N being the total number of training samples (image patches). Equation (11) should be minimized to obtain the trainable parameters θ.

The network shown in Fig. 2 includes three kinds of layers. The first one consists of convolution layer followed by ReLU nonlinearity (Conv + ReLU). Convolution is performed with 64 filters each of size 3×3 generating 64 feature maps. This layers takes an image patch of low-dose CBF map as the input. These patches are symmetrically zero padded before each stage of the network to result in same size. A nonlinear activation function was used after the convolution operation. The second layer is Conv+BN+ReLU. The parameters of this layer are similar to the first layer with the addition of BN between convolution layer and ReLU. The last layer is a simple convolution layer with filters of size $3\times3\times64$ to determine the residue from the noisy low-dose input. The convolved feature maps are summed to compute the output patch. Note that there are no pooling layers utilized in the network.

This proposed map DCNN (M-DCNN) has three noticeable features. One of them is combining convolution layer with ReLU nonlinearity separates the desired image characteristics from the noisy input. The other is the importance of integrating residual learning and BN. The combination of residual learning and BN speedup the training process and also improves the denoising performance [21]. Existing denoising algorithms are trained for a specific noise level, whereas the proposed M-DCNN network is generic and is capable to efficiently remove unknown noise (will also be shown later).

2) Residual Learning: Recent works have shown that using deep CNN improves the accuracy of the task at hand, including large-scale image recognition [42], [43]. The same deep CNN framework was utilized in the proposed method. One of the limitation of using deep network is vanishing gradient problem [44]. This problem can be solved using normalization layers, which helps to improve the

convergence [45]. Moreover, the training accuracy of deeper networks may start degrading and this effect is termed as degradation problem. It was shown that residual network helps to overcome the degradation problem providing better training accuracy and faster convergence [46], which was utilized in the proposed network.

3) Batch Normalization: Stochastic gradient descent (SGD) is a popular optimization algorithm to determine optimal weights of the network. It should be noted that training deeper networks is a slow process and requires an appropriate selection of learning rate and initialization scheme. Additionally, training models with saturating nonlinearities is challenging. This effect termed as covariant shift is due to variations in distributions of nonlinear inputs. BN layer helps to overcome these limitations due to internal covariant shift introduced in the previous layer by subtracting with mean and dividing with standard deviation [45]. There exists two learnable parameters γ and β to undo the effect of subtraction and division. The batch norm (bn) can be expressed as [45]

$$\text{bn} = \gamma \frac{z - \mu_t}{\sigma_t} + \beta \tag{12}$$

where μ_t and σ_t corresponds to mean and standard deviation over the batch, respectively, and z indicates the input to the BN layer. It works as a regularizer eliminating the need of dropout and is less sensitive to initialization, allowing to achieve higher learning rate. Note that integrating residual learning with BN provides faster training and better denoising performance.

4) Data Preparation: Training on large data boosts the performance of the deep network. Generating large amounts of data for medical imaging applications is always a challenging task. To generate training data, the digital brain perfusion phantom package that has the tools to create a realistic brain perfusion data was utilized [47]. The phantoms were used for reliable evaluation of different methods. The package contains tools that help us to choose reduced and severely reduced perfusion inside the segmented brain. With this toolbox, four phantoms of dimension $256 \times 256 \times 256$ were generated by considering various core and penumbra regions at each mAs. Data augmentation through image rotation was considered to generate large amounts of data. Note that data augmentation increases the efficiency of the task at hand [48]. A total of 453 120 training patches across all (mixed) mAs of low-dose and high-dose CBF map pairs were generated for training the proposed network. For a network with 12 blocks around 0.45 million parameters needed to be learned. Sample training images for various noise levels were shown in Fig. 3.

5) Implementation Details: The major steps involved in the proposed M-DCNN were summarized as method-II in Fig. 1. The convolutions performed in the network utilized 3×3 filters with the number of blocks (D) as shown in Fig. 2 varying from 9 to 15. The input low-dose (details of the data generation are given in the next section) and output residue have a dimension of 40×40 (patch size). For larger size images, denoising was performed in a patch-wise manner. Network was initialized with a robust method given in [49] that takes into account ReLU nonlinearities. This initialization enables to train deep networks efficiently. The proposed M-DCNN was trained using MatConvNet package [50]. An efficient mini-batch gradient descent (SGD) was employed to determine optimal weights for the network. Training was performed for 30 epochs on a batch-size of 256 with learning rate decaying exponentially from 1e-1 to 1e-4. The time taken for training the proposed network was about 36 h. All experiments were performed on a single Nvidia k-20 GPU. Note that the training time could be significantly reduced when multiple GPUs can be deployed or using more recent GPU architectures.

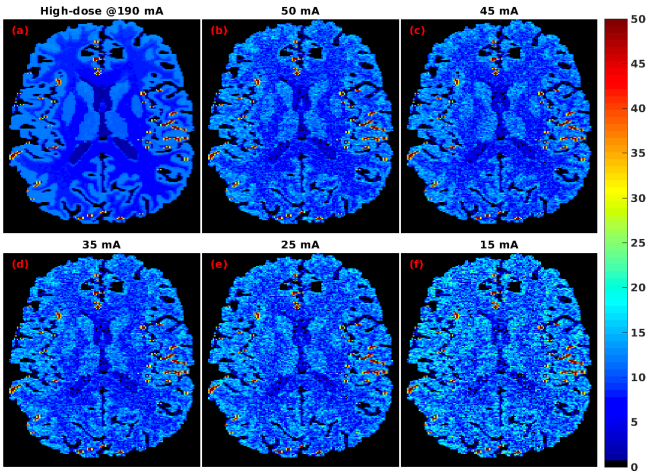

Fig. 3. Typical CBF maps that generated without any denoising using digital brain phantom data at different mAs using the methodology given in Section II-F. (a) CBF map obtained using high dose CT perfusion data acquired at X-ray tube current of 190 mA (expected output of the network, Fig. 2). This map was used as a reference to compute the figures of merit discussed in Section II-H. Low-dose CBF maps obtained using CT perfusion data (method-I in Fig. 1) with tube current of (b) 50 mA, (c) 45 mA, (d) 35 mA, (e) 25 mA, and (f) 15 mA.

F. Low-Dose Simulation

To simulate low-dose CT perfusion data, spatially correlated statistical noise was added to the reconstructed high-dose CT perfusion imaging data as described in [30] and [51]. In CT images, the tube current I (mA) is inversely proportional to noise standard deviation (σ_n) given as [8]

$$I = \frac{K^2}{\sigma_n^2}. \tag{13}$$

If σ_0 is the noise standard deviation corresponding to I_0 tube current, (13) can be rewritten as

$$\frac{I}{I_0} = \frac{\sigma_0^2}{\sigma_n^2}. \tag{14}$$

The standard deviation of the added noise can be expressed as

$$\sigma_n^2 = \sigma_0^2 + \sigma_{\text{noise}}^2. \tag{15}$$

The spectral properties of noise to be added were determined from the phantom data. As it was observed that the power spectrum of noise for different noise levels do not vary and thus the same spectral model was considered for all noise levels (similar to [30] and [51]). A convolution filter generated from a 11×11 window around the autocorrelation peak of the noise autocorrelation function was utilized for generating

colored noise. This convolution filter was applied on the white Gaussian noise and subsequently scaled to the desired standard deviation. The generated noise was then masked with a filter of valid pixels in the image. The resulting noise was then added to the image to simulate low-dose CT scan [30], [51]. Note that the simulated noise obtained from this procedure ensures that the spectral properties match to that observed in the actual low-dose CT perfusion data [30], [51].

Typically for whole-brain, CT perfusion data collected around tube current of 200 mA at 80 kVp with a scan time of less than a minute can be considered as high-dose and the data collected with varying only tube current between 15 and 100 mA can be considered as ultra to low dose CT perfusion data [52]. In this paper, the CT perfusion data collected between 15 and 50 mAs was considered as low-dose and the aim will be to denoise the CBF maps obtained using corresponding data.

G. Denoising Methods Utilized in This Paper

The denoising methods utilized in this paper were summarized in Fig. 1. Totally seven methods (including the one which does not use any denoising) were deployed in this paper for effective comparison. These methods are described below to bring out the differences between them.

Method-I—Low-Dose: In this case, the CT perfusion map (CBF) is generated from low-dose perfusion data using standard deconvolution method (Section II-B)

Method-II—Map Denoising (M-DCNN): For this, the maps generated using method-I become input to the deep CNN and the output of this method will be approximate solution to high-dose perfusion maps (Section II-E). The denoising of CBF is performed using deep CNN (Fig. 2 gives the network architecture). This method is proposed in this paper for robust denoising.

Method-III—T-DCNN: Similar to map denoising, in here the low-dose CT perfusion raw data (volume wise) becomes input to the deep CNN (architecture is same as Fig. 2) with high-dose CT perfusion raw data being the expected output. This denoised low-dose perfusion data is used as input to method-1 to result in CT perfusion map (CBF).

Method-IV—TM-DCNN: In here the result of method-III becomes input to the deep CNN (similar to method-II) with high dose CBF map being the expected output. This method is effectively combining methods-II and III, except that the part pertaining to method-II (map denoising) was retrained to include the maps generated using deep CNN-based denoised low-dose CT perfusion data. This method is proposed in this paper for comparison with only map denoising approach (method-II).

Method-V—T-BM3D: This method is similar to method-III except that instead of deep CNN for low-dose CT perfusion data denoising, in here BM3D filtering (Section II-D) was applied. This is a standard approach for denoising of low-dose CT perfusion raw data.

Method-VI—TM-BM3D: Same as method-IV except low-dose CT perfusion raw data denoising being performed using BM3D. This method is effectively combination of methods-II and V. Note that the deep CNN used for maps denoising was retrained here. This method is proposed in this paper for comparison with only map denoising approach (method-II).

Method-VII—Online-SPD: Online SPD-based method (Section II-C). This method was utilized in this paper as state-of-the-art for comparison with other methods described here.

Method-VIII—M-BM3D: This method is similar to method-II except that instead of using DCNN for low-dose map denoising, in here BM3D filtering (Section II-D) was applied.

Note all methods that utilize deep CNN (methods-II, III, IV, and VI) for either denoising the low-dose CT perfusion data or generated maps were all retrained for each case individually to map the low-dose input and high-dose outputs appropriately.

H. Figures of Merit

To evaluate the effectiveness of the deep CNN-based methods described above in terms of denoising the low-dose CBF map, two quantitative metrics were utilized. Note that the high-dose CBF maps obtained from CT perfusion data acquired at 190 mAs was considered as a reference.

1) Root-Mean-Square-Error: Root-mean-square-error (RMSE) measures the closeness of the denoised CBF map with the reference (high-dose CBF map). It is defined as

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(h_i - \bar{h}_i)^2} \tag{16}$$

where h_i represents the computed CBF map (either low-dose one or denoised output) and $\bar{h}_i$ corresponds to the high-dose reference CBF map with n representing the total number of pixels/voxels in the map.

2) Structural Similarity: Structural similarity (SSIM) index is a similarity measure between the reference and the distorted images, which is widely used across many image reconstruction/processing algorithms. This index provides an estimate of distortion in the reconstructed/denoised perfusion map. It is defined as [53]

$$\mathrm{SSIM} = \frac{\left(2\mu_{\bar{h}_i}\mu_{h_i} + C_1\right)\left(2\sigma_{\bar{h}_i h_i} + C_2\right)}{\left(\mu^2_{\bar{h}_i} + \mu^2_{h_i} + C_1\right)\left(\sigma^2_{\bar{h}_i} + \sigma^2_{h_i} + C_2\right)} \tag{17}$$

where $C_1 = K_1 L$, $C_2 = K_2 L$ with L denoting the dynamic range of pixel values and K_1 and K_2 are small constants, μ_{h_i} and $\mu_{\bar{h}_i}$ are the mean of h_i and $\bar{h}_i$, respectively, σ_{h_i} and $\sigma_{\bar{h}_i}$ are standard deviation of h_i and $\bar{h}_i$, respectively, and $\sigma_{\bar{h}_i h_i}$ is covariance of h_i and $\bar{h}_i$. The values of SSIM can vary from 0 to 1, the closer the value to one indicates the superior performance of the denoising algorithm.

III. Digital Phantom and In-Vivo Experiments

The methods discussed in this paper till now were systematically compared with the state-of-the-art online SPD method using digital brain phantom [47] and *in-vivo* clinical data. Digital brain phantom data was generated using the

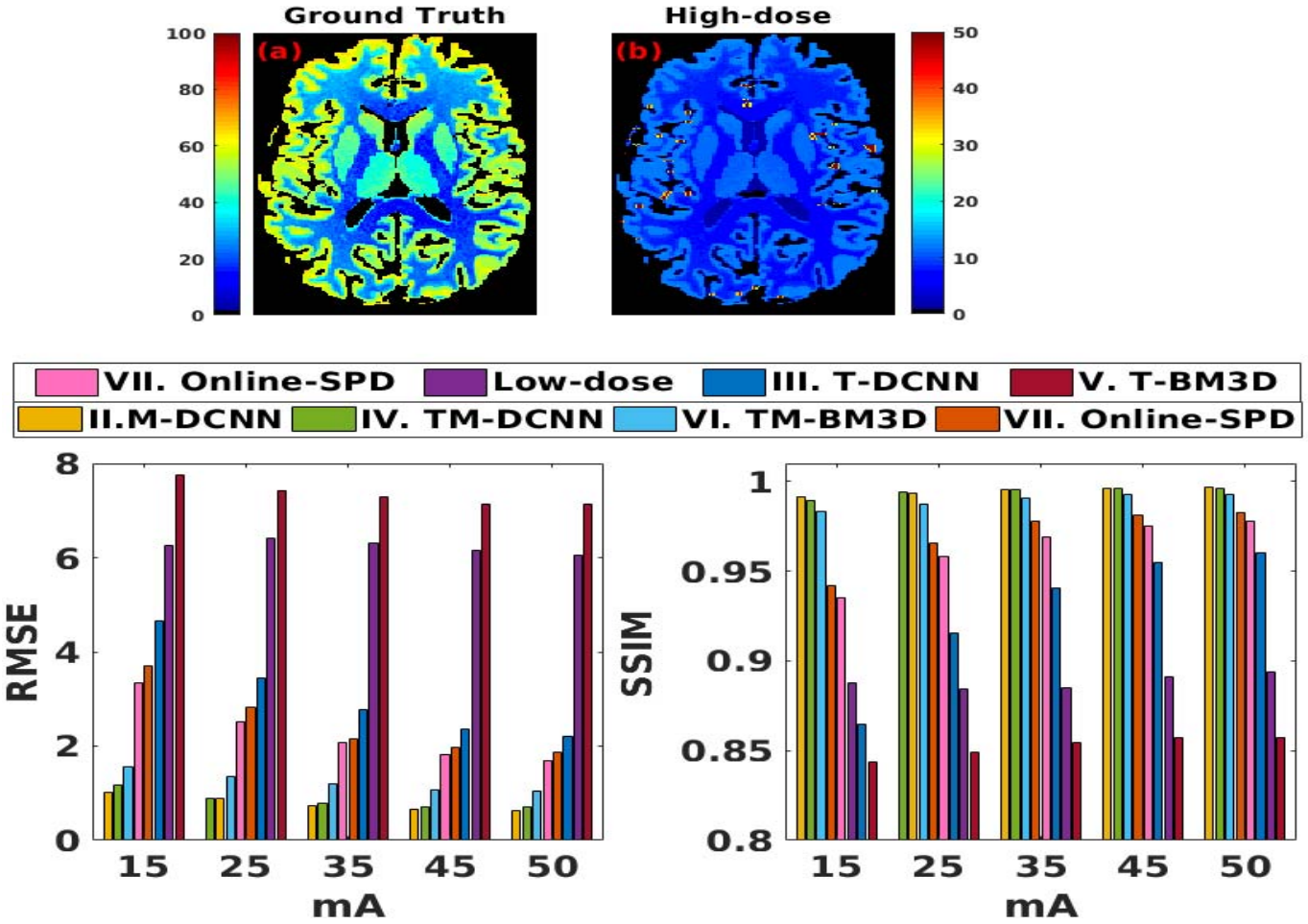

Fig. 4. CBF map. (a) Ground truth and (b) high-dose (reference), of middle slice of the digital brain CT perfusion phantom mimicking the normal CBF. The figures of merit, RMSE, and SSIM computed for CBF maps obtained using denoised methods discussed in Section II-G. (also shown in Fig. 1) are given in (c) and (d), respectively. The corresponding CBF maps pertaining to middle slice are given in Fig. 5.

digital brain perfusion phantom package [47] as described in Section II-E. There are five digital brain phantoms that were used for evaluating the effectiveness of the proposed method and this data was not utilized in the training of the network (Fig. 2). The clinical *in-vivo* data was also not used in the training.

Initially, only the digital brain phantom data was utilized to know the best performing method among the proposed methods that utilize DCNN (methods-II, III, IV, and VI) in comparison with online-SPD method (method-VII). After initial evaluation, only the method-II (M-DCNN), which was shown to outperform other DCNN-based methods was retained for comparison with online-SPD method.

The clinical data that was utilized in this paper was provided as an open-source as part of SPD toolbox [54]. The description of the *in-vivo* data acquisition was provided in [8]. It is briefly reviewed here for completeness. The *in-vivo* data was acquired with using GE Light speed or Pro-16 scanner with cine 4i scanning mode at 1 rotation per second. The X-ray tube was operated at 80 kVp and 190 mA to acquire the high-dose perfusion data. To minimize the radiation exposure, a scanning volume of 2.0 cm was utilized consisting of four slices at 5.0 mm thickness with its inferior selected at the level of basal ganglia above the orbits. A nonionic iodinated contrast of approximately 45 mL was intravenously injected using a power injected at a speed of 5 mL/s with a 5-s delay. The data corresponds to a 35-year-old female with left middle cerebral artery (LMCA) perfusion deficit caused by vasospasm in aneurysmal SAH. Note that the low-dose CT perfusion data is generated using the methodology given in Section II-F.

IV. Results

A. Digital Brain Phantom

The results pertaining to both digital brain phantoms that were not used in the training were shown in Figs. 4–6. The low-dose (method-I) results corresponding to 15, 25, 35, 45, and 50 mA are shown in the first row of Fig. 5. The results corresponding to only map denoising using the proposed network were shown in the second row of Fig. 5. The third row of Fig. 5 corresponds to the maps generated after denoising the temporal data using the network (method-III) given in Fig. 2. The maps generated are trained with the high dose maps as ground truth and the maps obtained after denoising both in the temporal and map domain (method-IV) are shown in the fourth row of Fig. 5. The denoised maps generated after temporal denoising using BM3D (method-V) are given in the fifth row of Fig. 5. The sixth row in Fig. 5 corresponds to the maps

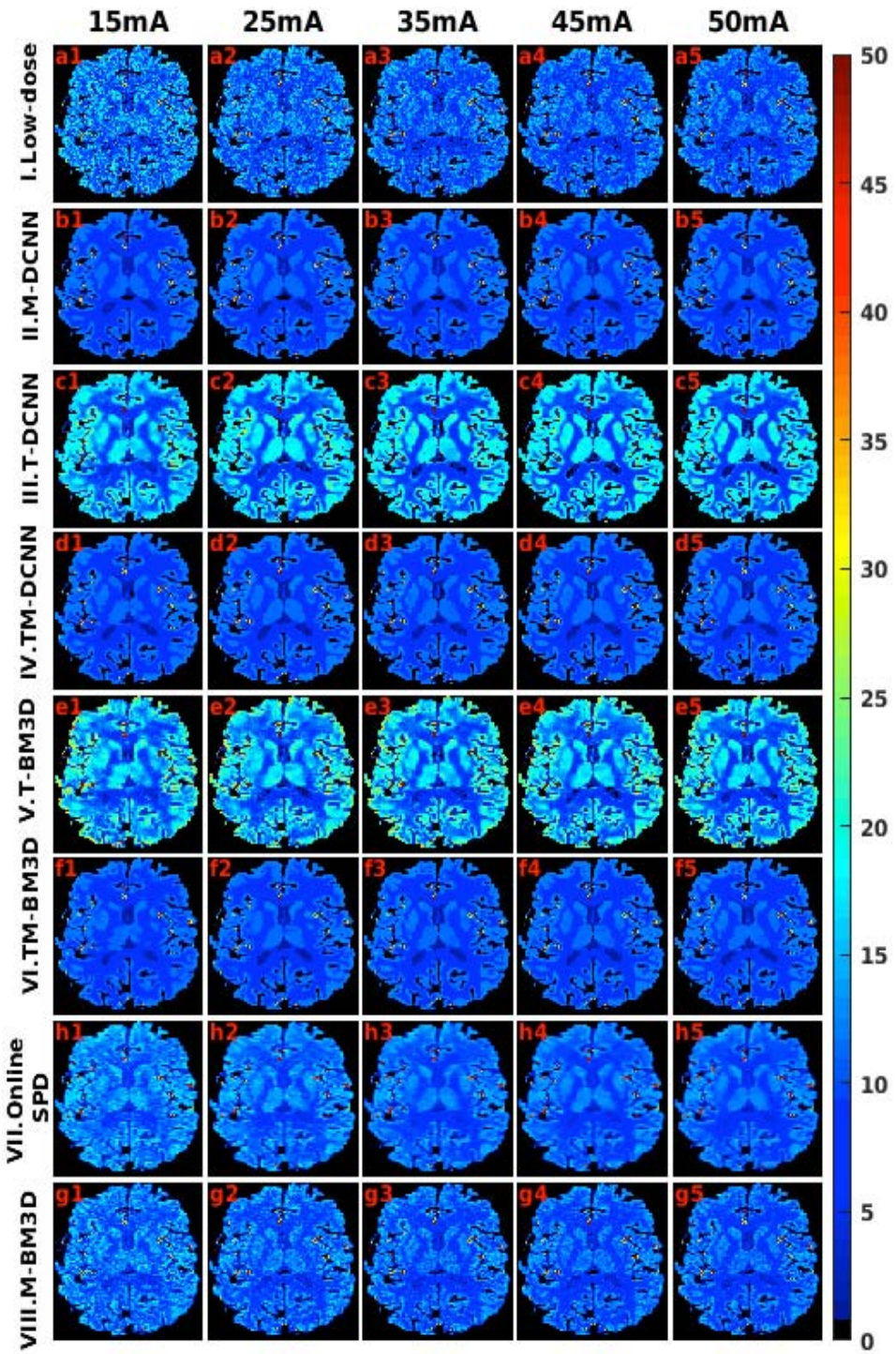

Fig. 5. Results pertaining to the middle slice of the digital brain CT perfusion phantom mimicking the normal CBF. The low-dose CBF maps (obtained using method-I described in Section II-G) computed using the perfusion data mimicking the X-ray tube current of (a1) 15 mA, (a2) 25 mA, (a3) 35 mA, (a4) 45 mA, and (a5) 50 mA. The proposed M-DCNN (method-II in Section II-G) results are given in (b1)–(b5) correspondingly. The corresponding denoised maps using method-III (T-DCNN) are shown in (c1)–(c5). The denoised results corresponding to method-IV (TM-DCNN) are given as (d1)–(d5), respectively. The maps generated from method-V (T-BM3D) are given in (e1)–(e5), respectively. The denoised maps corresponding to method-VI (TM-BM3D) are shown in (f1)–(f5). Corresponding denoised maps using the online SPD (method-VII in Section II-G) are given in (h1)–(h5). The denoised results corresponding to method-VIII (M-BM3D) are given as (g1)–(g5), respectively. The CBF maps corresponding to ground truth (mimicking healthy brain) and high-dose (reference) perfusion data are shown in the first row of Fig. 4. The figures of merit for these results are shown in second row of Fig. 4.

generated by denoising the BM3D temporal denoised maps using the network (method-VI). As stated earlier, the online SPD was chosen as the state-of-the-art method for comparing the results obtained using method-II (M-DCNN) and the corresponding results for the healthy brain phantom are shown in the seventh row of Fig. 5. The denoised maps generated

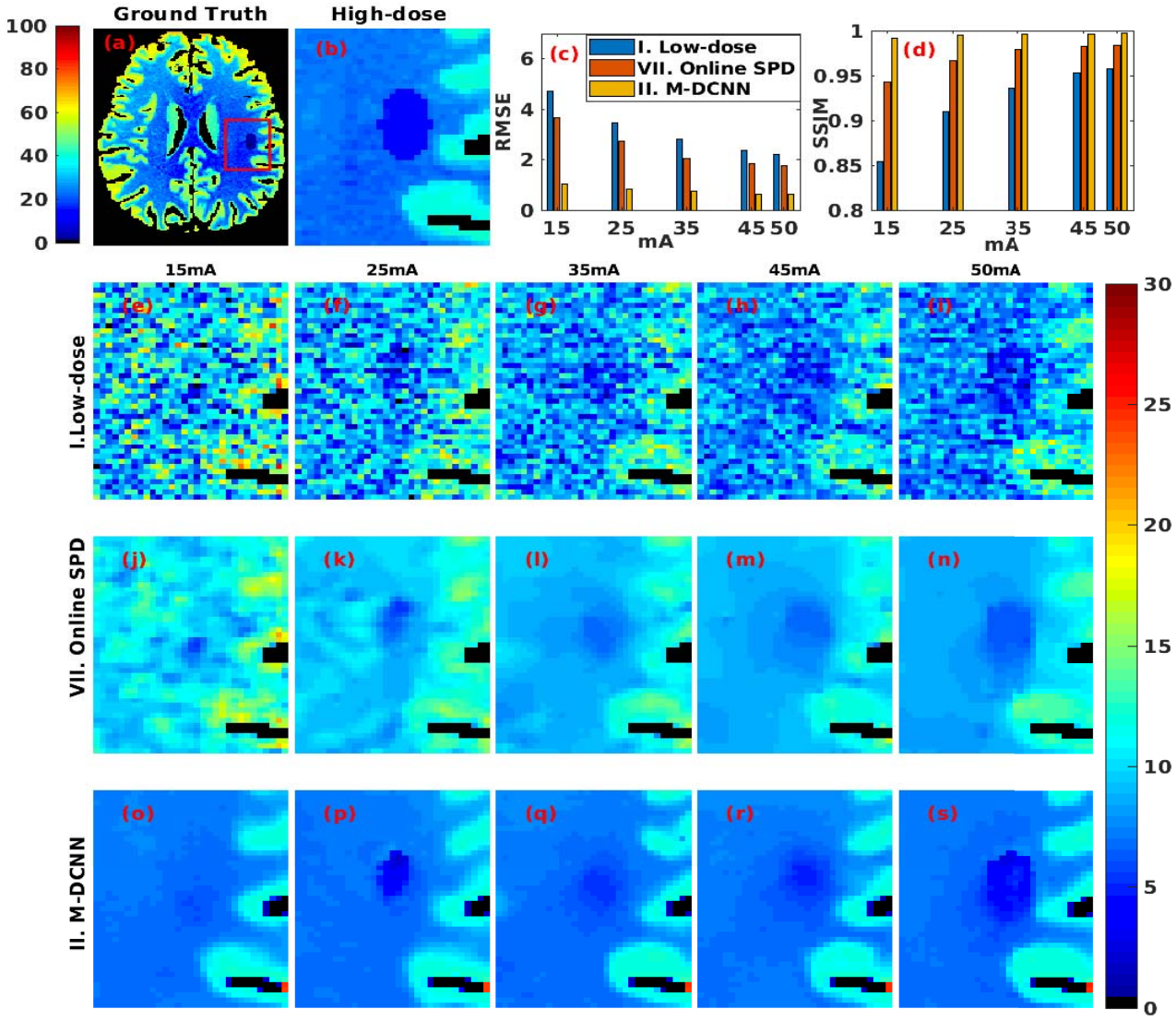

Fig. 6. Similar effort as Fig. 5 using digital brain CT perfusion phantom data for a case mimicking cerebrovascular abnormality (similar to stroke). For effective comparison, the abnormal (stroke mimicking) region of interest outlined as a box in the ground truth was zoomed in for other results. The figures of merit (RMSE and SSIM) were computed for the whole slice to show the improvement using the proposed M-DCNN method.

after map denoising using BM3D (method-VIII) are given in the last row of Fig. 5. Fig. 4 provides the ground truth along with CBF obtained with high-dose data. The figures of merit for the results shown in Fig. 5 are plotted in Fig. 4.

The results pertaining to healthy brain phantom case (Figs. 4 and 5) show that the proposed DCNN-based methods for map denoising outperforms the online SPD in terms of figures of merit. The RMSE values for 15 mAs CBF maps (shown in Figs. 4 and 5) are 4.7, 1.1, 6.5, 1.2, 7.7, 1.6, 3.7, and 3.3 for methods-I–VIII, respectively, indicating that the map denoising using DCNN improves the RMSE by atleast a factor of 2.5. Note that the parameters for online SPD algorithm were chosen for each specific mAs, its performance is better than the denoising methods applied in the raw data domain. Among the map denoising methods, the method-II (which only applies denoising in the map domain) has been consistently outperforming others across all mAs as well as other data. The results pertaining to other testing data has been shown later (Fig. 11) asserting that trends shown in these results were followed with other testing data establishing that method-II provides the best performance. From now on, only results pertaining to method-II (M-DCNN) were shown to compare with online-SPD-based method. In here when DCNN or BM3D was utilized for denoising of the raw data, it did not improve the derived maps significantly. Note that the X-ray tube currents considered here (range of 15–50 mAs) is typically labeled as ultra low-dose, which tend to be more noisy. Moreover, this can also be due to alteration in the underlying model (AIF) caused by the denoising process, which directly affect the deconvolution utilized in the map generation. Thus, methods-III (T-DCNN) and V (T-BM3D) are not effective. Note that the methods-IV (TM-DCNN) and VI (TM-BM3D), which utilize methods-III and V as initial steps, respectively, were able to improve the derived maps as they have additional step of map denoising using DCNN (similar to method-III), which is independent of any underlying model. Still errors

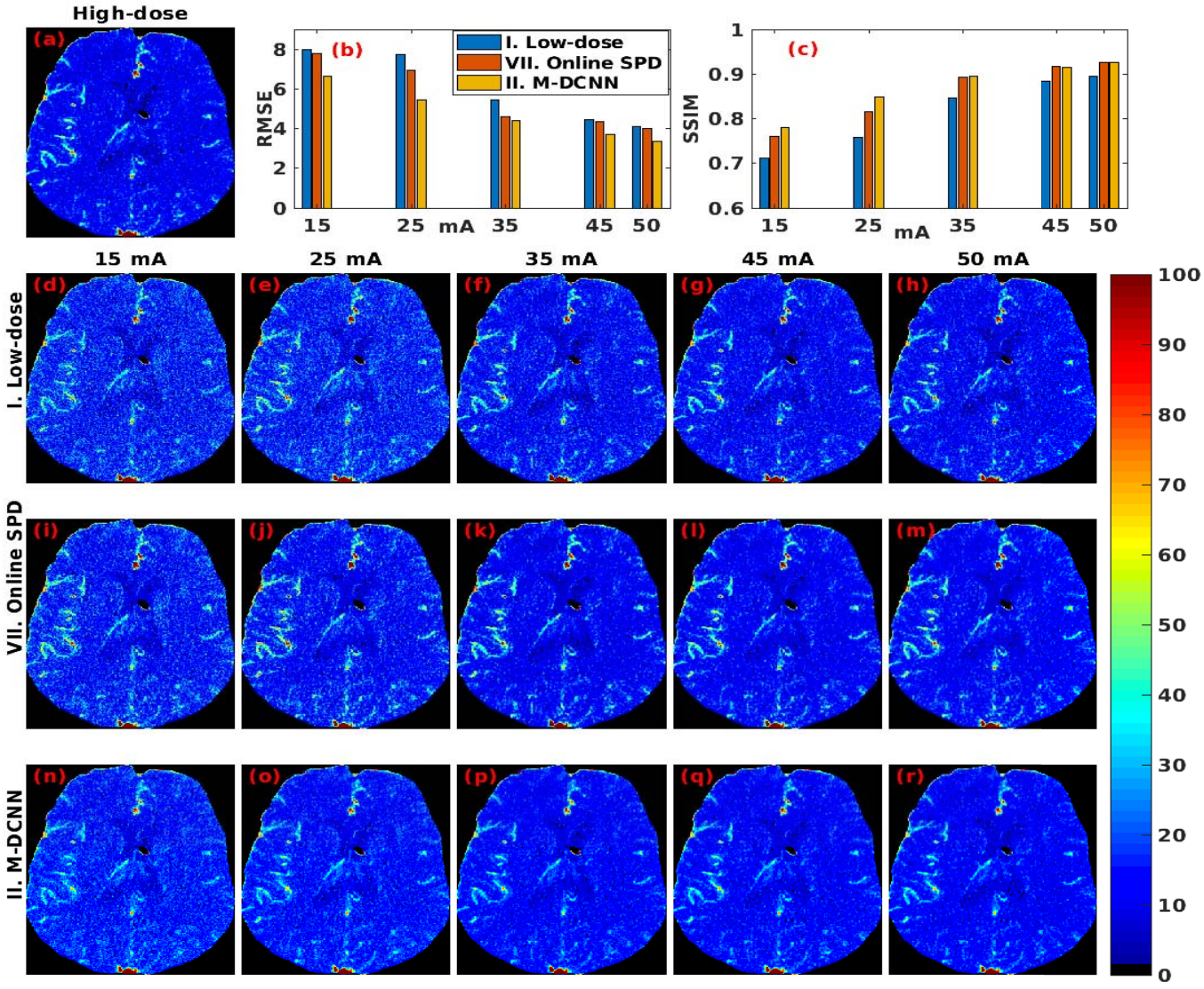

Fig. 7. Results pertaining to the middle slice of *in-vivo* CT perfusion data of a 35 years old female with vasospasm in aneurysmal SAH, which caused LMCA perfusion deficit. The reference CBF map corresponding to the high-dose CT perfusion data acquired at 190 mA is given in (a). The low-dose CBF maps (obtained using method-I described in Section II-G) computed using the analogous perfusion data corresponding to the X-ray tube current of (d) 15 mA, (e) 25 mA, (f) 35 mA, (g) 45 mA, and (h) 50 mA. Corresponding denoised maps using the online SPD (method-VII in Section II-G) are given in (i)–(m). The proposed M-DCNN (method-II in Section II-G) results are given in (n)–(r) correspondingly. The figures of merit, RMSE, and SSIM computed for last three rows of CBF maps are shown in (b) and (c), respectively.

induced by the raw-data denoising methods effect the final outcome, making them less attractive despite of additional denoising in the raw-data domain. It should be noted that the method-IV (TM-DCNN) stands out as the second best method next to method-II (M-DCNN) and the difference is less than 20% across all experiments performed in this paper.

B. Computational Time

Recorded computational time (in seconds) for obtaining results given in Figs. 4 and 5 for various methods with 50 time points are as follows. M-DCNN method took 0.6 s on CPU and 0.006 s on GPU. T-DCNN took 30 s with GPU and 0.3 s on GPU. TM-DCNN took 30.6 s on CPU and 0.306 s on GPU. T-BM3D took 3.5 s on CPU. TM-BM3D took 4.1 s on GPU and 3.506 s on GPU (temporal part on CPU). Online SPD method took 36.3 s and M-BM3D method took 0.07 s on GPU. All timings for DCNN methods are excluding the one time overhead time of training the network for 50 epochs on a batch-size of 256 which is approximately 72 h. This time could be further brought down by utilization of GPUs. The proposed method-II (M-DCNN) as well as M-BM3D (method-VIII) operates on the computed CBF map and is independent of the time points and marginally depends on the CBF map size. The CBF image size in case of digital brain phantom case is 132×184. It is also important to note that the training data and time required for online SPD is much smaller compared to the deep convolution network-based methods utilized in this paper.

C. Stroke Mimicking Digital Brain Phantom

For the case of stroke mimicking (Fig. 6), the results clearly indicate that the M-DCNN (method-II) method was able to consistently provide improved denoising results compared to

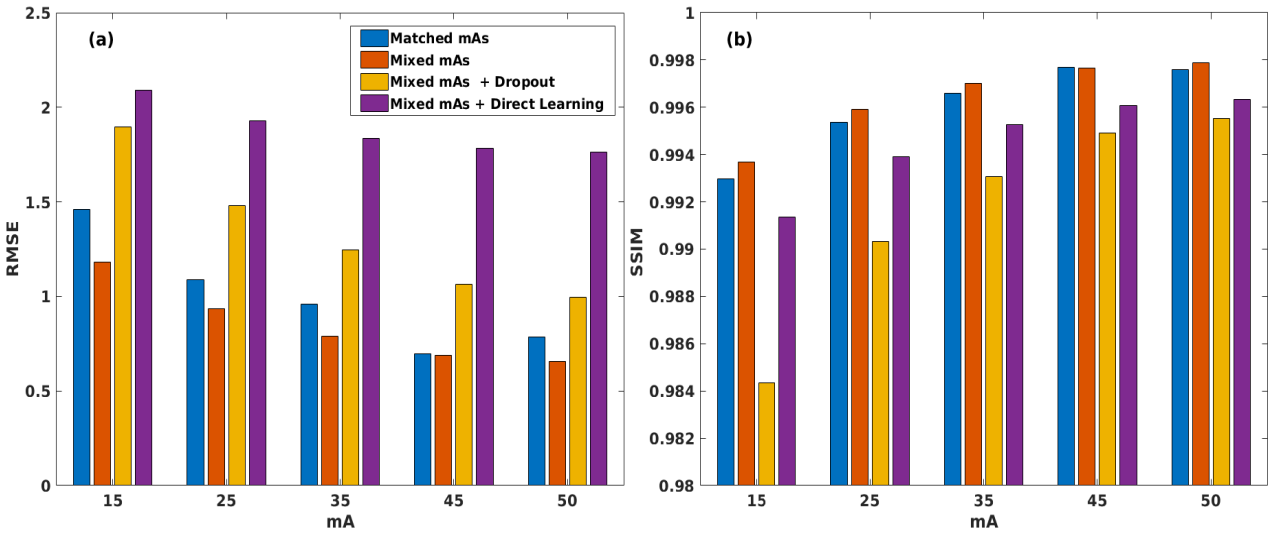

Fig. 8. Plot of averaged figures of merit. (a) RMSE and (b) SSIM on testing data (two unseen digital phantoms) with respect to different X-ray tube currents in method-II (M-DCNN). The different models used for training are listed in the legend. The basic block composition is given in Fig. 2. These results correspond to the number of blocks (referred as D blocks in Fig. 2) being 15.

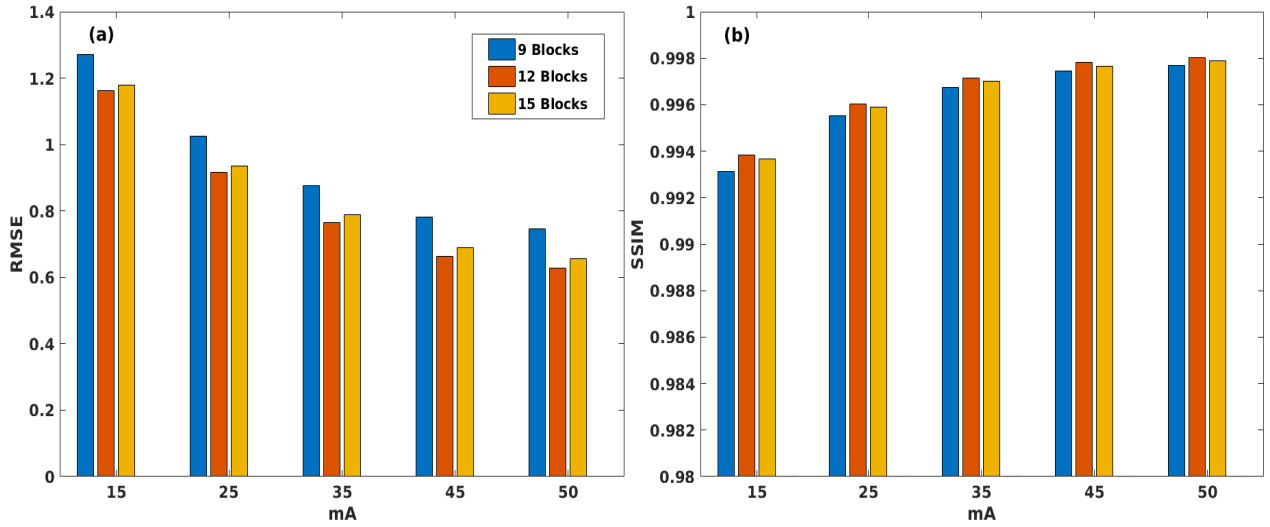

Fig. 9. Plot of averaged figures of merit. (a) RMSE and (b) SSIM on testing data (two unseen digital phantoms) with respect to different X-ray tube currents in method-II (M-DCNN) with varying number of blocks (referred as D blocks in Fig. 2) with mixed mAs model.

online SPD. Even in these cases, the figures of merit indicate that the M-DCNN was able to outperform the online SPD. The stroke mimicking lesion is more dispersed in the online SPD case, giving an impression of more stroke volume compared to the reference image. Note that results pertaining to other methods (discussed in Section II-G) were not presented here, but the trends observed for the healthy digital brain phantom case were valid here as well. As the input image size and number of time points are same as the results presented in Figs. 4 and 5, the computational times observed here are similar to the ones presented in the previous section.

D. In-Vivo Data

The denoised results corresponding to the *in-vivo* data are given in Fig. 7. The same trend as observed in the digital phantom cases was also observed here, with the proposed method-II (M-DCNN) outperforming the online SPD. In the 15 mA case, M-DCNN (method-II) outperforms the online SPD in terms of both figures of merit. It should be observed that, in terms of RMSE, the proposed M-DCNN outperforms the online SPD. In case of SSIM, the results of online SPD method for 35, 45, and 50 mAs closely match with proposed M-DCNN method, asserting that the similarity with high dose CBF map is same for both methods. As the number of time points in this case were 119 and the CBF map size being 367×280, the required computation time for the online SPD is 478.7 s and for method-II (M-DCNN) CPU implementation requires 2.2 s (with GPU, the recorded time is only 8 ms). As the online SPD computational time is highly dependent on the number of time points, this method turns out to be expensive in terms of computational time. Considering the training time, overall the online SPD still be an attractive option for performing the denoising step.

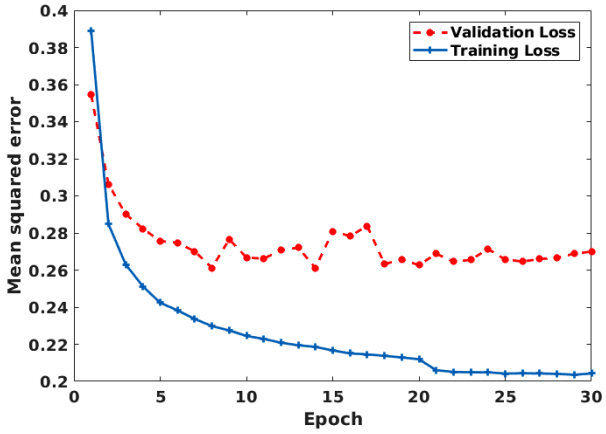

Fig. 10. Plot of training loss and validation loss with respect to the number of epochs for M-DCNN method (Fig. 1) with 12 blocks utilizing mixed mAs model. The network architecture is shown in Fig. 2.

V. Discussion

The online SPD built on underlying model of convolution for obtaining the tissue enhancement curves, may not be applicable for all cases of brain perfusion. For example, other models such as adiabatic approximation to tissue homogeneity [55], [56] cannot be directly utilized in the online SPD framework. The DCNN-based methods (including M-DCNN) is independent of the underlying model and requires only the low-dose perfusion map (which can be computed using any model), making it universally appealing.

A. Model Selection/Training

The proposed method-II (M-DCNN) was analyzed to determine effect of the underlying neural network model and the corresponding results were given in Fig. 8. Matched mAs correspond to the results generated when the trained model mAs and testing data mAs match. Mixed mAs represents the results generated by considering the training data with all mAs (this model was utilized for generating all results presented till now). Note that mixed mAs and matched mAs (the first two results) are without considering any dropout. It can be observed from Fig. 8 that the mixed model performs superior than the matched training model in all cases considered. Further analysis was performed with underlying model being mixed mAs. The results corresponding to adding dropout of 0.1 and direct learning (output of the network being clean map instead of noise) are also shown in Fig. 8. Note that the dropout was added after every intermediate layer excluding the first and last layers. From the results, it can be seen that the mixed model without dropout combined with residual learning outperforms the model with dropout or direct learning.

The training time for deep CNN-based methods could also be reduced by using less number of blocks in the network, but reducing the number of blocks can also reduce the performance. To study the effect of number of blocks on the performance of the trained network (specific to method-II), the number of blocks were varied from 9 to 15 using the mixed model as the reference. The results obtained in terms of the RMSE and the SSIM for various values of input mAs is reported in Fig. 9 as a function of model blocks (in here, they were varied as 9, 12, and 15). From Fig. 9, it can be ascertained that the model with 12 blocks was able to improve in all cases considered here. Thus, the trained model with 12 blocks was used for all results (Figs. 4–7) presented in this paper. The choice of number of layers/blocks is dependent on the data available for training and there is usually an optimal number of parameters in terms of number of blocks and this optimal parameter size is chosen based on the performance on a held out validation data-set. In this case, it was observed that increasing the number of blocks improved the performance initially. As increasing the depth of the network increases large number of trainable parameters to be estimated from the available data, there is always an optimal number of layers/blocks, which provide better approximation to the expected output for the validation data. In here, this optimal number of blocks were 12 as observed from results shown in Fig. 9.

As the proposed network utilizes the BN, it is invariant to the input dynamic range. This provides an inherent advantage of avoiding the data calibration across various dosages of iodine as well as varying sizes of brain. It is also shown in here that the chosen model corresponding to mixed training model was able to improve the SSIM and provide lower RMSE values for input CBF maps corresponding to 15, 25, 35, and 45 mAs (Figs. 5–7).

The plot of the training and validation loss over the number of epochs for the proposed method-II (M-DCNN) was shown in Fig. 10. It can be observed that the training and validation decreases with increase in epochs. As expected, the validation error is higher than the training error. After epoch 5, the validation has become consistent with increase in epochs. These results assert that the model did not over-fit the validation data and the training process followed the trends that are observable in the deep learning frameworks.

The results presented in Figs. 4–6 were limited to one slice. To show the improvement by utilizing the proposed method (M-DCNN), the average figures of merit RMSE and SSIM for two unseen testing phantoms (over the whole volume) for the methods discussed in Fig. 1 were shown in Fig. 11. It can be observed that the trends observed with respect to single slice were also valid here asserting that the improvement with respect to proposed M-DCNN method is across the whole volume.

This is the first work (as far as we know), where we are directly comparing the performance of denoising methods applied to raw-data as well as derived maps either separately or together in the low-dose CT perfusion imaging. The results presented here indicate that performing denoising in the derived map domain has larger applicability both in terms of improving the maps as well as offering computational efficiency especially with trained deep neural network models. The training time in the network models along with large data availability is a bottle neck for developing these type of models. With more realistic digital phantoms being available [57], this large data requirement can be largely met. The training time (in here being minimum of 36 h) can be reduced either by using multiple GPUs or implementing transfer learning framework [58]. As advances in GPU computing toward deep

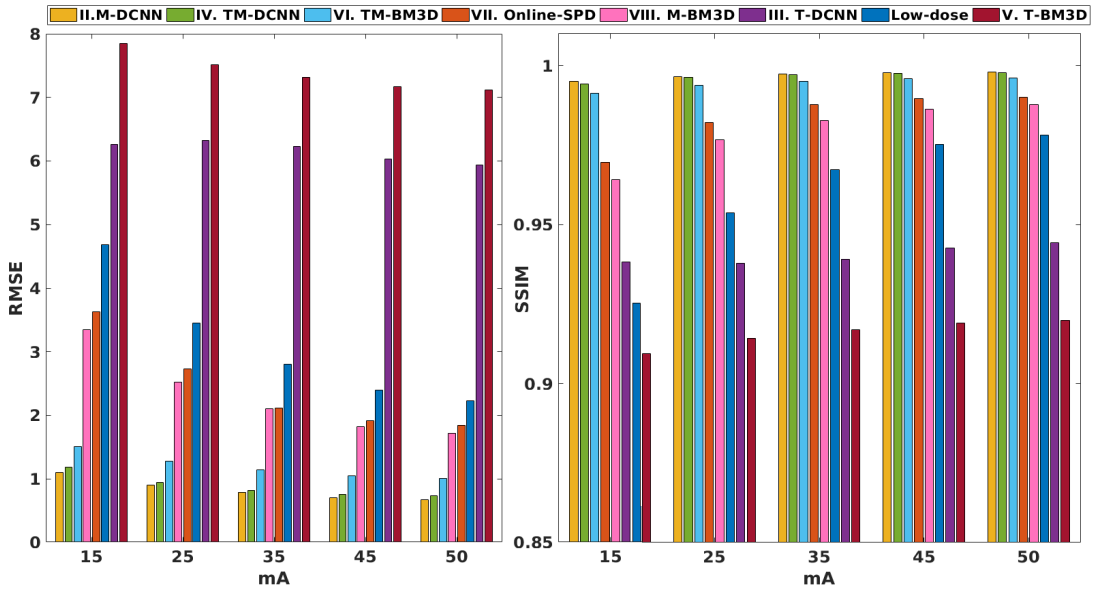

Fig. 11. Plot of average figures of merit. (a) RMSE and (b) SSIM on testing data (two unseen digital phantoms) using denoising methods discussed in this paper (Fig. 1) as a function of X-ray tube current (x-axis).

learning are becoming prominent, training DCNN on an optimized hardware (including multiple GPUs) should be able to bring down the training time to minutes [59]. It is also important to note that the training typically happens offline and is only performed once, the network is available online, requiring only couple of seconds on the CPU for performing the denoising (please refer to Section IV-B).

It is important to note that the residual network that was proposed here was inspired for performing denoising using DCNN with noise model being nonstandard additive white Gaussian [21]. The noise in the low-dose CT perfusion images is correlated and more importantly the noise model for the derived maps is unknown as computing them will involve a deconvolution process (solving an inverse problem) applied to the raw data. The standard denoising approaches (such as Wiener filtering) are not relevant in these cases and currently the state-of-the-art is to use sparse deconvolution approaches (such as the one used in here, namely, online SPD). The DCNNs provide generalizability in these scenarios to have larger applicability as they are purely data driven. It is possible to achieve the denoising of derived maps using other alternative deep CNN architectures. The results shown here indicate that deep CNN architectures applied in the map domain provides superior denoising performance compared to applying on the raw-data.

B. Future Extensions

Most works related to enhancing low-dose CT perfusion maps relied on denoising of raw-data (CTP data). The low-dose cases are typically mimicked by adding noise to high-dose CTP data with well established and calibrated statistical noise model for varying X-ray tube currents [8], [10], [27], [28], [30], [51]. This noise model with varying tube currents (15–50 mAs) was utilized in this paper to mimic the low-dose case. It is possible to add the noise in the sinogram space to mimic low-dose cases. This requires one to characterize the effect of tube current on the sinogram data to effectively compare the performance of various discussed methods in this paper. As CT reconstruction methods (especially advanced ones like statistical iterative reconstruction methods) are nonlinear in nature and the noise propagation through this reconstruction step can alter the CTP raw-data considerably [23], [25], [28], [29]. So this was not attempted in here and will be taken up as a future extension.

Even though the perfusion map that was considered here for denoising was only CBF, the other derived maps (such as CBV and MTT) can also be benefited in this framework. Also, the work presented here was focused toward CT perfusion, but it could be easily extended to the other perfusion imaging modalities, such as MRI, PET, SPECT, and ultrasound. The trends observed here should be applicable to these imaging modalities.

As low-dose CT imaging is receiving heightened interest in the recent past, methods such as the ones developed here should pave way for the low-dose CT imaging with an aim to move toward ultra low-dose CT imaging [9], [60]. In dynamic CT imaging, such as the perfusion, the contention for the low-dose CT imaging is its inability to provide quantitatively accurate parametric maps. This paper has shown that the deep learning could be the tool that can enable this without compromising the quantitativeness of the derived maps. Importantly, these deep networks can be implemented online to know the improvement in real-time (requiring negligible time than computation of derived maps), which can provide instant feedback for optimizing the image acquisition protocol with an aim to reduce the dose to the patient.

The future work also involves deriving perfusion maps directly from the low-dose CT perfusion raw data without

utilization of deconvolution. This approach requires a larger generalization as the number of time points in the raw data varies considerably across scan protocols [1], [2], [30] and will involve normalization across temporal domain to achieve high level of success in giving meaningful maps. The standard approach of mapping from temporal domain to map domain involves significant number of preprocessing steps other than deconvolution and each step utility in the direct CNN-based methods needs to be studied in detail to provide insights into solving the problem at hand.

VI. Conclusion

The derived maps from low-dose CT perfusion data are typically low-quality and requires a denoising to provide any quantitative information. The denoising methods that are currently available are either computationally complex or does not provide the required improvement in terms of image quality. In this paper, a data driven DCNN was proposed to perform denoising of low-dose CT raw-data or derived maps. It was shown through the experiments that the DCNN-based methods were able to improve the low-dose CT perfusion imaging with these being applied in the derived map domain providing superior performance. The network architecture is also insensitive to the input data dynamic range as it integrates residual learning with BN. The digital brain phantom as well as *in-vivo* data cases demonstrate the effectiveness of the DCNN applied in the map domain without adding any additional computational burden in performing the denoising step. Integrating this DCNN with the post-processing software platforms should be able to ease the work flow in terms of diagnosing as well as characterizing the cerebrovascular abnormalities using low-dose perfusion data. The developed MATLAB code was provided as an open-source for enthusiastic users [61].

Acknowledgment

V. S. Kadimesetty would like to thank Samsung Research and Development Institute India, Bengaluru, India, for sponsoring his Ph.D. in Indian Institute of Science, Bengaluru, India. The authors are thankful to Dr. R. Fang, who gave permission to utilize the *in-vivo* data for evaluation of the proposed algorithm as well as providing the SPD toolbox as an open-source.

References

[1] E. G. Hoeffner *et al.*, "Cerebral perfusion CT: Technique and clinical applications," *Radiology*, vol. 231, no. 3, pp. 632–644, Jun. 2004.

[2] M. R. Harrigan, J. Leonardo, K. J. Gibbons, L. R. Guterman, and L. N. Hopkins, "CT perfusion cerebral blood flow imaging in neurological critical care," *Neurocrit. Care*, vol. 2, no. 3, pp. 352–366, Jun. 2005.

[3] M. Wintermark and M. H. Lev, "FDA investigates the safety of brain perfusion CT," *Amer. J. Neuroradiol.*, vol. 31, no. 1, pp. 2–3, Jan. 2010.

[4] L. Ostergaard, R. M. Weisskoff, D. A. Chesler, C. Gyldensted, and B. R. Rosen, "High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: Mathematical approach and statistical analysis," *Magn. Reson. Med.*, vol. 36, no. 5, pp. 715–725, Nov. 1996.

[5] L. Ostergaard *et al.*, "High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part II: Experimental comparison and preliminary results," *Magn. Reson. Med.*, vol. 36, no. 5, pp. 726–736, Nov. 1996.

[6] H. J. Wittsack *et al.*, "CT-perfusion imaging of the human brain: Advanced deconvolution analysis using circulant singular value decomposition," *Comput. Med. Imag. Graph.*, vol. 32, no. 1, pp. 67–77, Jan. 2008.

[7] A. Buades, B. Coll, and J.-M. Morel, "A review of image denoising algorithms, with a new one," *Multiscale Model. Simulat.*, vol. 4, no. 2, pp. 490–530, Feb. 2005.

[8] R. Fang, T. Chen, and P. C. Sanelli, "Towards robust deconvolution of low-dose perfusion CT: Sparse perfusion deconvolution using online dictionary learning," *Med. Image Anal.*, vol. 17, no. 4, pp. 417–428, May 2013.

[9] A. Mnyusiwalla, R. I. Aviv, and S. P. Symons, "Radiation dose from multidetector row CT imaging for acute stroke," *Neuroradiology*, vol. 51, no. 10, pp. 635–640, Oct. 2009.

[10] R. Fang, T. Chen, and P. C. Sanelli, "Tissue-specific sparse deconvolution for low-dose CT perfusion," in *Proc. Int. Conf. Med. Image Comput. Comput. Assisted Interv.*, Sep. 2013, pp. 114–121.

[11] Y. LeCun, Y. Bengio, and G. Hinton, "Deep learning," *Nature*, vol. 521, no. 7553, pp. 436–444, May 2015.

[12] G. Litjens *et al.*, "A survey on deep learning in medical image analysis," *Med. Image Anal.*, vol. 42, pp. 60–88, Dec. 2017.

[13] D. Bahdanau, K. Cho, and Y. Bengio, "Neural machine translation by jointly learning to align and translate," *arXiv preprint arXiv:1409.0473*, Sep. 2014.

[14] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 1097–1105.

[15] R. Wu, S. Yan, Y. Shan, Q. Dang, and G. Sun, "Deep image: Scaling up image recognition," *arXiv preprint arXiv:1501.02876*, Jan. 2015.

[16] J. Portilla, V. Strela, M. J. Wainwright, and E. P. Simoncelli, "Image denoising using scale mixtures of Gaussians in the wavelet domain," *IEEE Trans Image Process.*, vol. 12, no. 11, pp. 1338–1351, Nov. 2003.

[17] M. Elad and M. Aharon, "Image denoising via sparse and redundant representations over learned dictionaries," *IEEE Trans. Image Process.*, vol. 15, no. 12, pp. 3736–3745, Dec. 2006.

[18] J. Xie, L. Xu, and E. Chen, "Image denoising and inpainting with deep neural networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2012, pp. 341–349.

[19] H. C. Burger, C. J. Schuler, and S. Harmeling, "Image denoising: Can plain neural networks compete with BM3D?" in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2392–2399.

[20] V. Jain and S. Seung, "Natural image denoising with convolutional networks," in *Proc. Adv. Neural Inf. Process. Syst.*, 2009, pp. 769–776.

[21] K. Zhang, W. Zuo, Y. Chen, D. Meng, and L. Zhang, "Beyond a Gaussian denoiser: Residual learning of deep CNN for image denoising," *IEEE Trans. Image Process.*, vol. 26, no. 7, pp. 3142–3155, Jul. 2017.

[22] R. Zhang *et al.*, "A Gaussian mixture MRF for model-based iterative reconstruction with applications to low-dose X-ray CT," *IEEE Trans. Comput. Imag.*, vol. 2, no. 3, pp. 359–374, Sep. 2016.

[23] H. Chen *et al.*, "Low-dose CT via convolutional neural network," *Biomed. Opt. Exp.*, vol. 8, no. 2, pp. 679–694, Feb. 2017.

[24] E. Kang, J. Min, and J. C. Ye, "A deep convolutional neural network using directional wavelets for low-dose X-ray CT reconstruction," *Med. Phys.*, vol. 44, no. 10, pp. e360–e375, Oct. 2017.

[25] H. Chen *et al.*, "Low-dose CT with a residual encoder–decoder convolutional neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2524–2535, Dec. 2017.

[26] K. H. Jin, M. T. McCann, E. Froustey, and M. Unser, "Deep convolutional neural network for inverse problems in imaging," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4509–4522, Sep. 2017.

[27] J. M. Wolterink, T. Leiner, M. A. Viergever, and I. Išgum, "Generative adversarial networks for noise reduction in low-dose CT," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2536–2545, Dec. 2017.

[28] W. Yang *et al.*, "Improving low-dose CT image using residual convolutional network," *IEEE Access*, vol. 5, pp. 24698–24705, 2017.

[29] D. Wu, K. Kim, G. E. Fakhri, and Q. Li, "Iterative low-dose CT reconstruction with priors trained by artificial neural network," *IEEE Trans. Med. Imag.*, vol. 36, no. 12, pp. 2479–2486, Dec. 2017.

[30] K. Juluru *et al.*, "Effects of increased image noise on image quality and quantitative interpretation in brain CT perfusion," *Amer. J. Neuroradiol.*, vol. 34, no. 8, pp. 1506–1512, Aug. 2013.

[31] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising with block-matching and 3D filtering," in *Proc. SPIE Electron. Imag.*, vol. 6064. San Jose, CA, USA, Feb. 2006, Art. no. 606414.

[32] K. Dabov, A. Foi, V. Katkovnik, and K. Egiazarian, "Image denoising by sparse 3-D transform-domain collaborative filtering," *IEEE Trans. Image Process.*, vol. 16, no. 8, pp. 2080–2095, Aug. 2007.

[33] O. Wu *et al.*, "Tracer arrival timing-insensitive technique for estimating flow in MR perfusion-weighted imaging using singular value decomposition with a block-circulant deconvolution matrix," *Magn. Reson. Med.*, vol. 50, no. 1, pp. 164–174, Jul. 2003.

[34] A. Cenic, D. G. Nabavi, R. A. Craen, A. W. Gelb, and T. Y. Lee "Dynamic CT measurement of cerebral blood flow: A validation study," *Amer. J. Neuroradiol.*, vol. 20, no. 1, pp. 63–73, Jan. 1999.

[35] A. Cenic, D. G. Nabavi, R. A. Craen, A. W. Gelb, and T. Y. Lee, "A CT method to measure hemodynamics in brain tumors: Validation and application of cerebral blood flow maps," *Amer. J. Neuroradiol.*, vol. 21, no. 3, pp. 462–470, Mar. 2000.

[36] 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.*, vol. 2011, May 2011, Art. no. 467563.

[37] J. Mairal, F. Bach, J. Ponce, and G. Sapiro, "Online dictionary learning for sparse coding," in *Proc. 26th Annu. Int. Conf. Mach. Learn.*, Jun. 2009, pp. 689–696.

[38] D. L. Donoho, "For most large underdetermined systems of linear equations the minimal ℓ_1-norm solution is also the sparsest solution," *Commun. Pure Appl. Math.*, vol. 59, no. 6, pp. 797–829, Jun. 2006.

[39] B. Efron, T. Hastie, I. Johnstone, and R. Tibshirani, "Least angle regression," *Ann. Stat.*, vol. 32, no. 2, pp. 407–499, Apr. 2004.

[40] V. Katkovnik, A. Foi, K. Egiazarian, and J. Astola, "From local kernel to nonlocal multiple-model image denoising," *Int. J. Comput. Vis.*, vol. 86, no. 1, pp. 1–32, Jan. 2010.

[41] *Image and Video Denoising by Sparse 3D Transform-Domain Collaborative Filtering*. Accessed: Mar. 14, 2018. [Online]. Available: http://www.cs.tut.fi/∼foi/GCF-BM3D/index.html#ref_software

[42] K. He, X. Zhang, S. Ren, and J. Sun, "Spatial pyramid pooling in deep convolutional networks for visual recognition," in *Proc. Eur. Conf. Comput. Vis.*, 2014, pp. 346–361.

[43] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *Proc. Int. Conf. Learn. Represent.*, 2015, pp. 1–14.

[44] Y. Bengio, P. Simard, and P. Frasconi, "Learning long-term dependencies with gradient descent is difficult," *IEEE Trans. Neural Netw.*, vol. 5, no. 2, pp. 157–166, Mar. 1994.

[45] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," in *Proc. Int. Conf. Mach. Learn.*, Jun. 2015, pp. 448–456.

[46] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Las Vegas, NV, USA, 2016, pp. 770–778.

[47] A. Aichert *et al.*, "A realistic digital phantom for perfusion C-arm CT based on MRI data," in *Proc. IEEE Nucl. Sci. Symp. Med. Imag. Conf. Rec. (NSS/MIC)*, Seoul, South Korea, 2013, pp. 1–2.

[48] J. Wang and L. Perez, "The effectiveness of data augmentation in image classification using deep learning," Class Notes CS 231n, Stanford Univ., Stanford, CA, USA, 2017. Accessed: Nov. 21, 2017. [Online]. Available: http://cs231n.stanford.edu/reports/2017/pdfs/300.pdf

[49] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2015, pp. 1026–1034.

[50] A. Vedaldi and K. Lenc, "MatConvNet: Convolutional neural networks for MATLAB," in *Proc. 23rd ACM Int. Conf. Multimedia (ACM)*, Oct. 2015, pp. 689–692.

[51] A. J. Britten, M. Crotty, H. Kiremidjian, A. Grundy, and 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," *Brit. J. Radiol.*, vol. 77, no. 916, pp. 323–328, Apr. 2004.

[52] R. Manniesing, M. T. H. Oei, B. Ginneken, and M. Prokop, "Quantitative dose dependency analysis of whole-brain CT perfusion imaging," *Radiology*, vol. 278, no. 1, pp. 190–197, Jun. 2016.

[53] Z. Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: From error visibility to structural similarity," *IEEE Trans. Image Process.*, vol. 13, no. 4, pp. 600–612, Apr. 2004.

[54] *Sparse Perfusion Deconvolution (SPD) Package*. Accessed: Nov. 21, 2017. [Online]. Available: https://github.com/ruogufang/SPD

[55] L. E. Kershaw and H. L. Cheng, "Temporal resolution and SNR requirements for accurate DCE-MRI data analysis using the AATH model," *Magn. Reson. Med.*, vol. 64, no. 6, pp. 1772–1780, Dec. 2010.

[56] K. S. Lawrence and T. Y. Lee, "An adiabatic approximation to the tissue homogeneity model for water exchange in the brain: I. Theoretical derivation," *J. Cerebr. Blood Flow Metab.*, vol. 18, no. 12, pp. 1365–1377, Dec. 1998.

[57] R. Manniesing, C. Brune, B. van Ginneken, and M. Prokop, "A 4D CT digital phantom of an individual human brain for perfusion analysis," *Peer J.*, vol. 4, Nov. 2016, Art. no. e2683.

[58] J. Yosinski, J. Clune, Y. Bengio, and H. Lipson, "How transferable are features in deep neural networks?" in *Proc. Adv. Neural Inf. Process. Syst.*, 2014, pp. 3320–3328.

[59] *Whitepaper: GPU-Based Deep Learning Inference: A Performance and Power Analysis*, NVIDIA, Santa Clara, CA, USA. Accessed: Nov. 21, 2017. [Online]. Available: http://www.nvidia.com/object/white-papers.html

[60] Y. W. Lui, E. R. Tang, A. M. Allmendinger, and V. Spektor, "Evaluation of CT perfusion in the setting of cerebral ischemia: Patterns and pitfalls," *Amer. J. Neuroradiol.*, vol. 31, no. 9, pp. 1552–1563, Oct. 2010.

[61] *Convolutional Neural Network-Based Robust Denoising of Low-Dose Computed Tomography Perfusion Maps*. Accessed: Mar. 14, 2018. [Online]. Available: https://sites.google.com/site/sercmig/home/ct-perfusion-denoising