

Deeply Accelerated Arterial Spin Labeling Perfusion MRI for Measuring Cerebral Blood Flow and Arterial Transit Time

Yiran Li  and Ze Wang , Senior Member, IEEE

Abstract—Cerebral blood flow (CBF) indicates both vascular integrity and brain function. Regional CBF can be non-invasively measured with arterial spin labeling (ASL) perfusion MRI. By repeating the same ASL MRI sequence several times, each with a different post-labeling delay (PLD), another important neurovascular index, the arterial transit time (ATT) can be estimated by fitting the acquired ASL signal to a kinetic model. This process however faces two challenges: one is the multiplicatively prolonged scan time, making it impractical for clinical use due to the escalated risk of motions; the other is the reduced signal-to-noise-ratio (SNR) in the long PLD scans due to the T1 decay of the labeled spins. Increasing SNR needs more repetitions which will further increase the total scan time. Currently, there lacks a way to accurately estimate ATT from a parsimonious number of PLDs. In this paper, we proposed a deep learning-based algorithm to reduce the number of PLDs and to accurately estimate ATT and CBF. Two separate deep networks were trained: one is designed to estimate CBF and ATT from ASL data with a single PLD; the other is to estimate CBF and ATT from ASL data with two PLDs. The models were trained and tested using the large Human Connectome Project multiple-PLD ASL MRI. Performance of the DL-based approach was compared to the traditional full dataset-based data fitting approach. Our results showed that ATT and CBF can be reliably estimated using deep networks even with one PLD.

Index Terms—ATT, CBF, deep learning, deep residual network, wide activation.

I. INTRODUCTION

CEREBRAL blood flow (CBF) is a physiological measure fundamental to brain neuro-vascular health and brain function. CBF can be measured with several neuroimaging techniques but most of them require exogenous radioactive tracers

Manuscript received 8 December 2022; revised 3 June 2023 and 4 August 2023; accepted 28 August 2023. Date of publication 9 October 2023; date of current version 6 December 2023. This work was supported by NIH under Grants R01AG060054, R01AG070227, R01EB031080-01A1, R21AG080518, P41EB029460-01A1, and 1UL1TR003098. (*Corresponding author: Ze Wang*)

This work did not involve human subjects or animals in its research.

The authors are with the Department of Diagnostic Radiology and Nuclear Medicine, University of Maryland School of Medicine, Baltimore, MD 21021 USA (e-mail: yiran.li@som.umaryland.edu; ze.wang@som.umaryland.edu).

This article has supplementary downloadable material available at <https://doi.org/10.1109/JBHI.2023.3312662>, provided by the authors.

Digital Object Identifier 10.1109/JBHI.2023.3312662

or side-effect causing contrast agents. Arterial spin labeling (ASL) perfusion MRI remains the only non-invasive and non-radioactive technique for measuring CBF in a quantitative unit (ml/100g/min) across the whole brain [1]. ASL MRI uses the magnetically inverted inflowing arterial blood as the endogenous tracer. The labeled arterial blood will be transited to the imaging place and exchange with tissue water and subsequently change the tissue signal. This signal change is in proportion to CBF and is encoded in the corresponding image acquired in the imaging place, which is often called the label (L) image. To remove the background signal, a paired image (the control (C) image) acquisition is often performed using the same spin labeling and imaging sequence after modulating the spin labeling pulses to not change the net magnetization of the arterial blood. Perfusion signal is subsequently extracted from the C-L difference, which can be converted into the quantitative CBF in a unit of ml/100g/min using a kinetic model [1]. Because arterial spins are often labeled in a proximal place away from the imaging site, a post-labeling delay time (PLD) is necessary to allow labeled spins to reach the imaging site. If PLD is shorter than the arterial transit time (ATT), CBF underestimation will occur due to the insufficient arrival of the labeled spins.

ATT is the time of the labeled blood traveling from the labeling plane to the tissue voxel [2], which differs by region and individual brain. By repeating data acquisitions at different PLDs, ATT can be estimated by fitting the ASL MRI perfusion signals at different PLDs to the kinetic model [1], [3]. A practical issue, however, is the substantially increased total scan time due to the repetitions of the entire imaging process at multiple PLDs. The long scan time not only increases the cost of imaging but also increases the risk of head motions, making the multiple PLD ASL MRI difficult to implement in clinical research [4]. Reducing the number of PLDs will partially address the challenge but will make the ATT estimation less stable since the curve fitting-based estimation is prone to noise. There is a need to develop a method to reliably estimate ATT but without significantly increasing the total scan time, which however is highly challenging using conventional imaging method. This dilemma can now be solved by deep learning (DL), which has been successively adopted to model various highly complex relationships using multiple layers of hierarchical neural networks.

We have recently proposed a DL-based algorithm to solve a similar problem of accelerating the MR imaging process without sacrificing parameter quantification accuracy in the

U.S. Government work not protected by U.S. copyright.

glutamate-weighted Chemical Exchange Saturation Transfer (Glu-CEST) MRI [5]. Accurate glutamate weighted signal estimation in Glu-CEST MRI depends on the CEST signal spectrum, the so-called Z-spectrum. In practice, the Z-spectrum is approximately depicted by MR signal measured by repeating the same CEST MRI sequence at different off-resonance frequencies. Densely sampling the Z-spectrum can better estimate the Z-spectrum and subsequently the Glu-CEST quantification but will inevitably increase the total acquisition time. Robustly estimating the glutamate-weighted signal from a parsimoniously sampled Z-spectrum is technically challenging and is difficult to achieve using traditional methods. Using DL, we have demonstrated that the total scan time of Glu-CEST can be reduced by >70% without sacrificing the Glu-CEST quantification accuracy. The Z-spectrum measurement acceleration in Glu-CEST MRI is very similar to the problem of accelerating the multiple-PLD ASL MRI. The goal of multiple-PLD ASL MRI is to measure the tracer dynamics and then estimate ATT at the peak position of the curve of the tracer dynamics. To get a high-quality estimation of the curve, ASL MRI needs to be repeated at many different PLDs. The acceleration goal is then to parsimoniously sample the tracer kinetic curve at a few PLDs without sacrificing ATT and CBF estimation. While reducing the number of PLDs to one or two seems too optimistic, the low-dimensional spatial variations of kinetic curve along time and across space make it possible to use the spatial correlations to recover the temporal information missed during acquisitions. By implementing this space-for-time conversion concept through a nonlinear DL-based multivariate CBF and ATT learning, we hypothesize that we can reliably estimate the CBF and ATT from one or two PLDs ASL MRI data as compared to those estimated from the full set of multiple-PLDs ASL data.

Our contributions in this paper include: 1) we proposed a DL model to learn a nonlinear transform from a few perfusion-weighted images (PWIs) images to both the CBF and ATT maps on a large-scale dataset; 2) our method reduced the acquisition time by >60% compared to the conventional multiple-PLD based ATT and CBF estimation method; 3) our method improved image quality as measured by three performance indices (SSIM, PSNR, and MAE).

II. RELATED WORK

Prior-data guided DL has been applied to various medical imaging processing fields [6]. In ASL MRI, DL has been used to denoise CBF images, accelerating the imaging process, improve image resolution [7], [8], [9], [10], [11], [12], or provide an alternative approach to estimate the physiological measurements encoded in the ASL MRI time series [13], [14]. DL in ATT and CBF estimation has been assessed in three studies. In a pilot study [15], a 3D convolutional neural network (CNN) was proposed to estimate ATT and CBF simultaneously using the perfusion weighted images (PWIs) at multi-PLDs. Because PWIs are relative measures, without calibrations from the M0 scan the corresponding model may not be generalizable to data acquired at a different scanner or at different time. In another preliminary study, the same research group reported that CBF

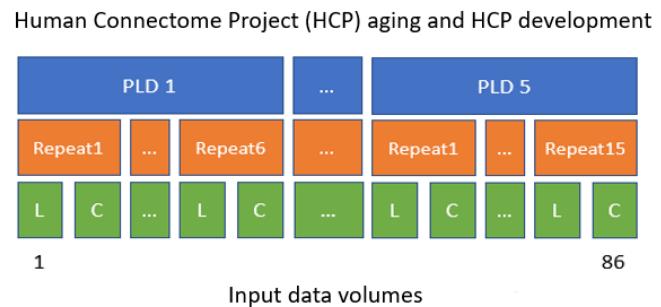


Fig. 1. HCP-A and HCP-D ASL acquisition scheme. Blue denotes the number of PLDs. Orange denotes the repetition acquisitions for each PLD. L denotes label, C denotes control.

and ATT can be estimated from ASL MRI at reduced number of PLDs without accuracy loss [16]. These two pilot studies were based on 12 subjects, which may substantially limit the generalizability of the trained network. In [17], a CNN and a UNet were implemented as an alternative to the traditional data fitting-based approach for estimating CBF and ATT from multiple-PLDs ASL data. The model was trained using data from 50 subjects. The authors showed that CBF and ATT estimation using DL can be stable for fewer (up to three missing PLDs) PLDs. Our work differs from the previous work by using a more sophisticated network, focusing on substantially shortening the total scan time by using one or two PLDs to estimate CBF and ATT maps, rigorously training and evaluating the networks using much larger training samples, and demonstrating model generalizability through a direct generalizability test by using the model trained using one dataset to a new dataset acquired from a totally different age group.

III. METHODOLOGY

A. Dataset Descriptions

Multiple-PLD ASL data were downloaded from Human Connectome Project aging dataset (HCP-A) and development dataset (HCP-D) [18]. Fig. 1 illustrates the HCP ASL MRI data acquisition protocol. The spin labeling time and imaging readout are the same for all L/C image pairs and all PLD blocks. Each PLD block corresponds to one ASL MRI acquisition scheme with a specific PLD. ASL data acquisitions were based on a pseudo-continuous arterial spin labeling (PCASL) and 2D multiband (MB)-echo-planar imaging (EPI) sequence [19], with following acquisition parameters: $86 \times 86 \times 65$ matrix, $3.5 \times 3.5 \times 3.5$ mm 3 voxel resolution, TR/TE = 3580/18.7 ms, pseudo-continuous labeling (1500 ms label duration. The same imaging process was repeated at 5 PLDs of 200, 700, 1200, 1700, and 2200 ms, containing 6, 6, 6, 10, 15 control-label image pairs, respectively, resulting in a total of 86 raw L or C images. Imaging slice acquisition time was 54 ms. Two M0 images for CBF quantification were acquired at the end of all the PLD acquisitions. To create the reference of CBF and ATT images, a voxel-wise ASL kinetic model fitting was applied using the PWIs from all the PLDs via BASIL [20] and confirmed with

ASLtbx [21], [22]. Slice-timing correction and motion correction were included via the toolbox as well. The raw ASL control and label images were successively subtracted and averaged to get the mean PWI, $\Delta M(t)$, for each PLD respectively. The kinetic model for pseudo-continuous ASL (pCASL) is described by the following equation [23]:

$$\Delta M(t) =$$

$$\begin{cases} 0 & 0 < t < \Delta t \\ 2M_0 f T'_1 \alpha e^{-\Delta t/T_{1b}} q_{ss}(t) & \Delta t < t < \tau + \Delta t \\ 2M_0 f T'_1 \alpha e^{-\Delta t/T_{1b}} e^{-(t-\tau-\Delta t)/T'_1} q_{ss}(t) & \tau + \Delta t < t \end{cases} \quad (1)$$

where M_0 is the fully relaxed MR signal, f is the magnitude of CBF, α is labeling efficiency, τ is labeling time, Δt is the ATT for the bolus to travel to the voxel, and

$$q_{ss}(t) = \begin{cases} 1 - e^{-(t-\Delta t)/T'_1} & \Delta t < t < \tau + \Delta t \\ 1 - e^{-\tau/T'_1} & \tau + \Delta t < t \end{cases} \quad (2)$$

$$\frac{1}{T'_1} = \frac{1}{T_1} - \frac{f}{\lambda} \quad (3)$$

where T_1 and T_{1b} are the longitudinal relaxation rate of tissue and blood, and λ is the blood/tissue water partition coefficient. The CBF and ATT maps are fit from variational Bayesian inference model voxel by voxel as the reference for the training [24]. T_1 is set to 1.5 s and T_{1b} is 1.65 s in 3T based on the literature.

The proposed full data supervised DL algorithm for estimating CBF and ATT from few PLD ASL MRI can be considered as an optimization problem aimed to learn a non-linear mapping function $F : x \rightarrow y$ using many training data samples $\{(x^{(i)}, y^{(i)}) : i = 1, \dots, N\}$.

$$F = \arg \min_{F \in \text{CNN}} \frac{1}{N} \sum_{i=1}^N \|g(x^{(i)}) - y^{(i)}\| \quad (4)$$

where CNN is a convolutional neural network with some tunable parameters. $\|\cdot\|$ is the L1 distance. F represents the implicit function for estimating CBF or ATT via deep learning. y is the reference, the training target ATT or CBF calculated from (1) to (3) with all the PLDs. x is the mean difference image ΔM at different PLDs. g means a fixed transform for removing the arbitrary image scale, which was the M_0 calibration in this paper. In this work, we aimed to recover y from x at one or two PLDs. For simplicity, the DL-based ATT estimation neural networks were dubbed as ATT-Nets.

B. Network Structures

1) Input Channel: ATT-Nets with one PLD and two PLDs were trained and tested separately. They differed mainly by the number of inputs. Fig. 2(a) shows the flow chart for preparing the input to the one-PLD ATT-Nets. The M_0 normalized mean and standard deviation of ΔM (PWI) images at one PLD ($= 1.7$ s in this work) were taken as the input. For two-PLD ATT-Nets, the input was the normalized ΔM and the PLD weighted PWIs (WD) [25] of two PLDs ($= 0.7/1.7$ s) (Fig. 2(b)). A brain mask was used to exclude outside brain voxels from model learning to reduce noise interference as those voxels only contain noise.

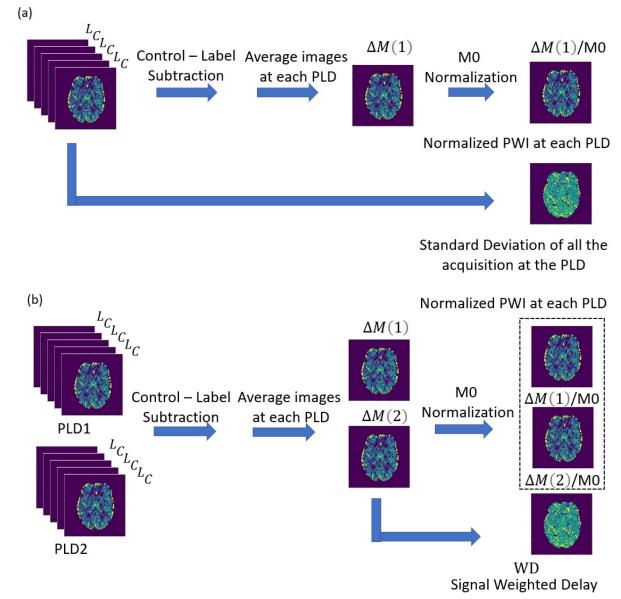


Fig. 2. Image input to the one-PLD ATT-Nets (a) and two-PLD ATT-Nets (b).

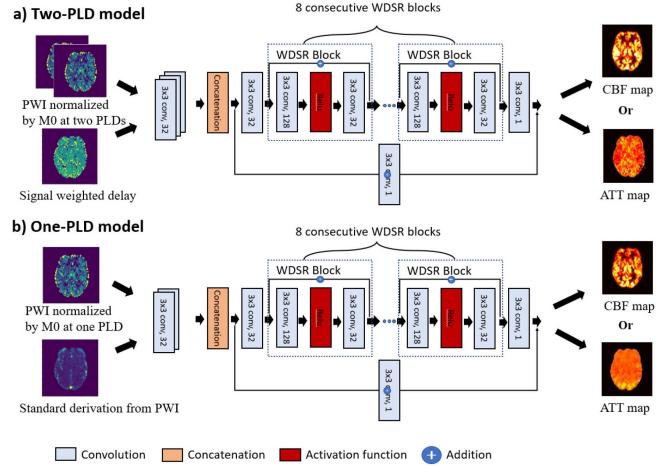


Fig. 3. Network architecture of the proposed ATT-Nets with one PLD (a) and two PLDs (b). The details of WDSR block are demonstrated in Fig. 4.

2) Network Architecture: Fig. 3 shows the architecture used in the one-PLD and two-PLD ATT-Nets. This architecture is an improved deep residual network [25]. The backbone of the network is the vanilla residual network while all the residual blocks are replaced by Wide-activation Deep Super-Resolution network (WDSR) blocks. WDSR uses more feature extraction filters while the residual identity mapping pathway needs to be slimmed. More filters (neurons) make the layer wider before the ReLU [26] activation layer in order to learn more features.

Fig. 4 illustrates the WDSR block. The width of the identity mapping pathway is assumed to be w_1 , and the width in the residual block before activation is w_2 . The expansion factor before activation is denoted as r thus $w_2 = r * w_1$. In the vanilla residual networks, we have $w_2 = w_1$ and the number

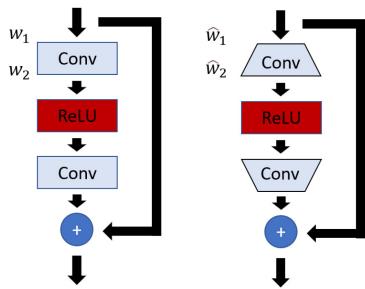


Fig. 4. Illustration of the vanilla (left) and wide activation block (right).

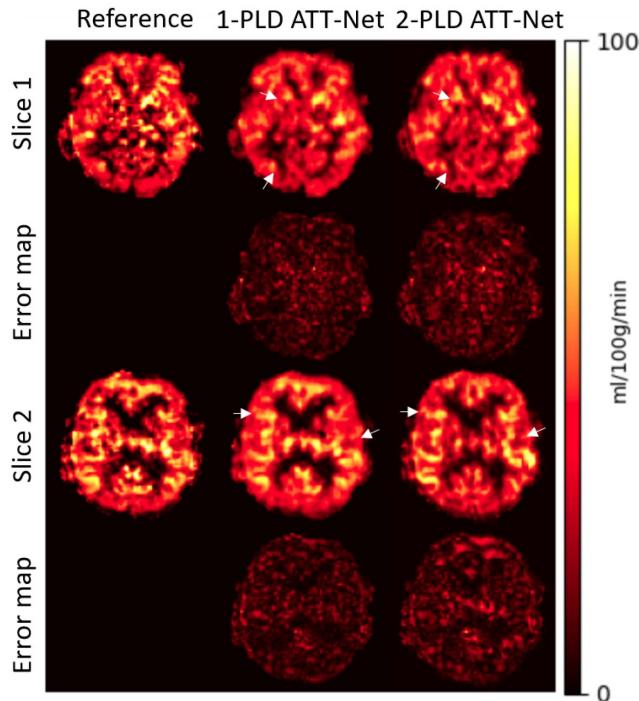


Fig. 5. Two CBF image slices from a randomly selected HCP-A subject. Arrows indicated the places with less distortion, better signal contrast, or suppressed artifacts.

of parameters is $2 \times w_1^2 \times k^2$ in each residual block where k is the number of input patches. If the input patch size is fixed, the computational complexity of the block is only related to w_1^2 . The WDSR block has same complexity $w_1^2 = \hat{w}_1 \times \hat{w}_2 = r \times \hat{w}_1^2$. Therefore, the residual identity mapping pathway needs to be reduced as a factor of \sqrt{r} and the filters can be expanded with \sqrt{r} times at the same time. Compared to the standard vanilla structure, WDSR does not change the number of parameters and computation complexity, but can provide better network performance [27]. We expected similar outcomes to predict ATT and CBF maps. The optimal number of WDSR blocks (8), the number of filters (32) and expansion ratio ($4\times$) for WDSR filters are also determined from this paper [27].

For both the one-PLD and two-PLD ATT-Nets, CBF and ATT were predicted from the input (described in Fig. 2) using a

separately trained model given the big difference of both image intensity and image contrast of the CBF and ATT maps. Each input channel was associated with a separate convolutional layer. The output of all input layers was concatenated into one channel as the input to the successive layer. Following each convolutional layer were eight consecutive WDSR blocks. In each block, wide activation was used to retain the high-frequency tissue boundary information. Following the eight consecutive blocks, another convolutional layer without any activation function was attached to the end to get the output with additional input from the second layer. The output would generate the CBF and ATT maps, respectively. Reference CBF and ATT images were generated using FSL [28] scripts using the PWIs from all the PLDs and were confirmed with the kinetic model fitting based approach implemented in ASLtbx.

3) Training Details: There are 679 qualified subjects in HCP-A dataset and 627 in HCP-D. For each dataset, 400 of them were randomly selected for training, another 100 subjects for testing, and the rest of the subjects for validation. The number of training and testing samples was kept the same for both datasets to conduct a fair comparison quantitatively.

The image size is $86 \times 86 \times 60$. For each subject, we extracted axial slices from slice 26 to slice 45 as high-quality training data. With an image size of 86×86 for each slice, 6000 image slices were used as the training set. All the training processes were implemented by using the entire image instead of patches. Each experiment was run by 300 epochs and a batch size of 32. The L1 loss function was used in training. The ADAM optimizer [29] with an initial learning rate of 0.001 was used to train the network. After 60 and 120 epochs, the learning rate was reduced by 0.1 and 0.05, respectively. All DL experiments were performed using framework of Keras and Tensorflow running on a Ubuntu18.04 system with a NVIDIA GTX 2080ti GPU card.

4) Evaluation: For each ATT-Nets, training was stopped after 300 epochs. The trained models were then tested using the test dataset that unseen by the model. The generated CBF and ATT maps were both registered to the corresponding T1-weighted images and then normalized into MNI space using scripts from ASLtbx, respectively. A couple of experiments were conducted to predict the CBF and ATT maps of all 100 subjects from testing sets on HCP-Aging and development, respectively. Some quantitative metrics and plots were used to evaluate the performance of the proposed method. Method performance was measured by visual image appearance, the structural similarity index (SSIM), peak signal-to-noise ratio (PSNR), and mean absolute error (MAE). Scatter plot of mean regional CBF and ATT was used to visualize the correspondence between the reference and DL output. Mean CBF and ATT values were extracted from 5 spatially distributed region-of-interest (ROIs): the dorsolateral prefrontal cortex (dlpfc), posterior cingulate cortex (dmnpcc), left amygdala (leftamy), left hippocampus (lelhipp), ventral striatum. Model generalizability was explicitly tested through direct model transferring. Specifically, the model trained with HCP-A data was directly applied to data from HCP-D to get CBF and ATT maps without retraining or fine-tuning the model with data from HCP-D. To evaluate whether the network output

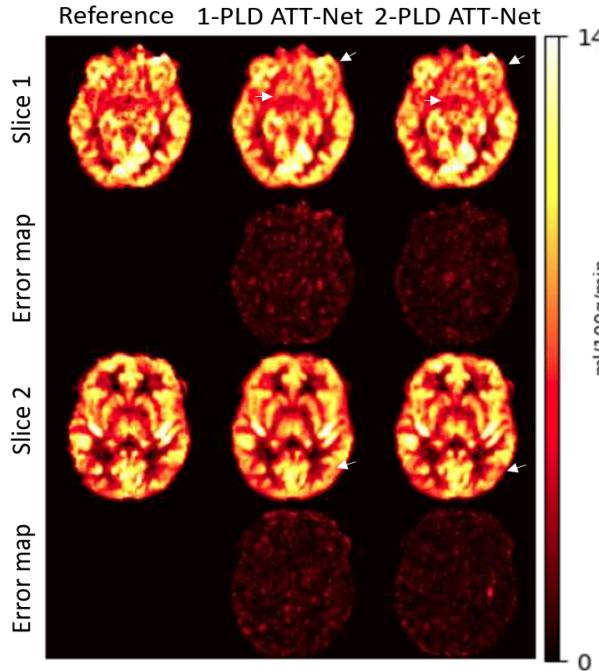


Fig. 6. CBF maps of randomly selected HCP-D subject. Two typical slices were displayed. Arrows indicated the places with less distortion, better signal contrast, or suppressed artifacts.

can preserve the physiological meanings of CBF and ATT, we compared the regional correlation between CBF/ATT and age calculated from the reference CBF/ATT and the network output.

IV. RESULTS

A. Qualitative Evaluations

Figs. 5 and 6 show the CBF estimation results for a representative subject of the HCP-A (Fig. 5) and HCP-D database (Fig. 6). As compared to the reference calculated by the conventional method in the first column, CBF images produced by the one-PLD and two-PLD ATT-Nets showed improved image quality in terms of less artifacts in grey matter, better signal contrast in the grey matter/white matter boundary as marked by the arrows in Figs. 5 and 6. CBF from the old subject (HCP-A) was lower than that of the young subject (HCP-D).

Figs. 7 and 8 show the ATT estimation results for the same subjects using separately trained one-PLD and two-PLD ATT-Nets. As compared to the reference, both the one-PLD and the two-PLD ATT-Nets showed better image contrast and reduced noise. The ATT maps of the one-PLD ATT-Nets showed bigger image context difference in the right side of the image in Fig. 7 and white matter in Fig. 8. The two-PLD ATT-Nets results were similar to the reference. ATT of the older subjects was higher than that of the young subject in grey matter.

B. Quantitative Evaluations I: SSIM, PSNR, and MAE

Tables I and II list the SSIM and PSNR values of the two DL models. For CBF, the two-PLD ATT-Nets performed slightly better than one-PLD though the difference is minor. For ATT, the

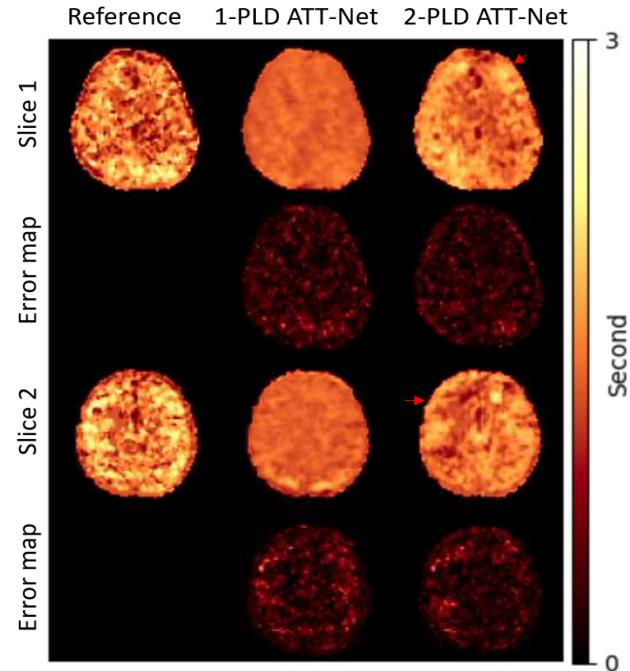


Fig. 7. ATT maps of randomly selected HCP-A subject. Two typical slices were displayed.

TABLE I
SSIM AND PSNR OF CBF MAPS

Method	HCP-A		HCP-D	
	SSIM	PSNR	SSIM	PSNR
One-PLD	0.8871±0.0606	34.42±16.54	0.8853±0.0668	32.57±20.93
Two-PLD	0.9069±0.0479	34.95±14.31	0.9167±0.0445	34.49±23.03

TABLE II
SSIM AND PSNR OF ATT MAPS

Method	HCP-A		HCP-D	
	SSIM	PSNR	SSIM	PSNR
One-PLD	0.8205±0.1087	66.53±13.32	0.8054±0.1165	66.99±14.97
Two-PLD	0.8763±0.0715	66.88±10.33	0.8909±0.0716	69.2948±15.85

two-PLD model significantly outperformed the one-PLD model significantly in terms of SSIM but not in PSNR. The PSNR of DL-estimated ATT was higher than that of the DL-estimated CBF because the quality of the original ATT images was worse than CBF, which was drastically improved by deep learning.

MAE was tested voxel by voxel from every testing subject. The two deep learning methods (one-PLD and two-PLD) for both ATT and CBF estimation demonstrated comparatively small values of MAE shown in Figs. 9 and 10. The one-PLD and two-PLD ATT-Nets showed no statistically significant difference in terms of CBF estimation on HCP-A dataset. For ATT estimation for HCP-A and HCP-D and CBF estimation for HCP-D, the two-PLD ATT-Nets outperformed the one-PLD

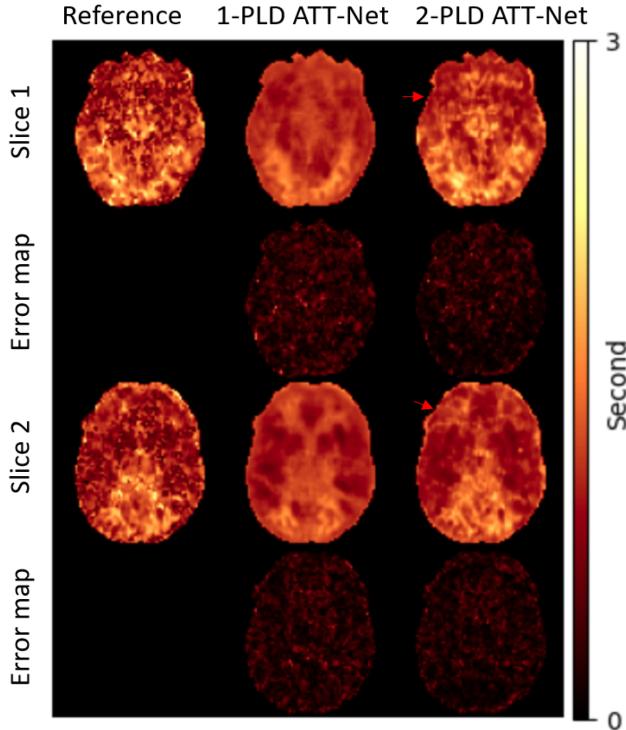


Fig. 8. ATT maps of randomly selected HCP-development subject. Two typical slices were displayed.

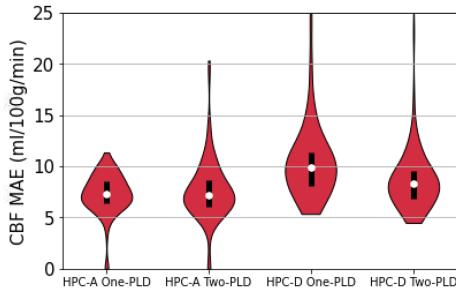


Fig. 9. MAE of CBF maps across different methods and datasets.

ATT-Nets with smaller average error and less standard deviation. In other words, the two-PLD model provided better and more stable predictions than the one-PLD model.

C. Scatter Plots of Mean ROI CBF and ATT

Figs. 11 and 12 show the mean ROI CBF and ATT of the reference and the network output. CBF and ATT estimated by different DL models are highly consistent with the ground truth though the ATT results showed bigger variations compared to CBF.

D. Direct Model Transferring

Figs. 13 and 14 show two slices of the CBF and ATT maps of a random testing subject selected from HCP-D for all methods. Compared to the reference, CBF images from the one-PLD and two-PLD ATT-Nets trained by HCP-A did not perform as well

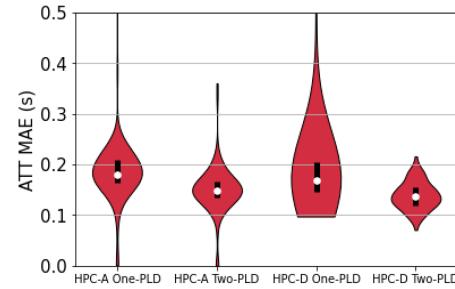


Fig. 10. MAE of ATT maps across different methods and datasets.

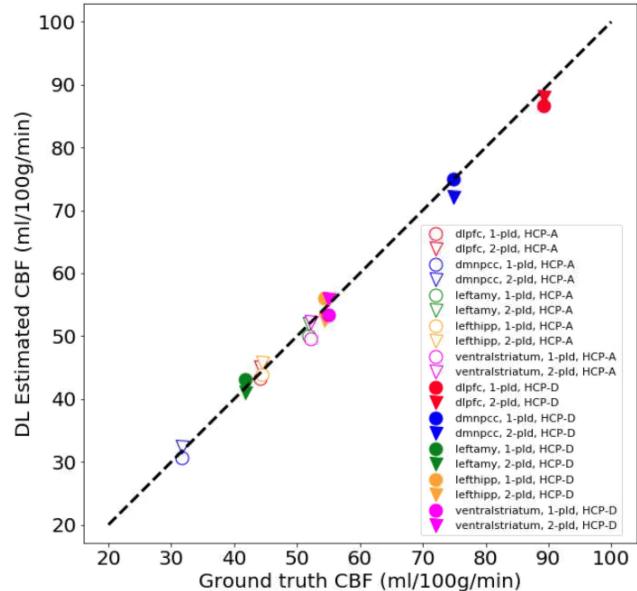


Fig. 11. Mean CBF value of 100 testing subjects in different ROIs. Colors denote different ROIs. Circle and triangle denote one-PLD and two-PLD models. Hollow and solid markers denote the HCP-aging and HCP-dev dataset, respectively.

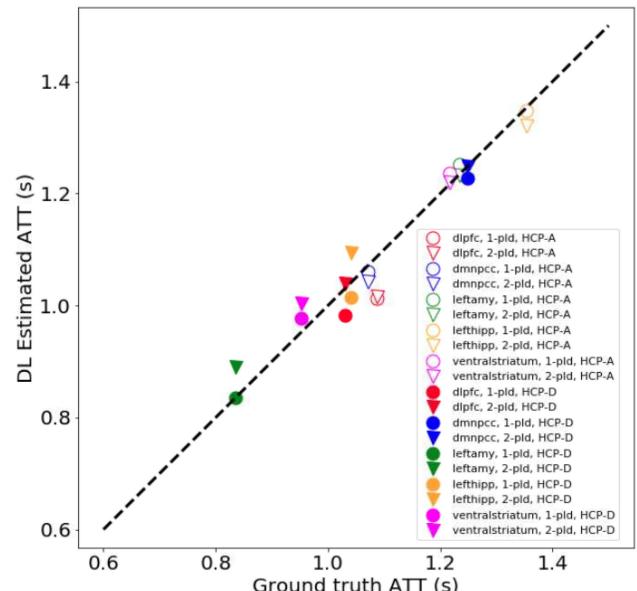


Fig. 12. Mean ATT value of 100 testing subjects in different ROIs.

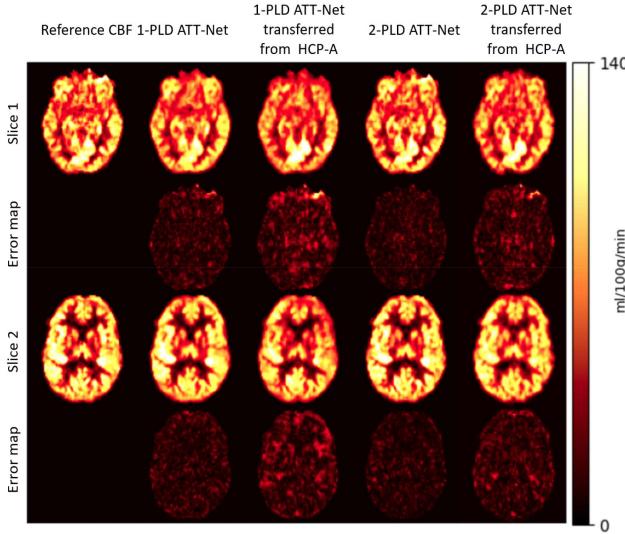


Fig. 13. Two CBF image slices of a representative HCP-D subject. “1-PLD ATT-Net transferred from HCP-A” means that the model was trained using HCP-A data rather than HCP-D data. Similar definition applies to “2-PLD ATT-Net transferred from HCP-A”.

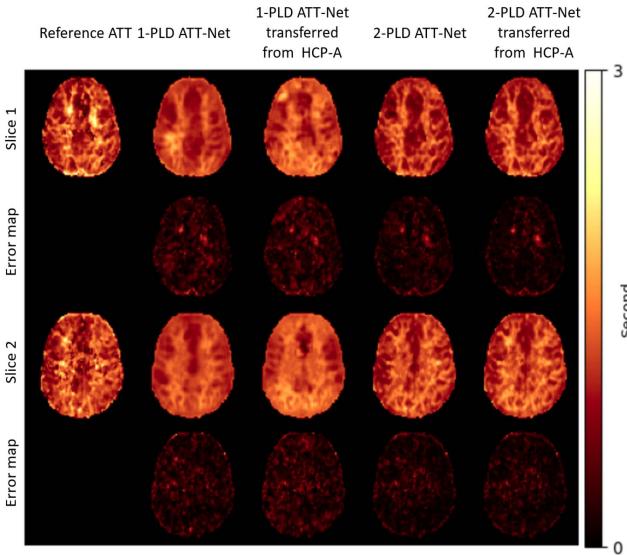


Fig. 14. Two ATT image slices of a representative HCP-D subject. The definition of the model name is the same as Fig. 13.

as the models from HCP-D. The overall error is slightly higher compared to the original models. However, it is difficult to see the decrease visually and the evaluation metrics remain at a high level. For ATT estimation, there was no significant difference between the models trained by HCP-A and HCP-D data, however, the two-PLD ATT-Nets still outperformed the one-PLD ATT-Nets in terms of visual perception and quantitative evaluation. Similar results are presented in Tables III and IV.

E. Voxel-Wise CBF/ATT vs Age Correlation Analysis

Figs. 15 and 16 show the correlation coefficient between age and CBF and ATT of each voxel. Statistical significance was

TABLE III
SSIM AND PSNR OF CBF MAPS

Method	Original HCP-D		Learned from HCP-A	
	SSIM	PSNR	SSIM	PSNR
One-PLD	0.8853±0.0668	32.57±20.93	0.8908±0.0643	32.41±18.48
Two-PLD	0.9167±0.0445	34.49±23.03	0.8098±0.0866	30.21±8.19

TABLE IV
SSIM AND PSNR OF ATT MAPS

Method	Original HCP-D		Learned from HCP-A	
	SSIM	PSNR	SSIM	PSNR
One-PLD	0.8054±0.1165	66.99±14.97	0.8095±0.1265	66.84±14.24
Two-PLD	0.8909±0.0716	69.29±15.85	0.8862±0.0698	67.80±9.71

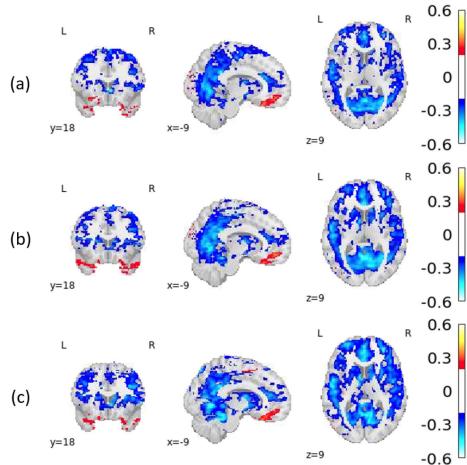


Fig. 15. Age vs CBF correlations for the reference CBF (a), and CBF estimated by the one-PLD ATT-Net (b) and two-PLD ATT-Net (c). Statistical threshold was $p < 0.05$.

defined by $p < 0.05$ ($r > 0.1966$). The two DL-models showed very similar age vs CBF and age vs ATT correlations. CBF was negatively correlated to age in most of gray matter and striatum. Positive age vs CBF correlations were found in orbito-frontal cortex and temporal cortex. ATT was positively correlated to age in most of frontal and temporal cortex, and the subcortical area but was negatively correlated to age in motor cortex, parietal cortex, and visual cortex.

V. DISCUSSION

In this study, we proposed ATT-Nets to accelerate multiple-PLD ASL MRI for measuring ATT and CBF. Compared to the full set of PLD ASL data fitting results, both the one-PLD ATT-Net and the two-PLD ATT-Net produced similar ATT and CBF estimation results. Two-PLD ATT-Net performed slightly better than one-PLD ATT-Net.

Additional experiments using two different network structures: the UNet and Resnet produced similar results to those

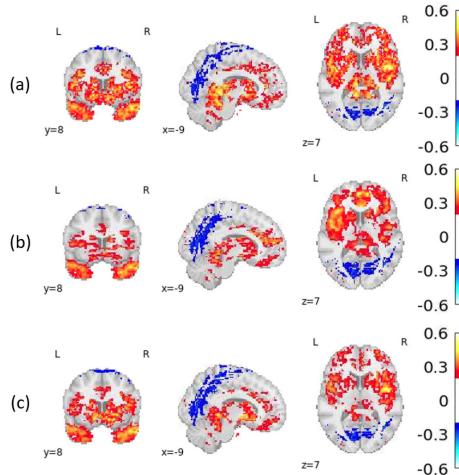


Fig. 16. Age vs ATT correlations for the reference ATT (a), and ATT estimated by the one-PLD ATT-Net (b) and two-PLD ATT-Net (c). Statistical threshold was $p < 0.05$.

reported above (the additional results were included in Supporting Information Table I), suggesting that our main findings are independent of specific network architectures and they suggest deep learning as a powerful tool for estimating CBF and ATT maps from limited number PLD ASL images.

We also conducted the experiment with two different number of measurements at each PLD: one with just a single repetition and the other with three repetitions instead of using all repetitions. The five-PLDs model was used to estimate ATT and CBF. We found that when employing fewer repetitions, the model's performance declined notably. The main reason is that the data we used was acquired with a 2D PCASL sequence without background suppression. SNR for this type of sequence is low, requiring many measurements to boost up SNR of the perfusion signal measured at different PLDs [12]. The detailed results can be found in Supporting Information as well.

One-PLD ATT-Net and two-PLD ATT-Net generated quite similar CBF images but their ATT output showed noticeable difference. Interestingly, the ATT map generated by one-PLD ATT-Net had slightly better grey matter white matter contrast while the ATT map of two-PLD ATT-Net showed higher similarity to the reference. One reason for this difference is that the reference ATT itself is quite noisy which can be improved using the state-of-art ASL sequence. HCP only acquired fewer than usual control/label image pairs for each PLD, resulting in a relatively low SNR in the perfusion weighted images and the corresponding tracer (the labeled arterial blood) kinetic curve fitting based ATT estimation. Acquiring more control/label image pairs and using background suppression can substantially boost up SNR, which however remains a future validation study as we do not have a large dataset acquired in that way. ATT in the HCP-D data showed reasonable contrast between grey matter and white matter with higher signal in white matter but lower ATT in grey matter. By contrast, the HCP-A data showed lower ATT in white matter which might be caused by the unreliable data

fitting in white matter in the aging population due to the further reduced SNR of ASL signal therein compared to young adults or children. Also, ATT in old subjects is typically longer than young subjects. It is even much prolonged in white matter. Likely most or a large portion of the labeled spins have not reached the white matter at all PLDs, which will subsequently reduce the ASL perfusion signal SNR and the reliability of the kinetic model fitting. This low reference quality can subsequently affect the model performance especially when the training sample size is limited. By using high quality reference, or increasing the training sample size, or enhancing the “learning-from-noise” capability of the network, we expect to improve the ATT estimation quality in future work.

In grey matter, the mean ROI CBF and ATT of the network output still showed high consistency to those extracted from the reference CBF and ATT images.

The domain adaptation results demonstrated the high generalizability of the ATT-Net. While we have only tested this using the HCP data, we expect to see similar generalizability for other datasets especially when the model will be fine-tuned by a few sets of new data as we recently demonstrated in the DL-based ASL denoising [11].

Generalizability is further demonstrated by the CBF/ATT versus age correlation analyses. Our results showed that output of ATT-Net had similar correlations with age to the reference ATT/CBF images, suggesting that ATT-Net preserved individual level biological variability of ATT/CBF, which is important for applications in translational or clinical research.

The network output showed some extent of image blurring, which is likely contributed by the lack of high resolution reference and the convolutional process in the networks. A non-convolutional network such as transformer might improve this issue.

We chose the WDSR network because it was successfully used in our previous CEST MRI study. In additional experiments comparing it with two other popular networks: the UNet and ResNet, we did not find a clear performance difference between those networks. We did not include these results in the main text because our focus was to prove the feasibility of estimating CBF and ATT from ASL data at fewer than usual PLDs rather than comparing different networks for estimating CBF and ATT.

VI. CONCLUSION

In conclusion, we proposed a novel network to estimate the CBF and ATT maps using fewer ASL MRI images on a large-scale dataset, which significantly accelerated the acquisition time and yield higher SNR than the conventional method. For CBF, the acquisition process can be reduced to one PLD while ATT estimation required two PLD to keep the image quality.

ACKNOWLEDGMENT

We thank Dr. Xiufeng Li for sharing his code for calculating CBF and ATT for the HCP data.

REFERENCES

- [1] D. C. Alsop et al., "Recommended implementation of arterial spin-labeled perfusion MRI for clinical applications: A consensus of the ISMRM perfusion study group and the European consortium for ASL in dementia," *Magn. Reson. Med.*, vol. 73, no. 1, pp. 102–116, 2015.
- [2] D. C. Alsop and J. A. Detre, "Reduced transit-time sensitivity in noninvasive magnetic resonance imaging of human cerebral blood flow," *J. Cereb. Blood Flow Metab.*, vol. 16, no. 6, pp. 1236–1249, 1996.
- [3] M. Günther, M. Bock, and L. R. Schad, "Arterial spin labeling in combination with a look-locker sampling strategy: Inflow turbo-sampling EPI-FAIR (ITS-FAIR)," *Magn. Reson. Med.: An Official J. Int. Soc. Magn. Reson. Med.*, vol. 46, no. 5, pp. 974–984, 2001.
- [4] J. A. Detre, H. Rao, D. J. Wang, Y. F. Chen, and Z. Wang, "Applications of arterial spin labeled MRI in the brain," *J. Magn. Reson. Imag.*, vol. 35, no. 5, pp. 1026–1037, 2012.
- [5] Y. Li et al., "Accelerating GluCEST imaging using deep learning for B0 correction," *Magn. Reson. Med.*, vol. 84, no. 4, pp. 1724–1733, 2020.
- [6] D. Shen, G. Wu, and H.-I. Suk, "Deep learning in medical image analysis," *Annu. Rev. Biomed. Eng.*, vol. 19, pp. 221–248, 2017.
- [7] D. Owen et al., "Deep convolutional filtering for spatio-temporal denoising and artifact removal in arterial spin labelling MRI," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2018, pp. 21–29.
- [8] D. Xie et al., "Denoising arterial spin labeling perfusion MRI with deep machine learning," *Magn. Reson. Imag.*, vol. 68, pp. 95–105, May 2020, doi: [10.1016/j.mri.2020.01.005](https://doi.org/10.1016/j.mri.2020.01.005).
- [9] C. Ulas, G. Tetteh, S. Kaczmarz, C. Preibisch, and B. H. Menze, "Deep-ASL: Kinetic model incorporated loss for denoising arterial spin labeled MRI via deep residual learning," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2018, pp. 30–38.
- [10] Z. Li et al., "A two-stage multi-loss super-resolution network for arterial spin labeling magnetic resonance imaging," in *Proc. Int. Conf. Med. Image Comput. Comput.-Assist. Interv.*, 2019, pp. 12–20.
- [11] L. Zhang et al., "Improving sensitivity of arterial spin labeling perfusion MRI in Alzheimer's disease using transfer learning of deep learning-based ASL denoising," *J. Magn. Reson. Imag.*, vol. 55, no. 6, pp. 1710–1722, Jun. 2022.
- [12] Z. Wang, "Arterial spin labeling perfusion MRI signal processing through traditional methods and machine learning," *Invest. Magn. Reson. Imag.*, vol. 26, no. 4, 2022, Art. no. 220.
- [13] H. Fan, P. Su, J. Huang, P. Liu, and H. Lu, "Multi-band MR fingerprinting (MRF) ASL imaging using artificial-neural-network trained with high-fidelity experimental data," *Magn. Reson. Med.*, vol. 85, no. 4, pp. 1974–1985, 2021.
- [14] Q. Zhang et al., "Deep learning-based MR fingerprinting ASL ReconStruction (DeepMARS)," *Magn. Reson. Med.*, vol. 84, no. 2, pp. 1024–1034, 2020.
- [15] D. Kim et al., "Parametric ATT and CBF mapping using a three-dimensional convolutional neural network," in *Proc. Annu. Meeting Int. Soc. Magn. Reson. Med.*, 2022, Art. no. 4904.
- [16] D. Kim et al., "A three-dimensional convolutional neural network in ASL with reduced number of inversion times or averages," in *Proc. ISMRM Workshop Perfusion MRI: From Head Toe*, 2022.
- [17] N. J. Luciw, Z. Shirzadi, S. E. Black, M. Goubran, and B. J. MacIntosh, "Automated generation of cerebral blood flow and arterial transit time maps from multiple delay arterial spin-labeled MRI," *Magn. Reson. Med.*, vol. 88, no. 1, pp. 406–417, 2022.
- [18] S. Y. Bookheimer et al., "The lifespan human connectome project in aging: An overview," *Neuroimage*, vol. 185, pp. 335–348, 2019.
- [19] M. P. Harms et al., "Extending the human connectome project across ages: Imaging protocols for the lifespan development and aging projects," *Neuroimage*, vol. 183, pp. 972–984, 2018.
- [20] M. A. Chappell, A. R. Groves, B. Whitcher, and M. W. Woolrich, "Variational Bayesian inference for a nonlinear forward model," *IEEE Trans. Signal Process.*, vol. 57, no. 1, pp. 223–236, Jan. 2009.
- [21] Z. Wang et al., "Empirical optimization of ASL data analysis using an ASL data processing toolbox: ASLtbx," *Magn. Reson. Imag.*, vol. 26, no. 2, pp. 261–269, 2008.
- [22] Y. Li, S. Dolui, D.-F. Xie, and Z. Wang, for the Alzheimer's Disease Neuroimaging Initiative, "Priors-guided slice-wise adaptive outlier cleaning for arterial spin labeling perfusion MRI," *J. Neurosci. Methods*, vol. 307, pp. 248–253, 2018.
- [23] R. B. Buxton, L. R. Frank, E. C. Wong, B. Siewert, S. Warach, and R. R. Edelman, "A general kinetic model for quantitative perfusion imaging with arterial spin labeling," *Magn. Reson. Med.*, vol. 40, no. 3, pp. 383–396, 1998.
- [24] W. Dai, P. M. Robson, A. Shankaranarayanan, and D. C. Alsop, "Reduced resolution transit delay prescan for quantitative continuous arterial spin labeling perfusion imaging," *Magn. Reson. Med.*, vol. 67, no. 5, pp. 1252–1265, 2012.
- [25] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2016, pp. 770–778.
- [26] V. Nair and G. E. Hinton, "Rectified linear units improve restricted boltzmann machines," in *Proc. 27th Int. Conf. Int. Conf. Mach. Learn.*, 2010, pp. 807–814.
- [27] J. Yu et al., "Wide activation for efficient and accurate image super-resolution," 2018, *arXiv:1808.08718*.
- [28] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "Fsl," *Neuroimage*, vol. 62, no. 2, pp. 782–790, 2012.
- [29] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," 2014, *arXiv:1412.6980*.