Biomedical Physics &
Engineering Express

Endorsed by IPEM Institute of Physics and Engineering in Medicine
PURPOSE-LED PUBLISHING™

PAPER • **OPEN ACCESS**

A deep learning approach for quantifying CT perfusion parameters in stroke

To cite this article: Wanning Zeng *et al* 2025 *Biomed. Phys. Eng. Express* **11** 035015

Biomedical Physics & Engineering Express

PAPER

A deep learning approach for quantifying CT perfusion parameters in stroke

Wanning Zeng[1], Yang Li[2], Jeff L Zhang[1], Tong Chen[2], Ke Wu[2] and Xiaopeng Zong[1,3,4]

[1] School of Biomedical Engineering, ShanghaiTech University, Shanghai, People's Republic of China
[2] United Imaging Healthcare Group, Shanghai, People's Republic of China
[3] State Key Laboratory of Advanced Medical Materials and Devices, ShanghaiTech University, Shanghai, People's Republic of China
[4] Shanghai Clinical Research and Trial Center, Shanghai, People's Republic of China

E-mail: zongxp@shanghaitech.edu.cn

Keywords: CT perfusion imaging, acute ischemic stroke, deep learning

OPEN ACCESS

RECEIVED
16 December 2024

REVISED
23 March 2025

ACCEPTED FOR PUBLICATION
7 April 2025

PUBLISHED
16 April 2025

Abstract

Objective. Computed tomography perfusion (CTP) imaging is widely used for assessing acute ischemic stroke. However, conventional methods for quantifying CTP images, such as singular value decomposition (SVD), often lead to oscillations in the estimated residue function and underestimation of tissue perfusion. In addition, the use of global arterial input function (AIF) potentially leads to erroneous parameter estimates. We aim to develop a method for accurately estimating physiological parameters from CTP images. *Approach*. We introduced a Transformer-based network to learn voxel-wise temporal features of CTP images. With global AIF and concentration time curve (CTC) of brain tissue as inputs, the network estimated local AIF and flow-scaled residue function. The derived parameters, including cerebral blood flow (CBF) and bolus arrival delay (BAD), were validated on both simulated and patient data (ISLES18 dataset), and were compared with multiple SVD-based methods, including standard SVD (sSVD), block-circulant SVD (cSVD) and oscillation-index SVD (oSVD). *Main results*. On data simulating multiple scenarios, local AIF estimated by the proposed method correlated with true AIF with a coefficient of 0.97 ± 0.04 ($P < 0.001$), estimated CBF with a mean error of 4.95 ml/100 g min^{-1}, and estimated BAD with a mean error of 0.51 s; the latter two errors were significantly lower than those of the SVD-based methods ($P < 0.001$). The CBF estimated by the SVD-based methods were underestimated by 10% ~ 15%. For patient data, the CBF estimates of the proposed method were significantly higher than those of the sSVD method in both normally perfused and ischemic tissues, by 13.83 ml/100 g min^{-1} or 39.33% and 8.55 ml/100 g min^{-1} or 57.73% ($P < 0.001$), respectively, which was in agreement with the simulation results. *Significance*. The proposed method is capable of estimating local AIF and perfusion parameters from CTP images with high accuracy, potentially improving CTP's performance and efficiency in diagnosing and treating ischemic stroke.

1. Introduction

Stroke is a leading cause of death worldwide. As the most common type of stroke (GBD 2019 Stroke Collaborators 2021, Pu *et al* 2023), ischemic stroke is characterized by arterial occlusion, which results in insufficient blood supply to certain downstream tissues of the brain. One important goal in the treatment of ischemic stroke is the timely recanalization of the occluded vessel, thereby preventing the salvageable ischemic penumbra from further developing into irreversible infarct core. Therefore, rapid and accurate assessment of the viability of potentially affected tissue is needed for optimal management of stroke patients.

Computed tomography perfusion (CTP) imaging has been widely used in clinical practice to reveal tissue hypoperfusion and thus assess the salvage potential of tissues in patients with ischemic stroke (Cereda *et al* 2016, de la Rosa *et al* 2021, Hakim *et al* 2021, Gava *et al* 2023, Luo *et al* 2024). In CTP image analysis, physiological parameters such as cerebral blood flow (CBF),

cerebral blood volume (CBV), mean transit time (MTT), and bolus arrival delay (BAD) are derived and used to determine whether the tissue is normal, hypoperfused or infarcted. The parameter derivation is based on the indicator dilution theory that the tracer concentration time curve (CTC) in tissue is the convolution of the arterial input function (AIF) and the flow-scaled residue function (Fieselmann *et al* 2011, Østergaard *et al* 1996). The residue function fully characterizes the microvasculature and vascular dynamics of tissue.

Methods for determining residue function from CTC generally fall into two categories: model-based and model-free (Khalifa *et al* 2014, Sourbron and Buckley 2013). In model-based methods, a mathematical model with a specific shape is predefined for the residue function, and representative examples include the Tofts model, the two-compartment exchange (2CXM) model and the adiabatic approximation of tissue homogeneity (AATH) model. Predefined shapes for residue function can simplify the computation process, but underlying assumptions of the model may not be applicable to all tissue types or pathological conditions. A popular model-free method is singular value decomposition (SVD) (Østergaard *et al* 1996, Wu *et al* 2003) for deconvolution. However, deconvolution is an inverse problem that is extremely sensitive to noise, which means that the residue function generated by the SVD often displays severe oscillations that are not physiologically meaningful. The CBF determined from such residue function may be underestimated, as observed in several studies (Fieselmann *et al* 2011, Mouridsen *et al* 2006, Shinohara *et al* 2010).

Recently, deep learning methods such as bi-CNN (Fang *et al* 2021, Ulas *et al* 2019), U-Net (Gava *et al* 2023), LSTM (Zou *et al* 2020), GRU (Ottens *et al* 2022), attention block (Zeng and Zhou 2021), PINN (van Herten *et al* 2022) and GAN (Huang *et al* 2023, Oh *et al* 2024) have shown great promise in quantifying parameters from dynamic images. However, CNN-based networks, while effective in capturing local features through convolutional kernels, struggle with capturing long-term dependencies and require fixed-length inputs, often necessitating padding or truncation that may lead to information loss or noise introduction. RNN-based networks are susceptible to gradient vanishing or explosion during training, and their sequential processing nature prohibits parallel computation. PINNs face challenges in simultaneously optimizing data fitting and physical constraint terms, resulting in high computational complexity and convergence difficulties. In GAN-based parameter estimation tasks, the adversarial training process may prioritize sample diversity over estimation accuracy, potentially leading to biased parameter estimates and unreliable results. Some of these methods use parameter maps obtained by conventional model-based methods as the ground truth for network training, and are thus essentially similar to model-based 'parameter mapping'. Other methods involve unsupervised learning that minimizes the difference between the reconstructed CTC and the measured one, but still tend to restrict the residue function to a specific family of shapes, and thus may limit the applicability of these methods.

Another challenge in quantifying CTP images is the accurate sampling of AIF (Calamante 2013). All of the methods described above use AIF sampled in a major artery (denoted as the global AIF). Using such AIF for analyzing signals of all tissues in the field of view, one ignores the potentially large difference in delay or dispersion of the tracer as it transits from the sampling point to the tissue of interest. In the case of acute ischemic stroke, vascular occlusion significantly delays and disperses tracer transit, and thus vasodynamics can be quite heterogeneous within the region of interest (Lorenz *et al* 2006).

In this study, we proposed a Transformer-based deep learning approach for estimating physiological parameters from CTP images. We chose the Transformer architecture due to its proven effectiveness in capturing long-range dependencies in sequences (Shamshad *et al* 2023, Vaswani *et al* 2017) , which enables native handling of variable-length inputs without the need for padding or truncation. Unlike traditional methods, our approach eliminated the requirement for pre-specifying a pharmacokinetic model to constrain the residue function to a specific family of shapes, thereby offering greater flexibility and generality. Inspired by recent work (Bae *et al* 2023) that utilized Vision Transformer for local AIF estimation in dynamic contrast enhanced MR images, we included the global AIF as a second input, as it contains information about arterial inputs, and applied non-negative, monotonically non-increasing and smoothness constraints on the shape of the residue function. For validation, we compared the proposed method with the SVD-based methods on simulated data and a public patient dataset of stroke, ISLES18. We hypothesized that the proposed method could estimate local AIF and quantify tissue perfusion parameters with higher accuracy than the SVD-based methods.

2. Materials and methods

In this section, we first briefly introduce the tracer kinetic theory and how to quantify physiological parameters from CTP images. The proposed method is then presented, including the network architecture and its training specifications. Lastly, we describe the data used for validating the method, including the simulated data and the public patient dataset of stroke, ISLES18, as well as the SVD-based methods used for comparison.

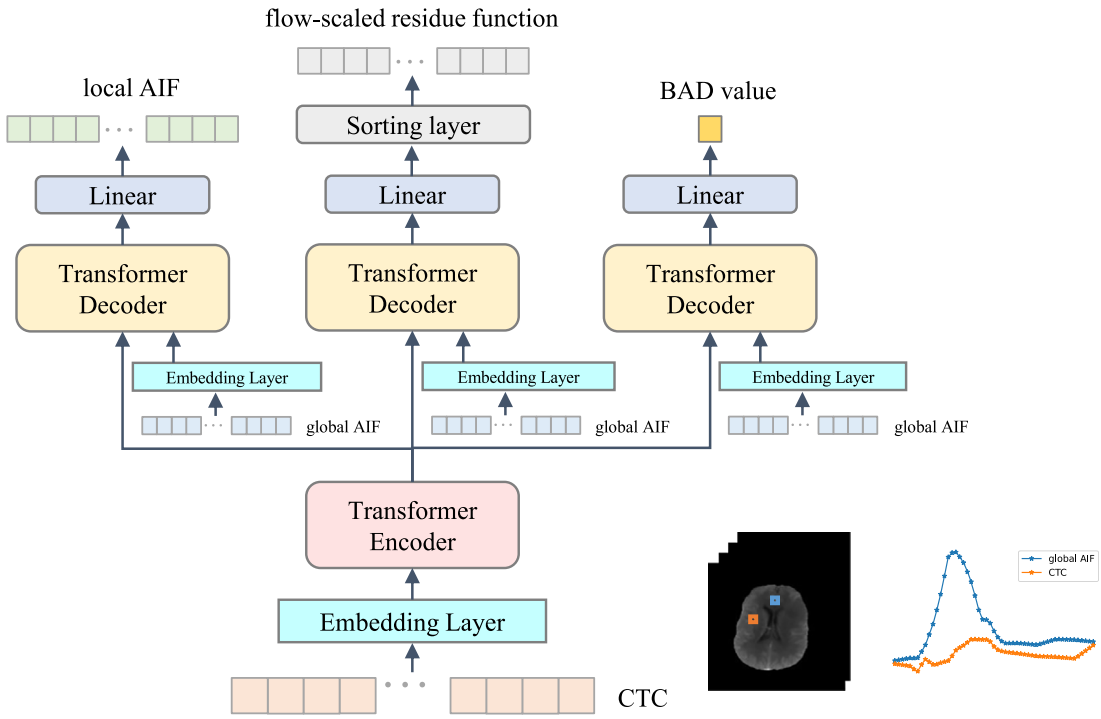

Figure 1. The proposed network consists of a shared encoder and three task-specific decoders. One CTC measured from the voxel of interest is input to the encoder, and the global AIF, either assumed or sampled from a large artery, is input to the three decoders. The embedding layer converts the input data to the Transformer dimension. The network outputs include local AIF, flow-scaled residue function and BAD value. Lower right are examples of global AIF and CTC measurements from a CTP image and the corresponding ROIs from which they were extracted.

2.1. Tracer kinetic theory

CTP imaging is dynamic CT scans performed over a period of time following intravenous injection of an iodinated tracer. It can be assumed that the tracer concentration per tissue voxel is proportional to the measured x-ray attenuation value (Fieselmann *et al* 2011). According to the indicator dilution theory (Østergaard *et al* 1996), if the blood–brain barrier is intact, the measured CTC, $C(t)$, is the convolution of the supplying AIF, $C_a(t)$, and the unknown flow-scaled residue function $R(t)$:

$$\kappa C(t) = C_a(t) \otimes R(t) = \mathrm{CBF} \cdot \int_0^t C_a(\tau) r(t-\tau) d\tau \tag{1}$$

where the residue function $r(t)$ characterizes the local vascular structure of the tissue of interest. The constant κ is usually set to 0.73, reflecting the average hematocrit and density of tissue (Fieselmann *et al* 2011, Mouridsen *et al* 2006).

For solving equation (1), the convolution integral is calculated using the trapezoidal approximation (Boutelier *et al* 2012, Koh *et al* 2011). Several physiologically meaningful parameters can be derived from the residue function $R(t)$. For example, CBF is computed as the maximum value of $R(t)$, CBV parameter as the ratio of integral of CTC over that of AIF over time, and MTT as ratio of CBV over CBF according to the central volume theorem. The parameter BAD is defined as the time required for arterial blood to reach the voxel of interest. These parameters are helpful in identifying abnormalities in brain tissue. It is found that the ischemic brain tissue in stroke patients usually shows significant decrease in CBF and CBV, and significant increase in MTT and BAD (de la Rosa *et al* 2021, Fieselmann *et al* 2011, Luo *et al* 2024).

In practice, the global AIF is usually sampled in a large artery due to limitations such as spatial resolution. However, potential delay and dispersion effects may occur during the transit from the AIF measurement point to the tissue of interest, introducing large errors in parameter quantification. The use of local AIF, which is theoretically closer to arterial inlet of the tissue voxel of interest, can mitigate this issue (Calamante 2013).

2.2. A transformer-based network for parameter estimation

In the following, we present a Transformer-based network for estimating flow-scaled residue function and local AIF from tissue's CTC and global AIF. In addition, the BAD is estimated as a separate output of the network, but not as part of estimated residue function or local AIF, in order to focus the network training on learning the complex shape features of the functions. As shown in figure 1, the network consists of a shared Transformer encoder and three task-specific decoders. The Transformer block extracts temporal features of the inputs by capturing the dependencies between values at different time points.

Similar to the original design (Vaswani *et al* 2017), our Transformer block consists of 3 identical layers, with the model dimension of 64 and the number of attention heads of 16. Each input data consists of a CTC and a global AIF, each being a one-dimensional vector. The CTC is fed to the encoder, and the global AIF to the decoder. Inspired by the previous approach for time series processing (Zhou *et al* 2021), the embedding layers consists of three components: data projection via one-dimensional convolutional filters, fixed position embedding and learnable temporal embedding. After each decoder, two linear layers are used to reshape the output from the Transformer block to the dimension of the final output. For physiological validity, a Softplus activation function is applied to all the outputs to ensure non-negativity, and a descending-sorting operation is applied to the residue function to maintain its monotonically non-increasing characteristics.

For network training, we used a loss function comprising four components: mean squared error (MSE) between the true BAD parameter δ and its estimate $\hat{\delta}$, MSE between the reconstructed CTC $\hat{C}(t)$ and the input CTC $C(t)$, MSE between the estimated local AIF $\hat{C}_{\text{a\_local}}(t)$ and its truth function $C_{\text{a\_local}}(t)$, and a regularization term that ensures smoothness of the flow-scaled residue function $R(t)$ (Boutelier *et al* 2012, Luo *et al* 2024). The latter two components are unsupervised. The loss function $L(w)$ is a weighted sum of the four components:

$$L(w) = \|\delta - \hat{\delta}\|_2^2 + \lambda_1 \|C_{\text{a\_local}}(t) - \hat{C}_{\text{a\_local}}(t)\|_2^2 + \lambda_2 \|C(t) - \hat{C}(t)\|_2^2 + \lambda_3 \left\| \frac{d^2R(t)}{dt^2} \right\|^2 \quad (2)$$

where λ_1, λ_2, and λ_3 denote the hyper-parameters that balance each loss component, with their relative contributions to the loss equal to approximately 1:1:7:1, respectively, to prioritize the CTC reconstruction error while maintaining a balanced contribution from the other terms.

For model training on the simulated dataset, we used the Adam optimizer to minimize the loss function with a batch size of 1024 over 100 epochs. The dataset was divided into training and validation sets at a ratio of 0.8/0.2. With a warm-up learning rate tuning strategy, we gradually increased the learning rate to 3e-5 in the first 10% of epochs, and then decreased it in cosine annealing in the following epochs. During the fine-tuning phase of the patient dataset, the network architecture and hyper-parameter settings remained unchanged. All weights were frozen except for the residue function estimation decoder, and the third and fourth loss components in equation (2) were considered. The entire network was implemented in Python 3.8.12 using Pytorch 1.12.0 and executed on a single NVIDIA A40 GPU (45 Gb).

2.3. Simulated and patient data for network training and testing

We used simulated data and a real patient dataset to validate the feasibility of the proposed method. In the following we first describe the process of generating the simulated data, as shown in figure 2.

First, since factors such as different injection schemes and the patient's cardiac activity can lead to different AIF shapes, we used two functions to simulate possible flow dynamics. Specifically, the gamma-variate function (Mouridsen *et al* 2006, Wu *et al* 2003) simulates the dynamics of the tracer's 'first pass' through the circulatory system, while the population-based AIF proposed by Parker *et al* (Parker *et al* 2006) additionally models the 'recirculation' phenomenon that may occur after the 'first pass'.

$$C_{\text{a\_gamma}}(t) = at^b e^{-\frac{t}{c}} \quad (3)$$

$$C_{\text{a\_parker}}(t) = \sum_{n=1}^{2} \frac{A_n}{\sigma_n\sqrt{2\pi}} e^{-(t-T_n)^2/2{\sigma_n}^2} + \frac{de^{-kt}}{1+e^{-s(t-\tau 1)}} \quad (4)$$

The parameter values of equations (3) and (4) were initially determined based on standard injection schemes described in the literature (Parker *et al* 2006, Wu *et al* 2003) to generate global AIFs with typical shapes and sizes. To more realistically simulate physiological variability, the AIF curves were scaled along both the time and value axes using uniformly distributed factors ranging from 70% to 130%.

The flow-scaled residue function $R(t)$ was derived from the integral of the transit time distribution $h(\tau)$ (Mouridsen *et al* 2006), shown in equation (5), which is modeled by a family of gamma distributions and consists of two parameters: the shape parameter α and the scale parameter β.

$$\begin{aligned} R(t) &= \text{CBF}(1 - \int_0^t h(\tau)d\tau) \\ &= \text{CBF}\cdot(1 - \int_0^t \frac{1}{\beta^\alpha \Gamma(\alpha)} \tau^{\alpha-1} e^{-\tau/\beta} d\tau) \end{aligned} \quad (5)$$

The residue function is approximately exponentially shaped for $\alpha = 1$ and box-car shaped for $\alpha = 100$. By definition, the MTT is the first moment of $h(\tau)$, which is equal to $\alpha\beta$.

Delay and dispersion effects occur as the tracer transits from a large artery, where global AIF is typically sampled, to a local capillary near the tissue of interest. The delay effect was modeled by shifting the global AIF by a certain time interval δ, and the dispersion effect was modeled by convolving the global AIF with the vascular transport function (VTF). Here, we used the exponential dispersion kernel (Calamante *et al* 2000, Pizzolato *et al* 2017, Sourbron and Buckley 2013) for VTF.

$$\text{VTF}(t) = \frac{e^{-t/T_A}}{T_A} \quad (6)$$

where T_A is the mean arterial transit time, i.e. the average time it takes to travel from the upstream site

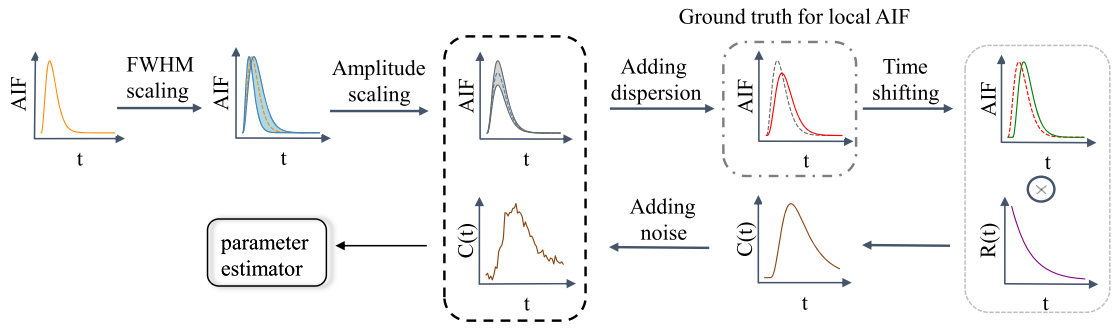

Figure 2. Simulated data generation process. Starting with an AIF with a typical shape and size, the possible flow dynamics are simulated through the steps of FWHM (full width at half maximum) scaling, amplitude scaling, dispersion addition and time shifting. The residue function is modeled using a family of gamma distributions, and different parameter settings result in different shapes of the residue function. The noisy CTC is generated after convolving the generated AIF and residue function as well as adding Gaussian noise.

where the AIF is measured to the local inlet of the voxel of interest. As the value of T_A increases, the dispersion effect becomes more pronounced, manifested by a widening of the peak width and a decrease in the peak amplitude. Considering the effects of delay and dispersion, equation (1) can be expressed as:

$$C(t) = (C_a(t - \delta) \otimes \mathrm{VTF}(t)) \otimes R(t) \qquad (7)$$

In order to simulate the inherent noise present in real patient data, zero-mean Gaussian noise was added to the simulated CTCs, resulting in noisy CTCs with different signal-to-noise ratios (SNR) (Brahma *et al* 2024, Welvaert and Rosseel 2013).

Following the process in figure 2, we simulated a training dataset consisting of 500,000 curves. During the convolution process, the curves were first generated at a small time interval to minimize the discretization error. Subsequently, the curves were resampled to the required temporal resolution, which was uniformly selected from 1 to 2 s, and the total number of time points is from 45 to 120. The parameter values for the simulation were uniformly sampled from physiologically realistic ranges in previous literature (Mouridsen *et al* 2006, 2014, Pizzolato *et al* 2017): SNR∈[1,30] dB, CBF∈[0,200] m1/100 g min^{-1}, MTT∈[3.43,24] s, α∈[0.5,100], T_A∈[1,4] s, δ∈[0,25] s. As shown in table 1, several simulated test datasets were constructed to evaluate the performance of the methods under different noise levels, delay levels, dispersion levels, sampling temporal resolutions, AIF types, and residue function shapes, with 10,000 curves generated independently for each variable. The CBV was set to 4% to simulate normal gray matter.

Our method was validated using a public patient dataset from the Ischemic Stroke Lesion Segmentation (ISLES) 2018 challenge (Cereda *et al* 2016, Hakim *et al* 2021). This dataset contains imaging data of acute ischemic stroke patients from multiple scanners and centers. We obtained the data directly through the ISLES challenge website (http://www.isles-challenge.org/). The dataset consists of 156 CTP acquisitions from 103 stroke patients, with the data divided into a training set (94 CTP volumes from 63 patients) and a test set (62 CTP volumes from 40 patients). Notably, the open database only provides binary-form infarct core segmentation masks (delineated in diffusion weighted imaging (DWI) images) for the training set. All images were motion-corrected, co-registered with DWI, and spatiotemporally resampled (resulting in a matrix size of 256 × 256 voxels with an in-plane resolution of 0.79–1.09 mm, slice thickness between 4.0 mm and 12.0 mm, and a temporal resolution of 1 s).

For each CTP volume in the patient dataset, in addition to acquiring a global AIF, it is necessary to acquire the venous output function (VOF) to compensate for the partial volume effect in AIF (de la Rosa *et al* 2021), which recalibrates the AIF peak by equalizing the area under the two curves, as illustrated in figure 3. Additionally, a whole-brain segmentation mask, a vascular region mask, and a cerebrospinal fluid (CSF) mask were generated. These pre-processing steps were facilitated by methods integrated in uOmnispace. CT (United-Imaging, Shanghai, China), an FDA-approved software package for CTP analysis. Guided image filtering (He *et al* 2013) was used as an additional step to improve the image quality, with the radius set to 2, the regularization parameter set to 0.01, and the time-averaged image of CTP volume selected as the guidance image. In addition, clinical CTP imaging is influenced by various factors, including differences in imaging devices, patient-specific pathologies, variations in imaging parameters, and artifacts (e.g., motion or metal artifacts). These factors contribute to more complex noise distribution and greater CTC variability in patient data compared to simulated data. For instance, the basic assumption for modeling simulated data is that the noise follows a Gaussian distribution and the residue function is characterized by a family of gamma distributions. This simplification may differ from the complex distributions in real CTP images (Li *et al* 2022, Luo *et al* 2024). Therefore, to improve the generality of the proposed network, the network was supplemented and fine-tuned using a

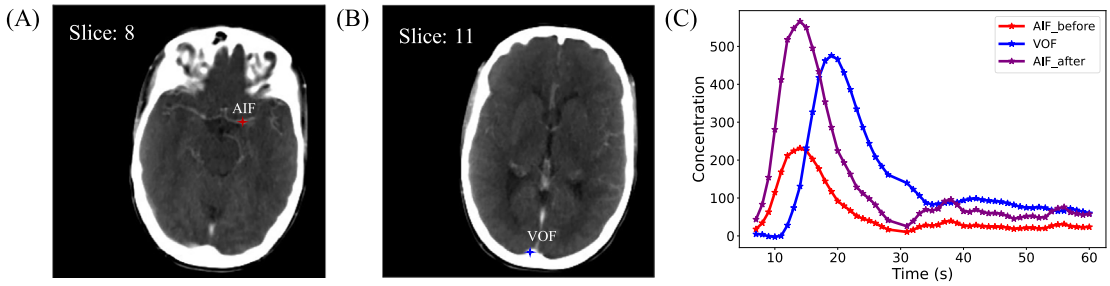

Figure 3. Illustration of VOF recalibration. (A) The AIF was obtained from the middle cerebral artery. (B) The VOF was extracted from the sagittal sinus, where the partial volumetric effects are considered negligible. (C) Amplitude correction of AIF using VOF, based on the assumption that the area under the AIF and VOF curves are equal.

small number of real data samples before applying the training weights from the simulated data to the patient dataset. Given that the ISLES challenge only provides binary-form DWI infarct core segmentation for some of the data, we used the original training set (a total of 94 cases) for real-world data test, and the data from the original test set for fine-tuning (a total of 62 cases).

2.4. Comparative techniques

The proposed method was compared with several commonly used deconvolution methods based on singular value decomposition (SVD), including the standard SVD (sSVD), the block-circulant SVD (cSVD) and the oscillation index SVD (oSVD). The main idea of the sSVD method is to decompose the AIF matrix into two orthogonal matrices and one diagonal matrix, thereby obtaining the solution of residue function (Fieselmann *et al* 2011, Ho *et al* 2016). In order to regularize the solution to reduce oscillations and obtain physiologically sound results, it is necessary to use a certain threshold to suppress the effect of small singular values (Mouridsen *et al* 2006, Østergaard *et al* 1996). The cSVD and oSVD methods both use a block-circulant AIF matrix and have been shown to provide flow estimates that are insensitive to BAD. The sSVD and cSVD both use a global threshold for regularization, while oSVD iteratively adjusts the threshold for each spatial voxel until the calculated oscillations index falls below a predefined one. Specifically, we used a regularization threshold of 20% for sSVD (Østergaard *et al* 1996), a threshold of 10% for cSVD, and set the oscillations index for oSVD to 0.035 (Wu *et al* 2003).

2.5. Statistical analysis

In the simulated dataset, we first evaluated the accuracy of the local AIF estimates by calculating the Pearson correlation coefficient (r) between the estimated local AIF and the ground truth. In addition, we assessed key characteristics of AIF, including magnitude of first-pass peak (M_{peak}), time to the peak (T_{peak}), and full width at half maximum (FWHM) of the peak, and calculated the difference of these parameters between the estimated AIF and the simulated true AIF. To evaluate the accuracy of parameter estimates (CBF and BAD) in the simulated datasets, we calculated the root mean square error (RMSE) of the parameter estimates by the different methods with respect to the simulated true parameter values for all simulated scenarios. These accuracy measures were plotted as a function of noise, delay, dispersion, temporal resolution, AIF type and residue function shape. In addition, we measured the consistency using the ratio between the estimated and true CBF, which is independent of the size of the true flow. Since the real-world data lack a gold standard for tissue perfusion estimation, in order to quantitatively explore the differences in estimated CBF between the different methods in the patient dataset (containing 94 CTP images), we calculated the CBF values for each method in normally perfused regions and ischemic regions (determined by DWI masks) respectively. Then we compared the deviation proportions of the mean CBF estimates of each method relative to the sSVD method in different brain regions. Paired Wilcoxon-test was performed, and the significance level was set to $\alpha = 0.05$.

3. Results

3.1. Validation with simulation

On simulated test datasets, a good correlation between the estimated local AIF and the ground truth was achieved ($r = 0.97 \pm 0.04$, $P < 0.001$). In addition, the results for the metrics used to evaluate the AIF characteristics were as follows: M_{peak} error of 0.02 ± 0.09, T_{peak} error of 0.12 ± 1.03 s, FWHM error of -0.01 ± 0.85 s. The errors of T_{peak} and FWHM were below the minimum 1 s temporal resolution set for the simulated dataset.

Figures 4, 5 illustrated the comparative performance of the four perfusion parameter estimation methods under different scenarios. The results indicated that the average estimation error of the proposed method was significantly lower than that of the SVD-based methods across six simulated test datasets ($P < 0.001$). Specifically, the average RMSE of CBF estimation for the network, sSVD, cSVD, and oSVD were 4.95 ml/100 g min^{-1}, 7.95 ml/100 g min^{-1},

Table 1. Generation of simulated test datasets. The main variables of each dataset are highlighted in bold, including different noise levels, delay levels, dispersion levels, sampling temporal resolutions, AIF types, and residue function shapes. $C_a(t)$: global AIF, $C_{a\_parker}(t)$: population-based AIF, $C_{a\_gamma}(t)$: gamma-variate AIF, CBV: cerebral blood volume, MTT: mean transit time, α: shape parameter controlling the residue function, δ: bolus arrival delay (BAD), Δt: temporal resolution, T: total duration, T_A: mean arterial transit time, SNR: signal-to-noise ratio.

Dataset	Setting
#1	$C_a(t)\in[C_{a\_parker}(t), C_{a\_gamma}(t)]$, CBV = 4%, MTT∈[3.43,24] s, $\alpha = 1$, $\delta\in[0,25]$ s, Δt = 1 s, T = 80 s, $T_A = 0$ s, **SNR = [5,7,10,15,20,25,30] dB**.
#2	$C_a(t)\in[C_{a\_parker}(t), C_{a\_gamma}(t)]$, CBV = 4%, MTT∈[3.43,24] s, $\alpha = 1$, $\boldsymbol{\delta = [0,1,2,3,4,5,6,7,8,9,10]}$ **s**, Δt = 1 s, T = 80 s, $T_A = 0$ s, SNR = 20 dB.
#3	$C_a(t)\in[C_{a\_parker}(t), C_{a\_gamma}(t)]$, CBV = 4%, MTT∈[3.43,24] s, $\alpha = 1$, $\delta\in[0,25]$ s, Δt = 1 s, T = 80 s, $\boldsymbol{T_A = [0,1,2,3,4,5]}$ **s**, SNR = 20 dB.
#4	$C_a(t)\in[C_{a\_parker}(t), C_{a\_gamma}(t)]$, CBV = 4%, MTT∈[3.43,24] s, $\alpha = 1$, $\delta\in[0,25]$ s, $\boldsymbol{\Delta t = [1,1.1,1.3,1.6,2]}$ **s**, T = 80 s, $T_A = 0$ s, SNR = 20 dB.
#5	$\boldsymbol{C_a(t) = [C_{a\_parker}(t), C_{a\_gamma}(t)]}$, CBV = 4%, MTT∈[3.43,24] s, $\alpha = 1$, $\delta\in[0,25]$ s, Δt = 1 s, T = 80 s, $T_A = 0$ s, SNR = 20 dB.
#6	$C_a(t)\in[C_{a\_parker}(t), C_{a\_gamma}(t)]$, CBV = 4%, MTT∈[3.43,24] s, $\boldsymbol{\alpha = [1,100]}$, $\delta\in[0,25]$ s, Δt = 1 s, T = 80 s, $T_A = 0$ s, SNR = 20 dB.

7.18 ml/100 g min^{-1}, 7.03 ml/100 g min^{-1} respectively; and the average RMSE of BAD estimation for the network, sSVD, cSVD, and oSVD were 0.51 s, 3.07 s, 3.27 s, and 2.82 s respectively. In addition, the average CBF ratios of each method were 0.95 ± 0.15 for the network, 0.85 ± 0.13 for sSVD, 0.90 ± 0.15 for cSVD, and 0.85 ± 0.11 for oSVD.

The results showed that the accuracy of parameter estimation for all methods was affected by different noise levels, where the estimation error decreased as the SNR increased, and stabilized for most of the methods except oSVD when the SNR exceeded 20 dB. Furthermore, the presence or absence of BAD had a great impact on the parameter estimation accuracy of the SVD-based methods, while the network exhibited a more stable estimation error variation. The enhanced dispersion effect led to an increase of parameter estimation error for all methods, while the network consistently maintained a relatively low error level. Moreover, the performance of all methods improved as the temporal resolution increased, with oSVD being more sensitive to this factor compared to sSVD and cSVD. In addition, the network was less susceptible to changes in the AIF assumptions than the SVD-based methods, and the parameter estimation error for the gamma AIF was generally lower than that for the parker AIF. Besides, the SVD-based methods displayed notably higher variation than the network in data with two different residue function shapes: exponential ($\alpha = 1$) and box-car ($\alpha = 100$), which indicated that the network exhibited greater robustness when confronted with data under different residue function shape assumptions.

3.2. Evaluation with patient data

To compare the residue function estimates of the different methods, figures 6, 7 illustrated an example of the estimation of the flow-scaled residue function and the reconstructed CTC within the brain ROIs using different methods. We first placed four 5 × 5 regions of interest (ROIs) in normal and ischemic regions (as determined by the DWI masks from ISLES18 dataset) of white matter and gray matter, respectively, then calculated the average of the estimated curves within these ROIs for each method. As depicted in figure 6, the reconstructed CTCs from our proposed network closely matched the observed CTCs. The local AIF estimated in the ischemic region exhibited a wider peak with a lower peak amplitude compared with the normally perfused region. In addition, the estimated residue function decayed more slowly in ischemic regions than in normal regions. And the results indicated that the CBF was higher in gray matter than in white matter, and higher in normally perfused regions than in ischemic regions. Figure 7 showed that there were significant oscillations in the residue function estimated by the SVD-based methods, alternating between positive and negative values in the late phase, in contrast to the smooth and monotonically non-increasing residue function predicted by the network. And these oscillations were more pronounced in the ischemic region than in the normally perfused region, especially when using the cSVD and oSVD methods. In addition, the network-estimated CBF values were higher than those estimated by the SVD-based methods.

Figure 8 showed examples of physiological parameter maps obtained by sSVD, cSVD, oSVD, and the proposed network, demonstrating the application on two different patients with ischemic stroke due to occlusion of the left middle cerebral artery. The generated perfusion parameter maps were helpful in differentiating between ischemic and normal tissues, with ischemic regions showing reduced CBF, prolonged MTT, and increased BAD compared to normally perfused regions. In addition, the network estimated higher CBF values, shorter MTT values, and shorter BAD values compared to the SVD-based estimates.

Table 2 showed that the CBF estimates by the network was significantly higher than those by the SVD-based methods ($P < 0.001$). Specifically, in the normally perfused region, the CBF estimates of cSVD,

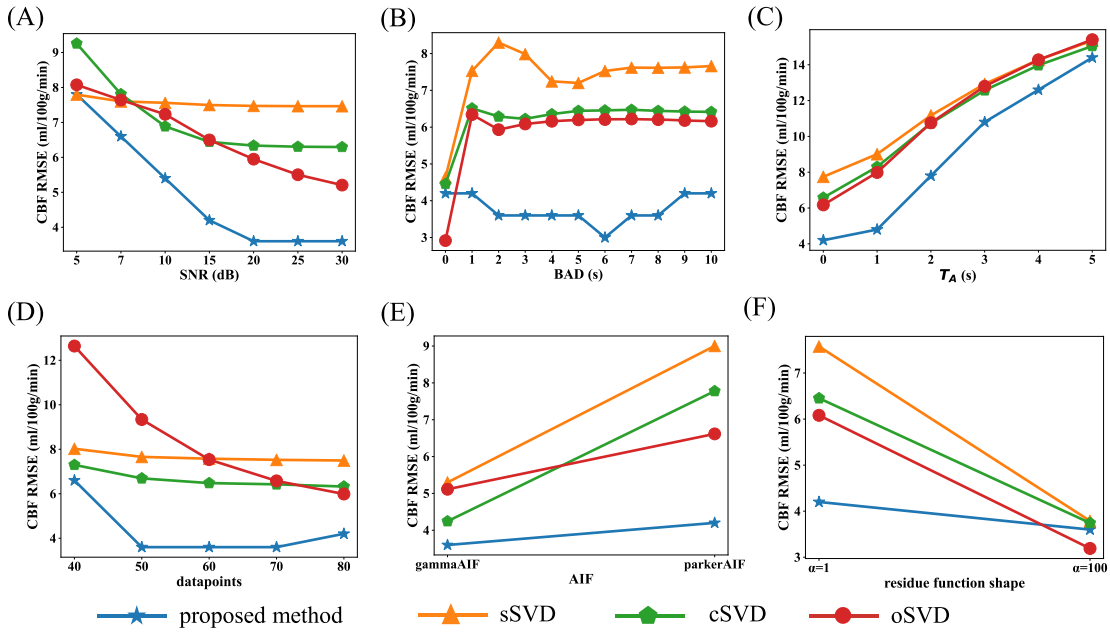

Figure 4. Root-mean-square errors between CBF estimates from different methods and the ground truth physiological parameters. The errors were plotted as a function of (A) noise, (B) delay, (C) dispersion, (D) temporal resolution, (E) AIF type and (F) residue function shape.

oSVD, and the proposed network were 1.32 ml/100 g/min lower (−3.75%), 0.41 ml/100 g/min lower (−1.17%), and 13.83 ml/100 g/min higher (39.33%) than that of sSVD, respectively. Additionally, in the ischemic region, the CBF estimates of cSVD, oSVD, and the proposed network were 1.16 ml/100 g/min higher (7.83%), 0.3 ml/100 g/min lower (−2.03%), and 8.55 ml/100 g/min higher (57.73%) than that of sSVD, respectively.

4. Discussion

In this study, we introduce a deep learning approach for automatic and accurate estimation of physiological parameters in CTP images. Using global AIF and voxel-wise CTC as inputs, the method was able to provide estimates of local AIF, flow-scaled residue function, and BAD value. The correlation coefficient between the local AIF estimated by the method and the true AIF was 0.97 ± 0.04 ($P < 0.001$) in simulated data. It estimated CBF with a mean error of 4.95 ml/100 g min^{-1} and BAD with a mean error of 0.51 s, both of which were significantly lower than the SVD-based methods ($P < 0.001$). In addition, the proposed method also yielded more accurate parameter estimates in the presence of poor data quality (e.g., noise, delay, dispersion, temporal resolution, bolus shape). Furthermore, when applied to real-world patient data, the proposed network had a distinct advantage over SVD-based methods, providing smooth non-increasing residue function estimates that were more physiologically interpretable. The network yielded higher CBF estimates than the SVD-based methods when evaluating normally perfused and ischemic tissues in patient data.

Regarding the network architecture, we adopted a modified Transformer framework tailored to our specific task requirements. While the vanilla Transformer architecture (Vaswani *et al* 2017) typically employs 6 identical layers for tasks such as machine translation, our input sequences consisted of at most 120 time points, making our task relatively lightweight in comparison. Through preliminary experiments comparing 6-layer and 3-layer configurations, we observed that the performance difference was not significant. The 3-layer configuration not only reduced the number of parameters, making the model less computationally expensive and easier to train, but also maintained sufficient modeling capacity for our task. The training process was completed in approximately one week using our hardware configuration. For inference, the model generated perfusion parameter maps (including CBF, BAD, MTT, CBV) in less than 1 min per slice for the patient dataset. We note that both training and inference times could be further reduced through the use of additional GPU resources and advanced parallel processing techniques.

When transferring the weights trained on the simulated dataset to the real patient dataset, we used a small number of real data samples to supplement and fine-tune the network. The main reason for this is the complexity of the noise distribution in real CTP datasets, which varies between frames and between different brain regions in the same frame (Li *et al* 2022). In addition, the actual residue function may not be fully represented by the family of gamma distributions. Therefore, it is challenging to use simulation to

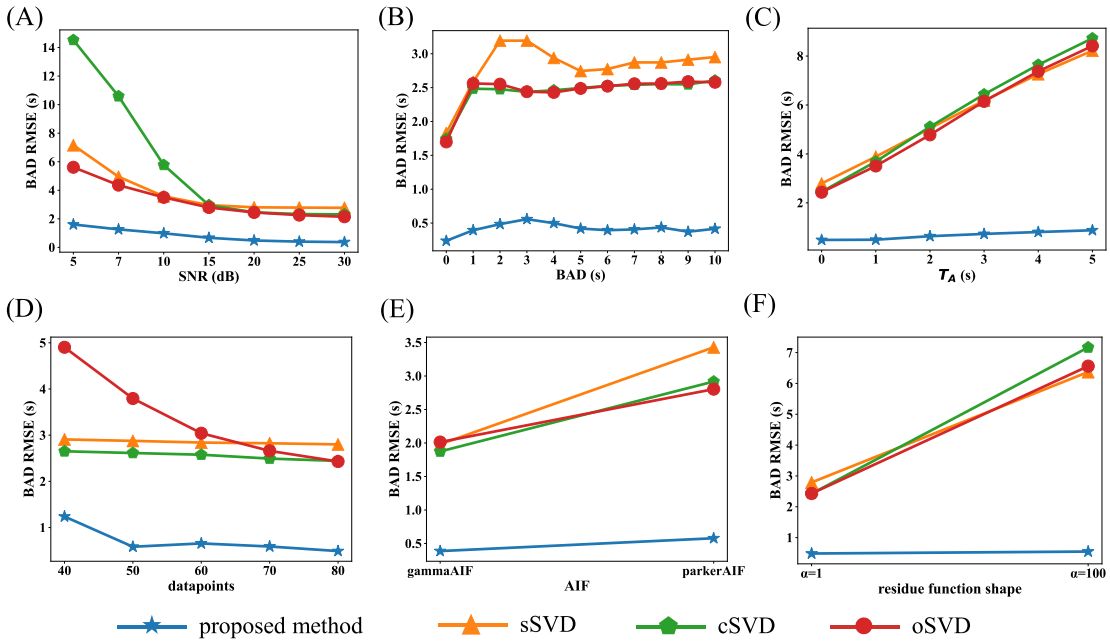

Figure 5. Root-mean-square errors between BAD estimates from different methods and the ground truth physiological parameters. The errors were plotted as a function of (A) noise, (B) delay, (C) dispersion, (D) temporal resolution, (E) AIF type and (F) residue function shape.

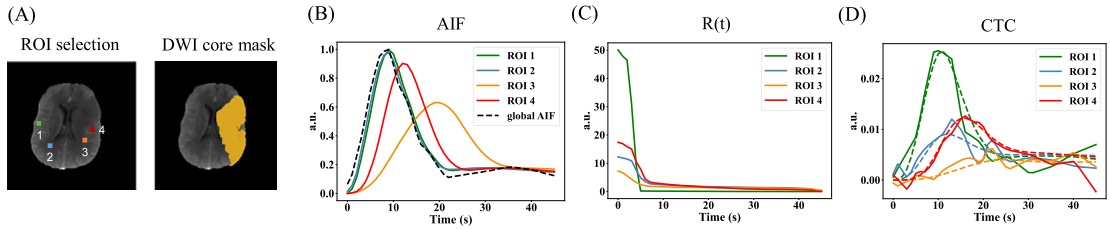

Figure 6. (A) Four regions of interest (ROIs) consisted of 5 × 5 voxels, each with an in-plane size of 0.86 mm × 0.86 mm: normal gray matter (ROI 1, green), normal white matter (ROI 2, blue), ischemic white matter (ROI 3, orange), and ischemic gray matter (ROI 4, red), with the DWI mask providing a delineation of normal and ischemic regions. (B) network-estimated local AIFs (solid line) and sampled global AIF (dashed line). (C) network-estimated flow-scaled residue functions. (D) network-reconstructed CTCs (solid line) and observed CTCs (dashed line) within different ROIs.

encompass the various complexities in real data (Luo *et al* 2024). Furthermore, since the flow-scaled residue function is unknown in reality, the estimation of the residue function is not directly supervised by the ground truth during training on simulated data, which ensures the feasibility of fine-tuning on the patient dataset.

The shape of the residue function reflects the retention of the tracer in the microvasculature. The results of the patient data indicated that the residue function decayed more slowly in ischemic regions compared with normally perfused regions. Residue functions obtained with the SVD-based methods exhibited significant oscillations in the late phase, alternating between positive and negative values. This phenomenon was more pronounced in ischemic regions. Such spurious oscillations may not be consistent with the true physiological behavior of the residue function and may lead to instability of the derived perfusion parameters. In contrast, our proposed method ensures the smoothness and monotonic non-increasing nature of the residue function, which is more consistent with the physiological interpretation and has the potential for more robust parameter estimation.

In practice, in order to achieve a reasonable SNR, AIF is not measured directly in the blood-supplying capillaries near the tissue of interest, but in larger arteries. This approach leads to a traveling time of the tracer bolus and a widening of the bolus shape from where the AIF is measured to the tissue of interest. It has been found that SVD-based methods tend to underestimate CBF and may exaggerate hypoperfusion if dispersion correction is not considered (Fieselmann *et al* 2011). In fact, it is infeasible to locally measure the concentration at the arterial inlet of tissue

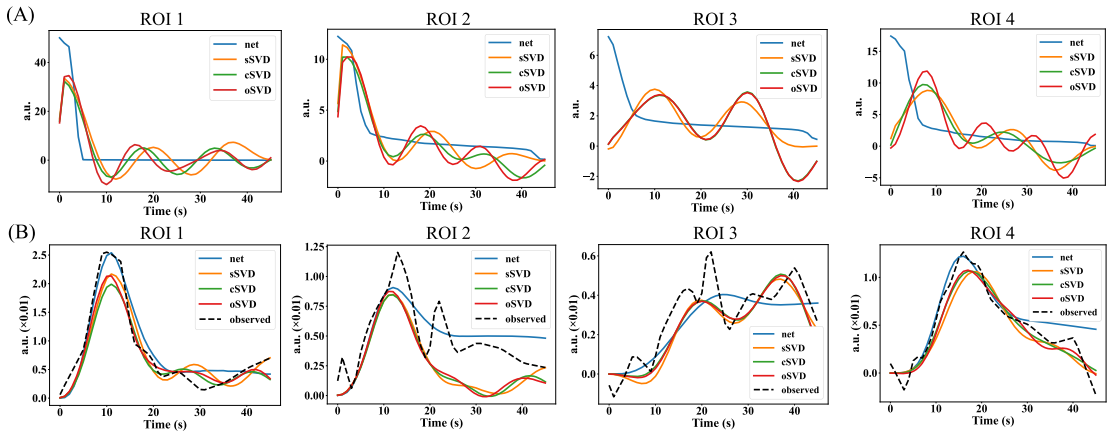

Figure 7. Comparison of the average estimates within different ROIs for the four methods: the proposed network (blue), sSVD (orange), cSVD (green), oSVD (red). (A) estimated flow-scaled residue functions. (B) reconstructed CTCs (solid line) and observed CTCs (dashed line).

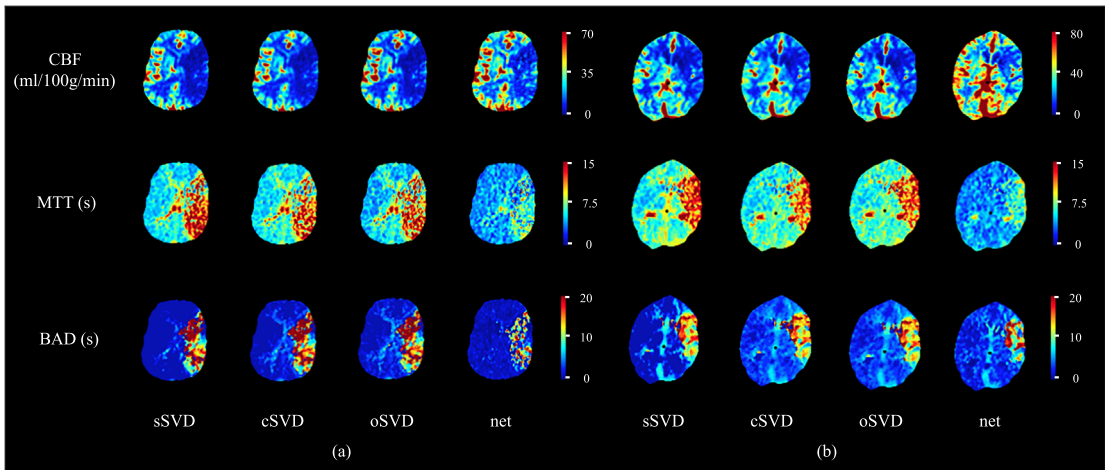

Figure 8. Examples of physiological parameter maps estimated by sSVD, cSVD, oSVD, and the proposed network. The left and right halves showed slices from two different ischemic stroke patients, respectively. Physiological parameters included CBF, MTT and BAD. Each parameter map estimated by all methods was visualized using the same window settings for each patient.

due to limited spatial resolution. Our network used local AIF to correct for the delay and dispersion effects in acute ischemic stroke. As a result, it could provide hemodynamic information at the capillary level, hypothesizing a more accurate estimation of CBF. Conventionally, the CBF measurements obtained by 15O-PET (positron emission tomography) serve as the gold standard for *in vivo* evaluations. Recent studies have shown that with the delay-compensated singular value decomposition method (SVD+) for CTP analysis, the mean value of PET-CBF was 41.13% higher in affected side and 44.03% higher in unaffected side compared to CTP-CBF (Shinohara *et al* 2010). Our results showed that the CBF estimates of the proposed network was 13.83 ml/100 g/min higher (39.33%) than that of sSVD in the normally perfused region, and 8.55 ml/100 g/min higher (57.73%) than that of sSVD in the ischemic region.

Our results showed that the BAD values based on the SVD-based methods were larger than those of the proposed network. The main reason is that the theoretical definition of BAD is the time it takes for arterial blood to reach the relevant voxel. In the SVD methods, the BAD parameter is usually estimated by calculating the time-to-maximum (Tmax) of the residue function. In practice, it has been found that Tmax reflects a combination of BAD, temporal dispersion, and to a lesser degree, mean transit time (Calamante *et al* 2010). This is why the Tmax estimated by the SVD method may be larger than the theoretical BAD. In addition, the sampling of TR data introduces discrete errors in the BAD estimation. Since the measured Tmax can only be a multiple of TR, a 'staircase' effect occurs. In contrast, the BAD estimated by our network is a continuous value, so the estimation is more realistic.

Table 2. Comparison of mean CBF estimates using sSVD, cSVD, oSVD, and the proposed network for brain regions with different perfusion states in a patient dataset containing 94 images. The delineation of normally perfused and ischemic regions was determined by the DWI masks provided by the ISLES18 dataset. The vascular and cerebrospinal fluid regions were excluded from the calculations.

CBF_mean (ml/100 g min^{-1})	sSVD	cSVD	oSVD	net
normally perfused region	35.16	33.84	34.75	48.99
ischemic region	14.81	15.97	14.51	23.36

The proposed method provides a new solution for perfusion image analysis that eliminates the need for presetting specific pharmacokinetic models. Only smoothness and monotonic non-increasing constraints are imposed on the residue function, thus enhancing its generality and interpretability. It is anticipated that this approach will be extended to perfusion analysis in other disease areas such as cancer, along with other imaging modalities including DCE-MRI, DSC-MRI and dynamic PET.

There are some limitations in our work. First, the voxel-level estimation is inherently susceptible to partial volume effects, and future studies could explore parameter estimation at the patch level, an approach that utilizes spatial information and helps to generate parameter maps with lower noise levels. Second, although delay and dispersion correction were addressed separately in this study, their interconnectedness in real-world scenarios requires a more sophisticated modeling framework. Additionally, the commonly used threshold for determining the ischemic penumbra and infarct core regions may not be optimal for the present method. Further refinement and evaluation are necessary for accurate delineation. Finally, the distribution of training data in deep learning can have a significant impact on performance, and adding more prior knowledge may further improve the robustness and accuracy of the proposed method.

5. Conclusions

The proposed method is a promising technique that provides a solution for parameter estimation from CTP images. The method is based on the general indicator dilution theory and directly estimates the flow-scaled residue function in a constrained form. As a result, it does not require predefined specific pharmacokinetic models and is more generalizable across different tissue types or pathological conditions. In addition, the method takes into account potential delay and dispersion effects that may occur in cases of ischemic stroke and is therefore more physiological interpretable and accurate in accessing tissue perfusion. Validated with both simulated data and patient data, the method could serve as a valuable tool for the clinical diagnosis and management of acute ischemic stroke.

Data availability statement

No new data were created or analysed in this study.

ORCID iDs

Wanning Zeng https://orcid.org/0009-0004-2048-8557
Yang Li https://orcid.org/0009-0007-4896-2494
Jeff L Zhang https://orcid.org/0000-0002-3982-0255
Xiaopeng Zong https://orcid.org/0000-0003-4235-6948

References

Bae J, Li C, Masurkar A, Ge Y and Kim S G 2023 Improving measurement of blood-brain barrier permeability with reduced scan time using deep-learning-derived capillary input function *NeuroImage* **278** 120284

Boutelier T, Kudo K, Pautot F and Sasaki M 2012 Bayesian hemodynamic parameter estimation by bolus tracking perfusion weighted imaging *IEEE Trans. Med. Imaging* **31** 1381–95 IEEE Transactions on Medical Imaging

Brahma S, Kofler A, Zimmermann F F, Schaeffter T, Chiribiri A and Kolbitsch C 2024 Robust myocardial perfusion MRI quantification with DeepFermi *IEEE Trans. Biomed. Eng.* **72** 1031–44 IEEE Transactions on Biomedical Engineering

Calamante F 2013 Arterial input function in perfusion MRI: a comprehensive review *Prog. Nucl. Magn. Reson. Spectrosc.* **74** 1–32

Calamante F, Christensen S, Desmond P M, Østergaard L, Davis S M and Connelly A 2010 The physiological significance of the time-to-maximum (Tmax) parameter in perfusion MRI *Stroke* **41** 1169–74

Calamante F, Gadian D G and Connelly A 2000 Delay and dispersion effects in dynamic susceptibility contrast MRI: simulations using singular value decomposition *Magn. Reson. Med.* **44** 466–73

Cereda C W *et al* 2016 A benchmarking tool to evaluate computer tomography perfusion infarct core predictions against a DWI standard *Journal of Cerebral Blood Flow & Metabolism* **36** 1780–9

de la Rosa E, Sima D M, Menze B, Kirschke J S and Robben D 2021 AIFNet: automatic vascular function estimation for perfusion analysis using deep learning *Med. Image Anal.* **74** 102211

Fang K *et al* 2021 Convolutional neural network for accelerating the computation of the extended Tofts model in dynamic contrast-enhanced magnetic resonance imaging *J. Magn. Reson. Imaging* **53** 1898–910

Fieselmann A, Kowarschik M, Ganguly A, Hornegger J and Fahrig R 2011 Deconvolution-based CT and MR brain perfusion measurement: theoretical model revisited and practical implementation details *Int. J. Biomed. Imaging* **2011** 467563

Gava U A *et al* 2023 Neural network-derived perfusion maps: a model-free approach to computed tomography perfusion in patients with acute ischemic stroke *Frontiers in Neuroinformatics* **17** 852105

GBD 2019 Stroke Collaborators 2021 Global, regional, and national burden of stroke and its risk factors, 1990-2019: a systematic analysis for the global burden of disease study 2019 *The Lancet. Neurology* **20** 795–820

Hakim A, Christensen S, Winzeck S, Lansberg M G, Parsons M W, Lucas C, Robben D, Wiest R, Reyes M and Zaharchuk G 2021 Predicting infarct core from computed tomography perfusion in acute ischemia with machine learning: lessons from the ISLES challenge *Stroke* **52** 2328–37

He K, Sun J and Tang X 2013 Guided image filtering *IEEE Trans. Pattern Anal. Mach. Intell.* **35** 1397–409 IEEE Transactions on Pattern Analysis and Machine Intelligence

Ho K C, Scalzo F, Sarma K V, El-Saden S and Arnold C W 2016 A temporal deep learning approach for MR perfusion parameter estimation in stroke *2016 23rd Int. Conf. on Pattern Recognition (ICPR)* 1315–20

Huang H, Liu B, Xu Y and Zhou W 2023 Synthetic-to-real domain adaptation with deep learning for fitting the intravoxel incoherent motion model of diffusion-weighted imaging *Med. Phys.* **50** 1614–22

Khalifa F, Soliman A, El-Baz A, Abou El-Ghar M, El-Diasty T, Gimel'farb G, Ouseph R and Dwyer A C 2014 Models and methods for analyzing DCE-MRI: a review *Med. Phys.* **41** 124301

Koh T S, Cheong D L H and Hou Z 2011 Issues of discontinuity in the impulse residue function for deconvolution analysis of dynamic contrast-enhanced MRI data *Magn. Reson. Med.* **66** 886–92

Li S, Zeng D, Bian Z and Ma J 2022 Noise modelling of perfusion CT images for robust hemodynamic parameter estimations *Phys. Med. Biol.* **67** 115016

Lorenz C *et al* 2006 Automated perfusion-weighted MRI using localized arterial input functions *J. Magn. Reson. Imaging* **24** 1133–9

Luo L *et al* 2024 CT perfusion parameter estimation in stroke using neural network with transformer and physical model priors *Comput. Biol. Med.* **182** 109134

Mouridsen K, Friston K, Hjort N, Gyldensted L, Østergaard L and Kiebel S 2006 Bayesian estimation of cerebral perfusion using a physiological model of microvasculature *NeuroImage* **33** 570–9

Mouridsen K, Hansen M B, Østergaard L and Jespersen S N 2014 Reliable estimation of capillary transit time distributions using DSC-MRI *Journal of Cerebral Blood Flow & Metabolism* **34** 1511–21

Oh G, Moon Y, Moon W-J and Ye J C 2024 Unpaired deep learning for pharmacokinetic parameter estimation from dynamic contrast-enhanced MRI without AIF measurements *NeuroImage* **291** 120571

Østergaard L, Weisskoff R M, Chesler D A, Gyldensted C and Rosen B R 1996 High resolution measurement of cerebral blood flow using intravascular tracer bolus passages. Part I: mathematical approach and statistical analysis *Magn. Reson. Med.* **36** 715–25

Ottens T, Barbieri S, Orton M R, Klaassen R, van Laarhoven H W M, Crezee H, Nederveen A J, Zhen X and Gurney-Champion O J 2022 Deep learning DCE-MRI parameter estimation: application in pancreatic cancer *Med. Image Anal.* **80** 102512

Parker G J M, Roberts C, Macdonald A, Buonaccorsi G A, Cheung S, Buckley D L, Jackson A, Watson Y, Davies K and Jayson G C 2006 Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI *Magn. Reson. Med.* **56** 993–1000

Pizzolato M, Boutelier T and Deriche R 2017 Perfusion deconvolution in DSC-MRI with dispersion-compliant bases *Med. Image Anal.* **36** 197–215

Pu L, Wang L, Zhang R, Zhao T, Jiang Y and Han L 2023 Projected global trends in ischemic stroke incidence, deaths and disability-adjusted life years from 2020 to 2030 *Stroke* **54** 1330–9

Shamshad F, Khan S, Zamir S W, Khan M H, Hayat M, Khan F S and Fu H 2023 Transformers in medical imaging: a survey *Med. Image Anal.* **88** 102802

Shinohara Y, Ibaraki M, Ohmura T, Sugawara S, Toyoshima H, Nakamura K, Kinoshita F and Kinoshita T 2010 Whole-brain perfusion measurement using 320-detector row computed tomography in patients with cerebrovascular steno-occlusive disease: comparison with 15O-positron emission tomography *J. Comput. Assist. Tomogr.* **34** 830–5

Sourbron S P and Buckley D L 2013 Classic models for dynamic contrast-enhanced MRI *NMR Biomed.* **26** 1004–27

Ulas C, Das D, Thrippleton M J, Valdés Hernández M del C, Armitage P A, Makin S D, Wardlaw J M and Menze B H 2019 Convolutional neural networks for direct inference of pharmacokinetic parameters: application to stroke dynamic contrast-enhanced MRI *Frontiers Neurol.* **9** 1147

van Herten R L M, Chiribiri A, Breeuwer M, Veta M and Scannell C M 2022 Physics-informed neural networks for myocardial perfusion MRI quantification *Med. Image Anal.* **78** 102399

Vaswani A, Shazeer N, Parmar N, Uszkoreit J, Jones L, Gomez A N, Kaiser Ł and Polosukhin I 2017 Attention is all you need *Proc. of the 31st Int. Conf. on Neural Information Processing Systems* 6000–10

Welvaert M and Rosseel Y 2013 On the definition of signal-to-noise ratio and contrast-to-noise ratio for fMRI data *PLoS One* **8** e77089

Wu O, Østergaard L, Weisskoff R M, Benner T, Rosen B R and Sorensen A G 2003 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.* **50** 164–74

Zeng Q and Zhou W 2021 An attention based deep learning model for direct estimation of pharmacokinetic maps from DCE-MRI images *2021 IEEE Int. Conf. on Bioinformatics and Biomedicine (BIBM)* 2368–75

Zhou H, Zhang S, Peng J, Zhang S, Li J, Xiong H and Zhang W 2021 Informer: beyond efficient transformer for long sequence time-series forecasting *Proceedings of the AAAI Conference on Artificial Intelligence* **35** 12

Zou J, Balter J M and Cao Y 2020 Estimation of pharmacokinetic parameters from DCE-MRI by extracting long and short time-dependent features using an LSTM network *Med. Phys.* **47** 3447–57