Unpaired deep learning for pharmacokinetic parameter estimation from dynamic contrast-enhanced MRI without AIF measurements

Gyutaek Oh [a], Yeonsil Moon [b], Won-Jin Moon [c,*], Jong Chul Ye [d,*]

[a] Department of Bio and Brain Engineering, Korea Advanced Institute of Science and Technology (KAIST), 291, Daehak-ro, Yuseong-gu, 34141, Daejeon, Republic of Korea
[b] Department of Neurology, Konkuk University Medical Center, 120-1, Neungdong-ro, Gwangjin-gu, 05030, Seoul, Republic of Korea
[c] Department of Radiology, Konkuk University Medical Center, 120-1, Neungdong-ro, Gwangjin-gu, 05030, Seoul, Republic of Korea
[d] Kim Jaechul Graduate School of AI, Korea Advanced Institute of Science and Technology (KAIST), 291, Daehak-ro, Yuseong-gu, 34141, Daejeon, Republic of Korea

ARTICLE INFO

ABSTRACT

DCE-MRI provides information about vascular permeability and tissue perfusion through the acquisition of pharmacokinetic parameters. However, traditional methods for estimating these pharmacokinetic parameters involve fitting tracer kinetic models, which often suffer from computational complexity and low accuracy due to noisy arterial input function (AIF) measurements. Although some deep learning approaches have been proposed to tackle these challenges, most existing methods rely on supervised learning that requires paired input DCE-MRI and labeled pharmacokinetic parameter maps. This dependency on labeled data introduces significant time and resource constraints and potential noise in the labels, making supervised learning methods often impractical. To address these limitations, we present a novel unpaired deep learning method for estimating pharmacokinetic parameters and the AIF using a physics-driven CycleGAN approach. Our proposed CycleGAN framework is designed based on the underlying physics model, resulting in a simpler architecture with a single generator and discriminator pair. Crucially, our experimental results indicate that our method does not necessitate separate AIF measurements and produces more reliable pharmacokinetic parameters than other techniques.

1. Introduction

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is an invaluable T1-weighted imaging technique utilized in the diagnosis and management of various diseases (Verma et al., 2012; Armitage et al., 2011; Van De Haar et al., 2016). In DCE-MRI, a contrast agent (CA) such as gadobutrol (Gd-BT-DO3 A) is injected, and the changes in the CA concentration within different tissues are captured through a series of T1-weighted images over time. This CA concentration data enables the estimation of pharmacokinetic (PK) parameters associated with vascular permeability and tissue perfusion (Sourbron and Buckley, 2013). These PK parameters offer valuable insights into the physiological characteristics of tissues.

There exist mathematical relationships between the concentration of CA and PK parameters, which are captured by tracer kinetic (TK) models such as the extended Tofts (eTofts) model (Tofts et al., 1999), Patlak model (Patlak et al., 1983), etc. Consequently, the estimation of PK parameters involves fitting one of these TK models to the time series of CA concentration for each voxel. The common approach for model fitting is the nonlinear least squares (NLLS) method (Branch et al., 1999), but it is computationally intensive due to a large number of iterations. More importantly, NLLS methods depend on inputs like the arterial input function (AIF), which, if obtained from inappropriate voxels or affected by high noise, can significantly compromise accuracy. Another method, linear least squares (LLS) (Flouri et al., 2016), was proposed to reduce computational complexity and runtime, but it exhibits decreased accuracy when noise levels are high.

Recently, deep learning methods have emerged for PK parameter estimation from DCE-MRI (Choi et al., 2020; Ulas et al., 2019; Fang et al., 2021). These methods offer reduced runtime and yield estimation results comparable to NLLS or LLS methods. However, existing deep learning methods rely on supervised learning, necessitating matched pairs of input DCE-MRI and label PK parameter maps (paired data). Obtaining these label PK parameter maps traditionally entails time-consuming NLLS or LLS estimation. Consequently, the challenge lies in obtaining a comprehensive set of label PK parameter maps to train supervised methods. In such scenarios, the supervised method must be
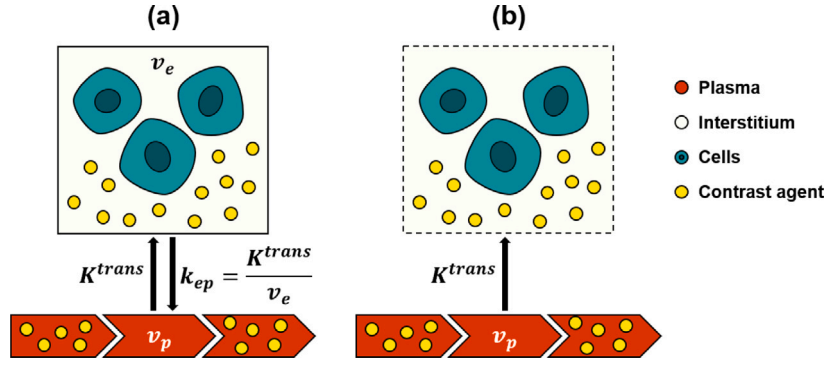
Fig. 1. Two commonly utilized tracer kinetic models in DCE-MRI: (a) the extended Tofts model, and (b) the Patlak model. The parameters in the models are as follows: K^{trans} represents the volume transfer constant, v_p denotes the fractional plasma volume, and v_e signifies the fractional interstitial volume.

trained with a limited dataset, even when a substantial amount of DCE-MRI data that lacks paired PK parameter maps exists. Furthermore, if the label data contains high noise levels, especially due to incorrect AIF estimation, the performance of supervised methods may be adversely affected.

To address the limitations of conventional and supervised methods, here we propose an unpaired deep learning approach for directly estimating PK parameters from DCE-MRI. Our method draws inspiration from the optimal transport-driven CycleGAN (OT-CycleGAN) (Sim et al., 2020) which is applicable in scenarios where matched input-label pairs are unavailable. Essentially, our approach allows the utilization of all available DCE-MRI data, even in the absence of paired PK parameter maps. Leveraging the known TK models relating PK parameters to DCE-MR images, our OT-CycleGAN employs a single generator and discriminator, resulting in reduced computational complexity. Another important contribution is that our model does not rely on AIF for PK parameter estimation as it allows for AIF extraction. Experimental results demonstrate that our proposed method, which does not necessitate separate AIF measurements, achieves competitive performance compared to supervised learning, despite being trained with an unpaired dataset.

This paper is organized as follows. First, we present a comprehensive review of the TK models and optimal transport-driven CycleGAN in Section 2. The theoretical foundations of our proposed method are elaborated upon in Section 3. Following this, we outline the experimental methods (Section 4) and present the corresponding results (Section 5). In Section 6, we engage in a thorough discussion of our findings, and finally, in Section 7, we provide a conclusive summary of our paper.

2. Backgrounds

2.1. Tracer kinetic models

Tracer kinetic (TK) models (Sourbron and Buckley, 2013; Heye et al., 2016) establish the connections between the concentration of the contrast agent (CA) and the pharmacokinetic (PK) parameters. In this study, we will examine two commonly employed TK models in DCE-MRI: the extended Tofts model and the Patlak model.

2.1.1. Extended Tofts model

The extended Tofts (eTofts) model (Tofts et al., 1999) is a TK model that assumes the highly perfused tissue condition (Fig. 1(a)). Specifically, the eTofts model assumes the presence of infinite plasma flow. According to the eTofts model, the CA concentration in the tissue, denoted as $C_t(t)$, is represented by the following equation:

$$C_t(t) = v_p C_p(t) + K^{trans} \int_0^t C_p(\tau) e^{-\frac{K^{trans}}{v_e}(t-\tau)} d\tau \tag{1}$$

where $C_p(t)$ represents the CA concentration in the plasma, v_p corresponds to the fractional plasma volume, v_e denotes the fractional

interstitial (or extravascular extracellular space) volume, and K^{trans} denotes the volume transfer constant from the plasma to the interstitium. Consequently, the estimation of the eTofts model involves determining three PK parameters: K^{trans}, v_p, and v_e.

2.1.2. Patlak model

When the backflux from the interstitium to the plasma is negligible, as depicted in Fig. 1(b), the eTofts model can be simplified to a two-parameter model, expressed as follows:

$$C_t(t) = v_p C_p(t) + K^{trans} \int_0^t C_p(\tau) d\tau. \tag{2}$$

Hence, this two-parameter model is known as the Patlak model (Patlak et al., 1983), which can be considered a special case of the eTofts model. In the Patlak model, K^{trans} and v_p are estimated.

2.1.3. Contrast agent concentration to DCE-MRI

If the time series DCE-MRI signal, denoted as $S(t)$, is acquired using the spoiled gradient echo, it can be converted to the corresponding CA concentration, denoted as $C_t(t)$, by (Di Giovanni et al., 2009; Chao et al., 2017)

$$C_t(t) = -\frac{1}{B}\left[A + \ln\left(\frac{\frac{S(t)}{S_0}\frac{1-e^{-A}}{1-\cos(\alpha)e^{-A}} - 1}{\frac{S(t)}{S_0}\cos(\alpha)\frac{1-e^{-A}}{1-\cos(\alpha)e^{-A}} - 1}\right)\right] \tag{3}$$

where S_0 is the pre-contrast image intensity, α is the flip angle, and

$$A = \frac{TR}{T_1}, \quad B = r_1 TR$$

where TR represents the repetition time, T_1 signifies the T_1 relaxation, and r_1 denotes the relaxivity of the contrast agent. Conversely, if we intend to obtain $S(t)$ from the CA concentration $C_t(t)$, derived from either the eTofts or Patlak model, Eq. (3) can be transformed as follows:

$$S(t) = \frac{(1 - e^{-A - B \cdot C_t(t)})(1 - \cos(\alpha)e^{-A})}{(1 - \cos(\alpha)e^{-A - B \cdot C_t(t)})(1 - e^{-A})} S_0. \tag{4}$$

2.2. Deep learning methods for pharmacokinetic parameter estimation

Due to the high computational complexity and additional requirements such as the arterial input function (AIF), deep learning methods have emerged as an alternative for PK parameter estimation. Choi et al. (2020) introduced conditional generative adversarial networks to translate AIF from DCE-MRI to AIF from dynamic susceptibility-enhanced MRI (DSC-MRI). By leveraging the higher reliability of AIF derived from T2*-weighted DSC-MRI, compared to the low signal intensity T1-weighted DCE-MRI, they demonstrated improved reliability of estimated PK parameters. However, this method does not directly estimate the PK parameters through deep learning.

Ulas et al. (2019) and Fang et al. (2021) proposed the use of convolutional neural networks (CNNs) for the direct estimation of PK parameters. Their approaches employed CNNs with local and global pathways, incorporating forward physical loss. These methods exhibited reliable results and reduced runtime compared to traditional NLLS or LLS methods. However, both approaches rely on supervised learning, necessitating paired input DCE-MRI images and corresponding labeled PK parameter maps. Acquiring of these labeled PK parameter maps for supervised training typically requires time-consuming NLLS or LLS estimations. Moreover, the accuracy of the label PK parameter maps heavily depends on the accuracy of AIF measurement. As a result, the process suffers from extended estimation time or degraded data quality.

To overcome these challenges, we proposed an unpaired deep learning method for the direct estimation of PK parameters using a physics-driven CycleGAN. This approach provides solutions to the challenges linked with obtaining labeled training data, facilitating direct approximation without requiring paired DCE-MRI images, matched PK parameter maps, and AIF measurements.

2.3. Optimal transport-driven CycleGAN

In our previous investigations (Sim et al., 2020; Lim et al., 2020), we uncovered the connection between the architecture of CycleGAN and the dual formulation of an optimal transport problem. In this section, we present a concise overview of the theoretical foundations explored in those studies.

Consider the following measurement model:

$$y = \mathcal{H}x + w \tag{5}$$

where $x \in \mathcal{X}$ and $y \in \mathcal{Y}$ denote the unknown image and the noisy measurement, respectively, w is the measurement noise, and $\mathcal{H} : \mathcal{X} \mapsto \mathcal{Y}$ is the known deterministic imaging operator. In supervised learning, the objective is to learn the relationship between image x and its corresponding measurement y pairs. However, in unpaired learning scenarios, there is no availability of matched image-measurement pairs. Consequently, the goal shifts towards matching the probability distributions instead of establishing direct correspondences between individual samples from the unpaired sets of images and measurements. This is accomplished by identifying transportation maps that facilitate the transportation of probability distributions between the two spaces.

Let us consider the target image space $\mathcal{X}$ and the measurement space $\mathcal{Y}$, which are associated with probability measures μ and ν, respectively. To facilitate the transportation of mass from $(\mathcal{X}, \mu)$ to $(\mathcal{Y}, \nu)$, we employ the forward operator $\mathcal{H}$. Conversely, the generator $G_\Theta : \mathcal{Y} \mapsto \mathcal{X}$, parameterized by Θ, performs the mass transport from $(\mathcal{Y}, \nu)$ to $(\mathcal{X}, \mu)$. In this context, the following transportation cost was proposed for the optimal transport problem (Sim et al., 2020):

$$c(x, y; \Theta) := \|y - \mathcal{H}x\| + \|G_\Theta(y) - x\|, \tag{6}$$

which denotes the sum of the distance between a training sample and a transported sample in each space. The objective of the optimal transport is not limited to minimizing the sample-wise cost computed using Eq. (6). Instead, its goal is to minimize the average transport cost. Specifically, the optimal transport problem aims to identify the joint distribution π that results in the minimum average transport cost. Thus, the problem is formulated as finding the π that minimizes the average transport cost:

$$\inf_{\pi \in \Pi(\mu,\nu)} \int_{\mathcal{X} \times \mathcal{Y}} c(x, y; \Theta) d\pi(x, y), \tag{7}$$

where $\Pi(\mu, \nu)$ denotes the set of joint measures whose marginal distributions in $\mathcal{X}$ and $\mathcal{Y}$ are μ and ν, respectively.

Using the transportation cost in (6), the Kantorovich dual formulation (Villani, 2008) is given by (Sim et al., 2020)

$$\min_\Theta \mathbb{K}(\Theta, \mathcal{H}) = \min_\Theta \max_\Phi \ell(\Theta, \Phi), \tag{8}$$

where

$$\ell(\Theta; \Phi) = \gamma \ell_{cycle}(\Theta) + \ell_{WGAN}(\Theta; \Phi). \tag{9}$$

Here, γ is a suitable hyperparameter, ℓ_{cycle} is the cycle-consistency loss, and ℓ_{WGAN} is the Wasserstein GAN (WGAN) loss (Martin Arjovsky and Bottou, 2017). More specifically, ℓ_{cycle} and ℓ_{WGAN} are given by

$$\ell_{cycle}(\Theta) = \int_{\mathcal{X}} \|x - G_\Theta(\mathcal{H}x)\| d\mu(x) + \int_{\mathcal{Y}} \|y - \mathcal{H}G_\Theta(y)\| d\nu(y), \tag{10}$$

$$\ell_{WGAN}(\Theta; \Phi) = \int_{\mathcal{X}} \varphi_\Phi(x) d\mu(x) - \int_{\mathcal{Y}} \varphi_\Phi(G_\Theta(y)) d\nu(y). \tag{11}$$

It is worth noting that in the proposed architecture, there is only one discriminator, denoted as φ_Φ, due to the known nature of the forward operator $\mathcal{H}$. As a result, there is no need for the discriminator to compete with the forward operator. This simplicity in the CycleGAN architecture will be further elucidated in the subsequent explanation.

3. Theory

Now, we will employ the optimal transport-driven CycleGAN for PK parameter estimation. In our specific scenario, a time series DCE-MRI signal $S \in \mathcal{S}$ and a PK parameter map $P \in \mathcal{P}$ correspond to a measurement and an unobserved image, respectively. To be more precise, we can formulate the forward model mapping PK parameter maps to DCE-MRI as follows:

$$S = f_{TK}(P, C_p) \tag{12}$$

where f_{TK} is the function composed by the TK models (Eqs. (1) or (2)) and (4). Here, the availability of the arterial plasma concentration $C_p \in \mathcal{C}$ is necessary, and it is feasible to utilize $C_p(t)$ derived from the measured AIF.

However, if the measured AIF is affected by noise or obtained from inappropriate voxels, it can significantly degrade the performance of our model. To address this challenge, we propose a framework where the generator simultaneously estimates the PK parameter maps and the arterial plasma concentration C_p. This is illustrated in Fig. 2. Additionally, in the lower branch of the CycleGAN framework depicted in Fig. 2, we apply a cycle-consistency loss for C_p to ensure that the estimated AIF remains consistent and avoids unusual estimations.

By identifying $(P', C_p') = G_\Theta(S)$, $S' = f_{TK}(G_\Theta(S))$ and $(P'', C_p'') = G_\Theta(f_{TK}(P, C_p))$, the cycle-consistency loss and WGAN loss for our OT-cycleGAN can be represented by

$$\ell_{cycle}(\Theta) = \int_{\mathcal{P}} \|P - P''\| d\mu(P) + \int_{\mathcal{S}} \|S - S''\| d\nu(S) \\ + \rho \int_{\mathcal{C}} \|C_p - C_p''\| d\kappa(C_p), \tag{13}$$

$$\ell_{WGAN}(\Theta; \Phi) = \int_{\mathcal{P}} \varphi_\Phi(P) d\mu(P) - \int_{\mathcal{S}} \varphi_\Phi(P') d\nu(S), \tag{14}$$

where ρ is a hyperparameter to control the weight of cycle-consistency loss, κ denotes the probability measure for C_p distribution. Since C_p has a different range of values compared to DCE-MRI or PK parameter maps, we adjust the weight of cycle-consistency loss for C_p by ρ.

While the direct application of OT-CycleGAN can be utilized for PK parameter estimation, we choose to employ the least squares GAN (LSGAN) (Mao et al., 2017) instead of WGAN to ensure more stable training. The relationship between WGAN and LSGAN was previously discussed in previous work (Lim et al., 2020).

$$\ell_{LSGAN}(\Theta; \Phi) = \frac{1}{2} \int_{\mathcal{P}} [\varphi_\Phi(P)]^2 d\mu(P) + \frac{1}{2} \int_{\mathcal{S}} [\varphi_\Phi(P') - 1]^2 d\nu(S). \tag{15}$$

Consequently, the final cost function for OT-CycleGAN can be formulated as

$$\ell(\Theta, \Phi) = \gamma \ell_{cycle}(\Theta) + \ell_{LSGAN}(\Theta; \Phi). \tag{16}$$
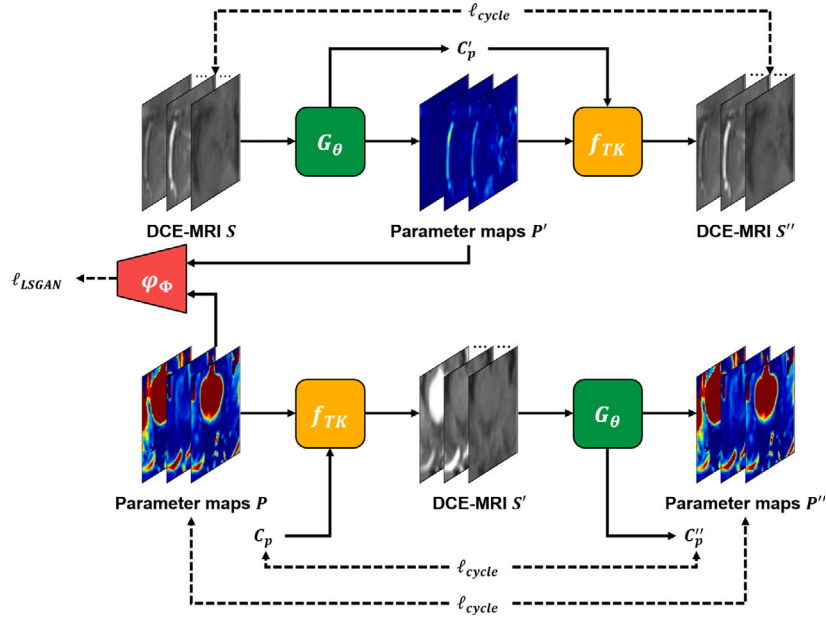
Fig. 2. The architecture of the proposed CycleGAN for PK parameter estimation consists of key components: G_θ representing the generator, f_{TK} denoting the known tracer kinetic model, φ_Φ representing the discriminator, and C_p representing the arterial plasma concentration. The proposed CycleGAN utilizes a single generator and a single discriminator due to the deterministic nature of the physics model mapping PK parameters to DCE-MRI. DCE-MRI S and parameter maps P are unpaired.

4. Methods

4.1. Experimental dataset

In our experiments, we utilized two datasets acquired from Konkuk University Medical Center. This study was approved by the institutional review board of Konkuk University Medical Center (IRB No. 2021-06-036) and was performed in accordance with the ethical standards as laid down in the 2013 Declaration of Helsinki and its later amendments. Because of its retrospective nature, the requirement for prior consent was waived by the institutional review board. Instead, written consent was obtained from all patients for the use of their data for research and publication purposes.

For both datasets, we employed gadobutrol as the contrast agent with a relaxivity value of $r_1 = 3.47$ mM^{-1}s^{-1} (Szomolanyi et al., 2019). A standard dose of gadobutrol (0.1 mmol/kg body weight; Bayer Healthcare) was administered with a 30 mL saline flush using an automatic injector, after the fourth dynamic scan, at a flow rate of 2 mL/s. To account for T1 inhomogeneity during kinetic parameter calculations, T1 mapping was generated using a pre-contrast T1-weighted gradient echo series with six different flip angles (2–12°).

4.1.1. DCE-MRI for brain tumor

The first dataset, referred to as the '**Tumor Data**', consisted of DCE imaging data acquired from 116 patients with brain tumors. An axial 3D DCE sequence was employed, comprising a dynamic series of 65 individual scans. The imaging parameters for this sequence were as follows: field strength = 3T, TR = 2.80 ms, TE = 0.90 ms, flip angle = 10°, average = 1, field-of-view = 222 × 240 mm^2, slice thickness = 4 mm, acquisition matrix = 160 × 148, voxel size = 0.75 ×0.75× 4.00 mm^3, acquisition time = 7 min, and time resolution = 6.5 s. Among the patients, data from 98 individuals were utilized for training, while the remaining data were used for inference purposes in simulation and in vivo studies.

4.1.2. DCE-MRI for normal control and MCI workup

The second dataset referred to as the '**MCI Data**', comprised imaging data obtained from subjects with normal cognition and mild cognitive impairment (MCI). A coronal 3D DCE sequence was employed to acquire a dynamic series of 60 individual scans. The imaging parameters for this sequence were as follows: field strength = 3T, TR = 3.72 ms, TE = 1.04 ms, flip angle = 10°, average = 1, field-of-view = 225 × 240 mm^2, slice thickness = 3 mm, acquisition matrix = 180 × 192, voxel size = 1.25 ×1.25× 3.00 mm^3, acquisition time = 10 min, and time resolution = 10 s. The acquisition time of 10 min was chosen considering both patient compliance and mathematical modeling requirements. The coronal plane was selected for DCE imaging as it provides an optimal assessment of the hippocampi and temporal lobes. For training purposes data from 196 subjects were utilized, while the evaluation of our method was performed using data from 36 subjects in simulation and in vivo studies.

4.1.3. DCE-MRI analysis

Postprocessing and region of interest selection in the DCE imaging data were performed using NordicICE software (Version 4.1.3), with T1-volume imaging utilized for structural imaging. The analysis was conducted by a trained researcher under the supervision of an expert neuroradiologist. Both the researcher and supervisor were blinded to the clinical diagnosis.

For the conventional approaches, the arterial input function (AIF) was obtained using the semiautomatic method within the NordicICE software. Following motion correction and noise modulation, automatic AIF search box detection was applied. In the first dataset, the AIF search box was placed in the area of the internal carotid artery on the axial slice, while in the second dataset, it was placed in the area of the superior sagittal sinus on the coronal plane. The final AIF curve was determined by calculating the mean of one to five different curves (up to 5 pixels) within the search box. For each patient, three to five (or one to five in the second dataset) pixel-time curves were selected, exhibiting expected input function properties such as low first moment, high peak enhancement, and a satisfactory area under the curve. After obtaining the AIF, we converted it to plasma concentration using the formula $C_p(t) = C_b(t)/(1-\text{Hct})$, where $C_b(t)$ represents the measured AIF and Hct denotes the blood hematocrit. Consistent with previous studies (Heye et al., 2016; Ulas et al., 2019), we used a value of Hct = 0.45.

For the 'Tumor Data', the eTofts permeability model was employed to estimate standard DCE perfusion parameters in the tissue (Padhani, 2002). Maps of K^{trans}, v_p, and v_e derived from the eTofts model were
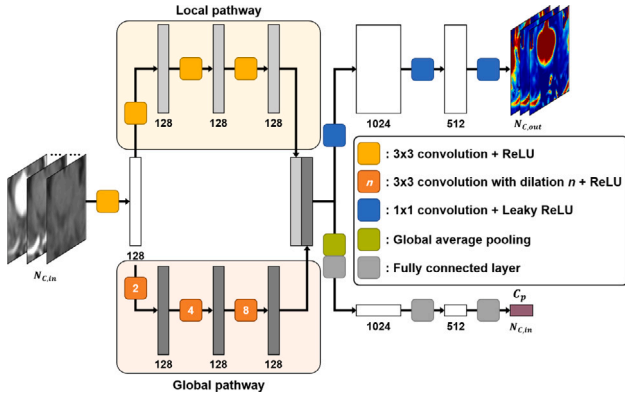
Fig. 3. The architecture of the generator comprises two pathways: the local pathway and the global pathway. The local pathway focuses on capturing local features, while the global pathway aims to capture global context information. The local and global features are then merged, and the split pathways estimate the PK parameters and C_p. The number below each block indicates the number of channels.

Fig. 4. The supervised methods for comparison. '**Supervised**' refers to the supervised method trained with L1 loss between the output and label, while '**Supervised (physics)**' denotes the supervised learning approach incorporating the physics model.

calculated following AIF acquisition. Conversely, for the 'MCI Data', the Patlak model, considered optimal for low-leakage conditions (Heye et al., 2016; Thrippleton et al., 2019), was utilized. Specifically, the two parameters, K^{trans} and v_p, were calculated using the Patlak model. In the NordicICE software, the LLS method (Flouri et al., 2016) was employed for estimating the PK parameters.

4.2. Network architecture

The CycleGAN architecture for PK parameter estimation is illustrated in Fig. 2. Unlike the vanilla CycleGAN, which employs two generators and two discriminators, our CycleGAN, designed for the deterministic physics model from PK parameters to DCE-MRI, utilizes only one generator and one discriminator. To generate DCE-MRI signals from PK parameter maps, the pre-contrast image intensity S_0, T_1 map, C_p, and scan parameters such as repetition time, flip angle, the time interval between frames, and relaxivity of the contrast agent are required. Furthermore, it is important to note that DCE-MRI S and parameter maps P are not paired, and the proposed cycleGAN is trained in an unpaired manner. On the other hand, during inference, only the time series DCE-MRI is required as input to the generator, eliminating the need for additional measurements such as T_1 or C_p.

The generator architecture is depicted in Fig. 3. The input consists of concatenated time frames of DCE-MRI, resulting in a channel dimension equal to the number of time frames $N_{C,in}$. The initial convolutional layer extracts features from the input. The generator is then divided into two pathways, as seen in previous works (Ulas et al., 2019; Fang et al., 2021). The local pathway employs three 3×3 convolutional layers to capture local details, while the global pathway utilizes three dilated convolutional layers with different dilations to capture global information. The outputs of the local and global pathways are concatenated, and this combined feature is fed into two separate pathways. The first pathway estimates the PK parameter maps using three 1×1 convolutional layers. The number of channels in the output $N_{C,out}$ corresponds to the number of parameters to be estimated, which is 3 for the eTofts model and 2 for the Patlak model. In the second pathway, C_p is estimated using global average pooling and fully connected layers. The number of channels in C_p matches the number of input channels ($N_{C,in}$), which represents the number of time frames.

For the discriminator, we adopt the PatchGAN discriminator introduced in Zhu et al. (2017). It consists of 4×4 convolutions, instance normalization (Ulyanov et al., 2016), and leaky ReLU activation. The initial number of filters in the discriminator is set to 32.
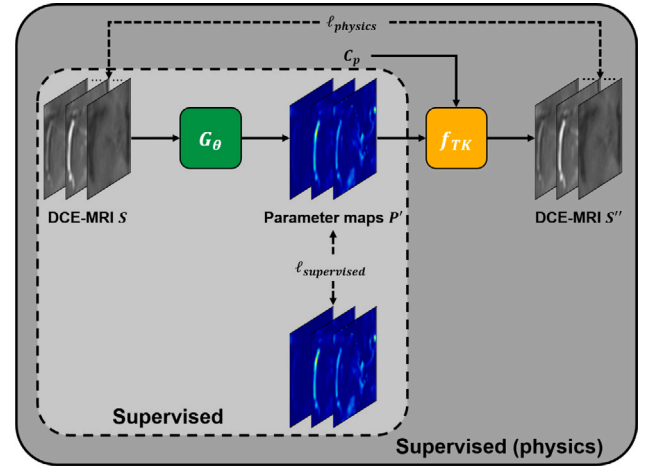
4.3. Implementation details

To handle PK parameters that may exhibit values outside the expected range in extra-brain regions, we employed the brain extraction tool (Smith, 2002) to extract brain masks, using only PK parameters within the brain regions. During training, small patches were cropped from DCE-MRI images or PK parameter maps. The patch size was set to 48×48, and for every training step, patches were randomly cropped from various locations. To augment the data, random flipping was applied. Additionally, the PK parameters were scaled by appropriate factors to match their range. Specifically, K^{trans}, v_p, and v_e in the eTofts model were scaled by 20, 40, and 4, respectively, while K^{trans} and v_p in the Patlak model were scaled by 40 and 8, respectively. During inference, the entire DCE-MRI images were fed into the network.

For the regularization parameters, we chose $\gamma = 10$ (Eq. (16)), $\rho = 1$ for the 'Tumor Data' and $\rho = 0.1$ for the 'MCI Data (Eq. (13)). We utilized the Adam optimizer with $\beta_1 = 0.5$, $\beta_2 = 0.999$, and a batch size of 32 to train our networks. The initial learning rate was set to 10^{-5} with linear decay applied. Our CycleGAN was trained for 200 epochs and implemented in Python using PyTorch.

4.4. Comparison methods

To validate the effectiveness of the proposed methods, we utilized two types of supervised learning approaches as comparative baselines. Firstly, we employed a supervised method trained with a simple L1 loss, defined as follows:

$$\ell(\Theta) = \ell_{supervised}(\Theta) \tag{17}$$

where $\ell_{supervised}$ is the L1 loss between the network output and the ground truth. Furthermore, we utilize the same network architecture as the generator in the CycleGAN, along with the other predefined hyperparameters.

In addition, we incorporate the physics model into the supervised learning method, following the approach employed in existing methods for PK parameter estimation (Ulas et al., 2019; Fang et al., 2021). In the supervised method with the physics model, the network output is used to reconstruct a time series of DCE images by incorporating the physics model. Subsequently, an additional loss, $\ell_{physics}$, is computed, and the supervised learning with the physics model is trained using the following loss function:

$$\ell(\Theta) = \ell_{supervised}(\Theta) + \ell_{physics}(\Theta) \tag{18}$$

where $\ell_{physics}$ is the L1 loss between the input DCE images and the reconstructed DCE images. The overall flow of the supervised methods is illustrated in Fig. 4.

4.5. Evaluation methods

4.5.1. Simulation studies

Although the PK parameter maps obtained from the LLS method are used for training the methods, there are cases where these parameters may not be reliable, especially when the arterial input function (AIF) or other measured parameters are inaccurate. Therefore, a direct quantitative comparison between the LLS method and deep learning methods may not be appropriate, as it does not account for the potential inaccuracies in the LLS method. Even if deep learning methods better capture the physics model and yield improved results, the quantitative evaluation alone may not reflect their actual performance.

To address this limitation, we generated simulated DCE time series images based on the physics model and the PK parameters obtained from the LLS method. More specifically, we computed $C_t(t)$ using PK parameters estimated by the LLS method with Eqs. (1) or (2). Here, we employed AIFs obtained from the NordicICE software to characterize $C_p(t)$. Subsequently, we generated the simulated time series DCE-MRI $S(t)$ using Eq. (4), which was then employed as input data for the supervised methods and the proposed method. After generating simulated DCE-MRI data, we introduced Gaussian noise to enhance the realism of the simulation.

Since the simulated data is generated based on the physics model, we can consider the PK parameters from the LLS method as ground truth. Consequently, deep learning methods that take the simulated data as input should yield similar results to the LLS method, as they effectively capture the underlying physics model. Therefore, we quantitatively evaluated the deep learning methods using the simulated data. For this evaluation, we employed two commonly used metrics: the peak signal-to-noise ratio (PSNR) and the structural similarity index measure (SSIM). The PSNR can be calculated as follows:

$$\text{PSNR} = 20 \log_{10} \left(\frac{MAX}{\sqrt{MSE}} \right), \tag{19}$$

where MAX is the maximum possible pixel value of the image, and MSE is the mean squared error between the ground truth image and the estimated image. Next, SSIM can be calculated as follows:

$$\text{SSIM} = \frac{(2\mu_z \mu_{\hat{z}} + c_1)(2\sigma_{z,\hat{z}} + c_2)}{(\mu_z^2 + \mu_{\hat{z}}^2 + c_1)(\sigma_z^2 + \sigma_{\hat{z}}^2 + c_2)}, \tag{20}$$

where z is the ground truth image, $\hat{z}$ is the estimated image, μ_z, $\mu_{\hat{z}}$ are the average of each image, σ_z, $\sigma_{\hat{z}}$ are the variance of each image, $\sigma_{z,\hat{z}}$ is the covariance of images, $c_1 = (k_1 L)^2$, $c_2 = (k_2 L)^2$ are variables to stabilize the division with weak denominator, where L is the dynamic range of pixel values, and $k_1 = 0.01$, $k_2 = 0.03$. We assessed statistical differences in quantitative results among various methods using a Friedman test.

Furthermore, a Friedman test was executed to compare the estimated PK parameter values, focusing specifically on the contrast-enhancing tumor regions within the simulated 'Tumor Data' and the hippocampus region of the simulated 'MCI Data'. The region of interest for tumors was defined by radiologists based on contrast-enhanced regions. This involved a comprehensive examination of results obtained from all patients, utilizing the Friedman test to identify statistical differences among the methods. Subsequently, the assessment of p-values facilitated the identification of specific statistical variations between the methods.

4.5.2. In vivo studies

Subsequently, we conducted experiments using in vivo DCE-MRI data. As the accuracy of the LLS method relies on additional inputs such as AIF, its results may not always be reliable. Hence, the utilization of in vivo data allows us to verify whether our method can offer more accurate results in comparison to the LLS method.

The quality of the arterial input function (AIF) is significantly influenced by the data quality or voxel selection. To mitigate the potential performance degradation caused by AIF quality, our model not only estimates PK parameters but also jointly estimates the AIF. This integrated approach helps maintain the overall model performance. To verify that the estimated AIF by our model improved the performance of our model itself, we evaluated the quality of the estimated AIF.

To assess the efficacy of our AIF estimation, we conducted a clinical evaluation to evaluate the quality of estimated AIFs. The AIF was evaluated for quality by a neuroradiologist with 23 years of experience. It was assessed using a 3-point scoring system: 1 = poor, 2 = fair, and 3 = good. The visual assessment partly referred to the quality assurance for perfusion CT as mentioned in Chung et al. (2021). The AIF shape was comprehensively checked for early tracer arrival, high peak height, and a quick wash-out pattern (Mouridsen et al., 2006). Cases that showed a distinct sharp and single peak were rated as 'good,' while those that exhibited multiple peaks or a jagged time course were rated as 'poor'.

5. Experimental results

5.1. Simulation studies

Qualitative results. Fig. 5 presents the results of PK parameter estimation using various methods on simulated 'Tumor Data'. The simulation assumes that the LLS method provides ground truth results. Hence, the estimated results from deep learning methods are expected to be consistent with those of the LLS method. However, as depicted in Fig. 5, the supervised methods, both with and without the physics model, exhibit differences in the estimation of v_p and v_e compared to the LLS method. For instance, the supervised method provides higher estimates of v_e in the tumor region in comparison to the LLS method, whereas the supervised learning method with the physics model yields relatively lower estimates of v_e in the tumor region when compared to the LLS method. On the other hand, our proposed method exhibits PK parameter values estimated in the tumor region that are similar to those obtained with the LLS method. Fig. 6 presents the results of PK parameter estimation using various methods on simulated 'MCI Data'. Again, our CycleGAN produces results that closely resemble the ground truth data.

Quantitative results. The obtained quantitative results of the proposed method and supervised methods are presented in Table 1. Notably, supervised learning demonstrates inferior results across the quantitative metrics. This observation suggests that conventional supervised learning approaches do not effectively incorporate the underlying physical model. While applying the physics model to supervised learning leads to some improvement in the quantitative results, the gains are relatively modest. Conversely, our CycleGAN method exhibits notably superior quantitative results to the supervised methods. In addition, we verified the presence of statistical differences in the quantitative results between our method and the supervised methods using the Friedman test. This outcome strongly suggests that our proposed approach successfully estimates PK parameters by explicitly reflecting the physics model.

Statistical results. The results of the Friedman test for the contrast-enhanced tumor regions in the simulated 'Tumor Data' are presented in Table 2. In Table 2, PK parameter values for each type of tumor are provided, confirming that the proposed method exhibits similar PK parameter values to the LLS method in most cases. Furthermore, no statistical differences were observed between the ground truth LLS method and the proposed method.

It is noteworthy that, for v_e values in brain metastasis, all deep learning methods show lower values compared to the LLS method.
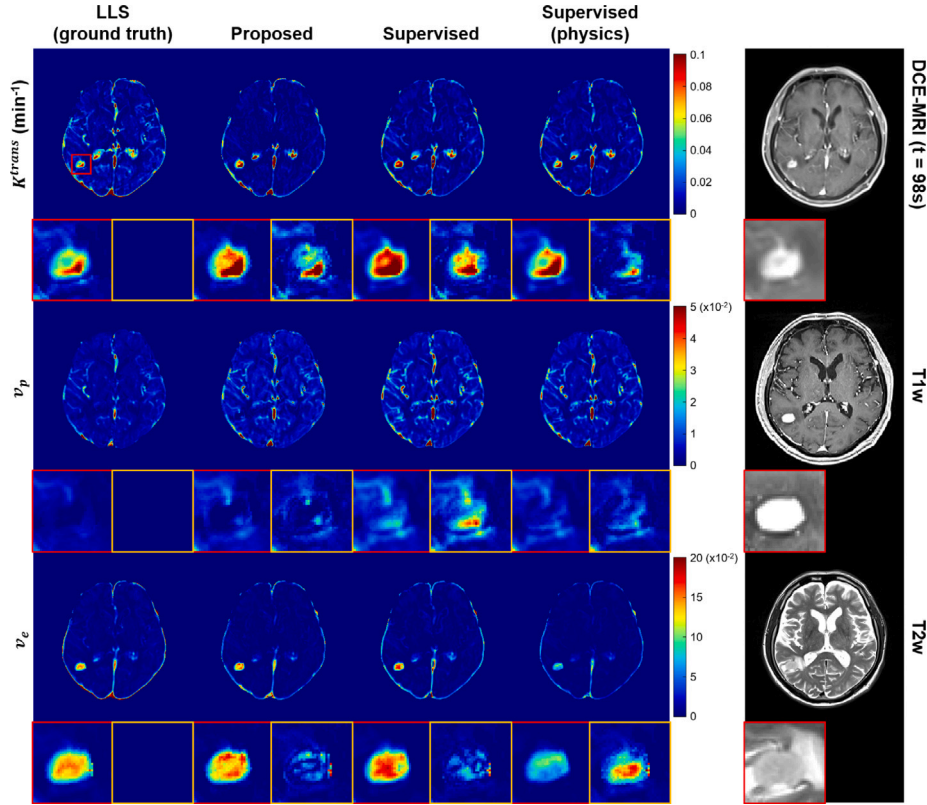
Fig. 5. Results of PK parameter estimation methods for simulated 'Tumor Data' (eTofts model). Here, the LLS corresponds to the ground truth data, from which dynamic images are generated using the eTofts model for inference by various models. Images outlined in red represent enlarged images that emphasize the tumor (glioma) region. Images outlined in yellow are difference maps illustrating the absolute variance between each image and the LLS image. The difference maps are scaled by a factor of 2 to enhance visibility. The scale bar for each parameter and difference map is located to the right of the parameter maps. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)
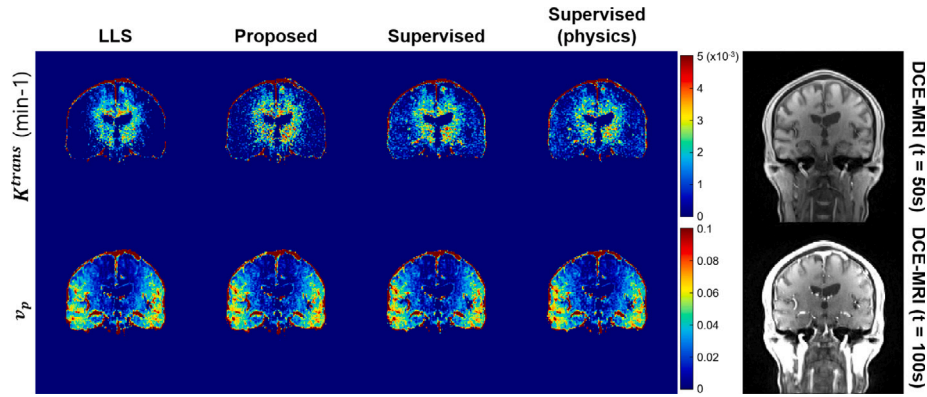

Fig. 6. Results of PK parameter estimation methods for simulated 'MCI Data' (Patlak model). Here, the LLS corresponds to the ground truth data, from which dynamic images are generated using the Patlak model for inference by various models. The scale bar for each parameter is located to the right of the parameter maps.

Firstly, it is crucial to emphasize that the v_e values in Table 2 specifically represent measurements in the contrast-enhancing tumor region, reflecting the characteristics of solid tumors. These observed v_e values are within a clinically acceptable range (Kang et al., 2021), considering the high cellularity commonly associated with solid tumors.

Secondly, the observed discrepancy is attributed to some overestimated v_e values that fall outside the common range with the LLS method. Notably, v_e values in brain metastasis for the LLS method are relatively higher than those in other regions.

Table 3 presents the average PK parameter values in the hippocampus region of the simulated 'MCI Data'. Consistently, both the LLS method and our proposed method demonstrate similar average values,

indicating that our approach effectively incorporates the underlying physics model during PK parameter estimation. However, the supervised methods, with or without the physics model, yield results that differ from those obtained by the LLS method.

Fig. 7 displays Bland–Altman plots illustrating the agreement between the PK parameter values in tumor regions of simulated 'Tumor Data'. As already shown in Table 2, our proposed method demonstrates the minimum mean differences with the ground truth in K^{trans} and v_p when compared to supervised methods (Fig. 7). Additionally, for 'MCI Data', the proposed method again exhibits the minimum mean differences with the ground truth compared to supervised methods as illustrated in Fig. 8.

Table 1

Quantitative evaluation results of various methods using simulated data (average ± standard deviation).

Data	PK parameter	Method	PSNR (dB)	SSIM	Different from ($P < 0.05$) (PSNR & SSIM)
Tumor Data (eTofts)	K^{trans}	Proposed	37.45 ± 4.22	0.968 ± 0.02	Supervised, Supervised (physics)
		Supervised	37.61 ± 3.91	0.969 ± 0.02	Proposed, Supervised (physics)
		Supervised (physics)	**37.65** ± 3.94	**0.969** ± 0.02	Proposed, Supervised
	v_p	Proposed	**39.48** ± 4.37	**0.975** ± 0.02	Supervised, Supervised (physics)
		Supervised	39.16 ± 4.54	0.973 ± 0.02	Proposed, Supervised (physics)
		Supervised (physics)	39.39 ± 4.44	0.975 ± 0.02	Proposed, Supervised
	v_e	Proposed	**43.41** ± 4.01	**0.986** ± 0.01	Supervised, Supervised (physics)
		Supervised	43.26 ± 4.08	0.985 ± 0.01	Proposed, Supervised (physics)
		Supervised (physics)	43.35 ± 4.04	0.986 ± 0.01	Proposed, Supervised
MCI Data (Patlak)	K^{trans}	Proposed	**40.50** ± 3.91	0.977 ± 0.02	Supervised, Supervised (physics)
		Supervised	40.12 ± 2.86	0.976 ± 0.02	Proposed, Supervised (physics)
		Supervised (physics)	40.25 ± 2.80	**0.978** ± 0.01	Proposed, Supervised
	v_p	Proposed	**37.96** ± 3.45	**0.958** ± 0.03	Supervised, Supervised (physics)
		Supervised	37.62 ± 3.56	0.952 ± 0.04	Proposed, Supervised (physics)
		Supervised (physics)	37.61 ± 3.62	0.951 ± 0.04	Proposed, Supervised

Table 2

PK parameter values in tumor regions of simulated 'Tumor Data' (n=11) (average ± standard deviation).

PK parameter	Method	Brain Metastasis (n=5)	Glioma (n=3)	Meningioma (n=3)	Total (n=11)	Different from ($P < 0.05$) (Total (n=11))
K^{trans} (min^{-1})	LLS	0.017 ± 0.019	0.011 ± 0.002	0.012 ± 0.003	0.014 ± 0.013	–
	Proposed	0.009 ± 0.012	0.010 ± 0.003	0.029 ± 0.020	0.015 ± 0.016	–
	Supervised	0.017 ± 0.025	0.014 ± 0.006	0.037 ± 0.020	0.022 ± 0.022	–
	Supervised (physics)	0.016 ± 0.023	0.014 ± 0.006	0.030 ± 0.016	0.019 ± 0.019	–
v_p (×10^{-2})	LLS	0.451 ± 0.487	0.258 ± 0.020	0.708 ± 0.332	0.469 ± 0.407	Supervised
	Proposed	0.374 ± 0.276	0.435 ± 0.112	0.934 ± 0.597	0.543 ± 0.440	Supervised
	Supervised	0.406 ± 0.211	0.619 ± 0.223	0.871 ± 0.315	0.591 ± 0.313	LLS, Proposed
	Supervised (physics)	0.331 ± 0.131	0.524 ± 0.210	0.569 ± 0.141	0.449 ± 0.192	–
v_e (×10^{-2})	LLS	6.876 ± 7.094	1.391 ± 0.420	3.987 ± 1.379	4.592 ± 5.358	Supervised (physics)
	Proposed	1.010 ± 1.208	1.503 ± 0.494	3.915 ± 2.207	1.937 ± 1.889	Supervised
	Supervised	2.121 ± 2.945	1.988 ± 0.705	8.033 ± 4.804	3.697 ± 4.174	Proposed, Supervised (physics)
	Supervised (physics)	1.349 ± 1.894	1.131 ± 0.361	5.723 ± 3.992	2.483 ± 3.156	LLS, Supervised

5.2. In vivo studies

Qualitative results. Fig. 9 presents the estimated PK parameters of the in vivo 'Tumor Data' using both the traditional LLS method and deep learning methods. This specific example highlights the superior reliability of the proposed method compared to the LLS method. The tumor type depicted in Fig. 9 is a meningioma, where known characteristics include relatively higher K^{trans} and v_p values (Bazyar et al., 2016).

In Fig. 9, the LLS method exhibits elevated v_p values, while the K^{trans} estimated by the LLS method is relatively low in the tumor region. Moreover, the LLS method shows notably low K^{trans} and v_e values in normal brain tissues, a result attributed to the inappropriate

Table 3

PK parameter values in hippocampus regions of simulated 'MCI Data' (n=31) (average ± standard deviation).

| PK parameter	Method	Average ± Standard deviation	Different from ($P < 0.05$)
K^{trans} (min^{-1}) ($\times 10^{-2}$)	LLS	0.035 ± 0.035	Supervised, Supervised (physics)
	Proposed	0.039 ± 0.042	Supervised, Supervised (physics)
	Supervised	0.046 ± 0.039	LLS, Proposed
	Supervised (physics)	0.047 ± 0.042	LLS, Proposed
v_p ($\times 10^{-2}$)	LLS	1.748 ± 1.336	Supervised, Supervised (physics)
	Proposed	2.327 ± 0.926	Supervised (physics)
	Supervised	2.532 ± 0.717	LLS
	Supervised (physics)	2.648 ± 0.685	LLS, Proposed
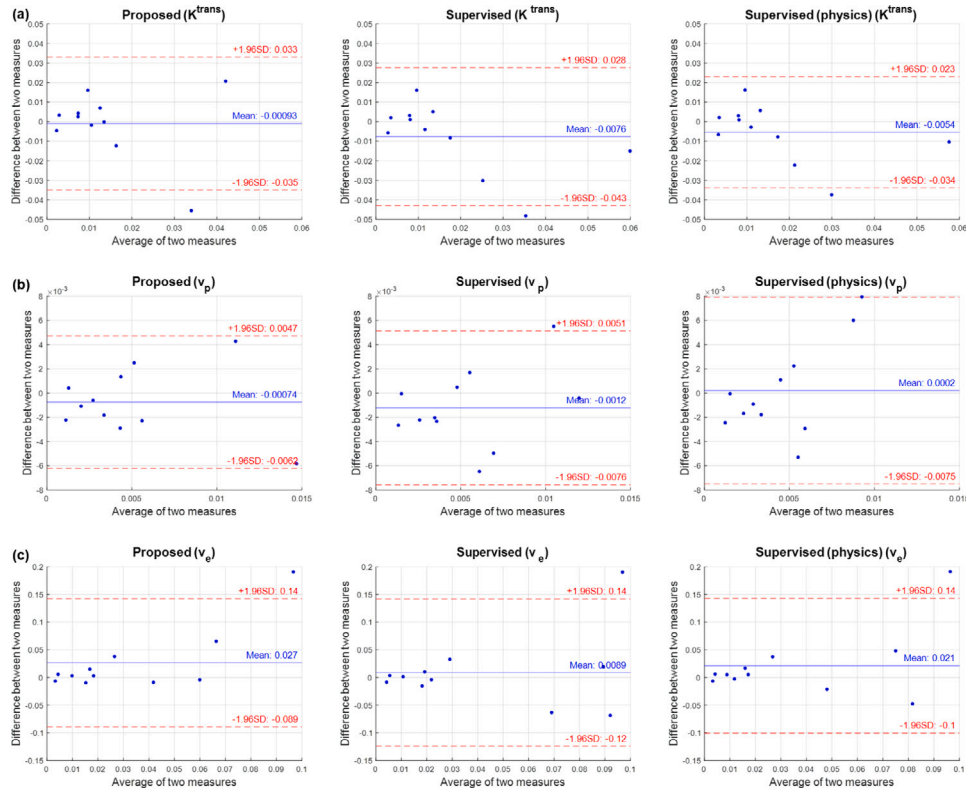
Fig. 7. Bland–Altman plots illustrating the agreement between PK parameter values in the tumor regions of simulated 'Tumor Data'. The PK parameter values of the LLS method correspond to the ground truth. (a) K^{trans}, (b) v_p, and (c) v_e, respectively.

AIF used for the LLS method, as depicted in Fig. 10(a). Consequently, the K^{trans} and v_e values from the LLS method are notably low not only in normal tissue but also in the tumor region.

Previous studies (Cha et al., 2006; Bazyar et al., 2016) have demonstrated that K^{trans} in meningioma is expected to increase due to heightened permeability resulting from the loss of the blood–brain barrier. Therefore, in Fig. 9, the K^{trans} and v_p values from the proposed method offer more reliable results that capture the biological features of the tumor.

On the other hand, the supervised learning approach yields excessive v_e values compared to other methods. When the supervised learning approach is augmented with the physics model, it achieves improved estimation results for v_e compared to standard supervised learning. Nonetheless, the supervised learning with the physics model exhibits an abnormally low estimation of v_p in the tumor region, in contrast to the v_p estimated by the LLS method or our proposed method.

In Fig. 9, high v_p values are observed in some thin sulci regions. The high v_p values observed in the thin sulci can be attributed to a partial volume effect, where the values of the gray matter on both sides of the gyrus overlap. Therefore, the occurrence of high v_p values in the sulci region aligns with normal expectations.

AIF estimation. Table 4 displays the results of the AIF quality evaluation. As depicted in Table 4, our proposed method consistently delivers
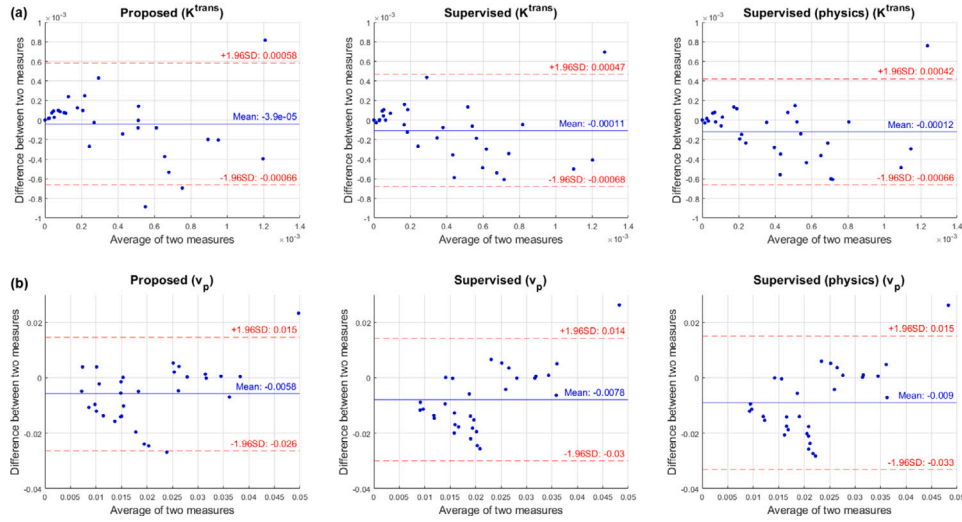
Fig. 8. Bland–Altman plots illustrating the agreement between PK parameter values in the hippocampus regions of simulated 'MCI Data'. The PK parameter values of the LLS method correspond to the ground truth. (a) K^{trans}, and (b) v_p, respectively.
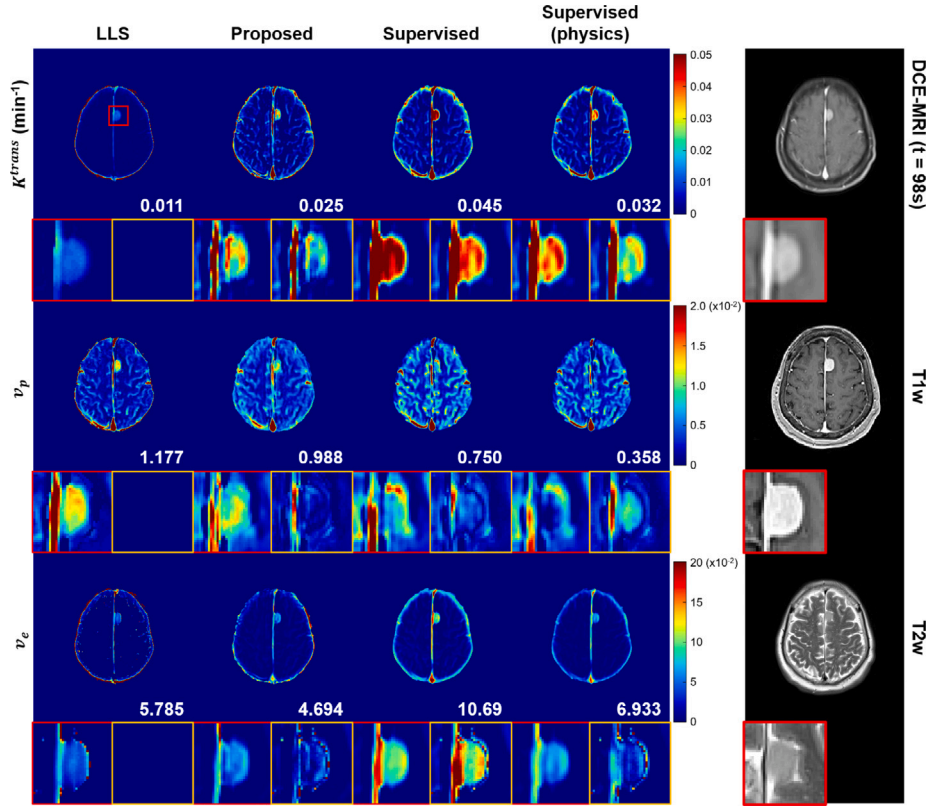

Fig. 9. Results of PK parameter estimation methods for in vivo 'Tumor Data' (eTofts model). Images outlined in red represent enlarged images that emphasize the tumor (meningioma) region. Images outlined in yellow are difference maps illustrating the absolute variance between each image and the LLS image. The difference maps are scaled by a factor of 2 to enhance visibility. The number in the bottom right corner indicates the mean parameter values at the tumor region. The scale bar for each parameter and difference map is located to the right of the parameter maps. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

higher-quality AIF estimations in cases where the AIF derived from the software exhibits poor quality.

Fig. 10 compares the AIFs by the software and our proposed method. AIFs in Fig. 10(a) are from the same patient in Fig. 9. In Fig. 10(a), our method provides more accurate AIF estimations compared to the NordicICE software. In particular, the CA concentration in the AIF generated by the proposed method decreases to approximately 1 mM after a few minutes. Conversely, the AIF obtained from the software exhibits

a CA concentration surpassing 2 mM, even after the passage of a few minutes, which could be considered an inappropriate AIF estimation. This inappropriate estimation of AIF contributes to inaccuracies in the PK parameter estimation results for the LLS method, as depicted in Fig. 9. While it is noted that there are cases (Fig. 10(b)) where the proposed method yields poor results compared to the software, such occurrences were infrequent.
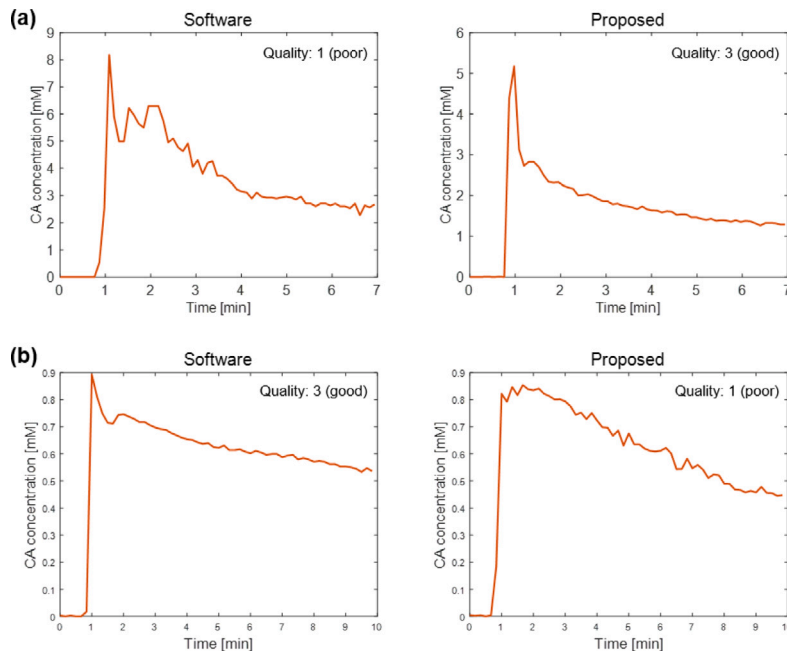
Fig. 10. Examples of the estimated AIFs generated by the NordicICE software and proposed method. Both AIFs are from the same in vivo data. The corresponding quality scores by radiologists are indicated at the top right of each graph. (a) An example where the AIF from the proposed method was better than the AIF from the NordicICE software. (b) An infrequent example where the AIF from the NordicICE software was better than the AIF from the proposed method.

Table 4
AIF quality evaluation results of the software and proposed method using in vivo 'Tumor Data' and in vivo 'MCI Data'.

Method	Software			
	Quality	Poor (1)	Fair (2)	Good (3)
Proposed	Poor (1)	0	2	2
	Fair (2)	3	4	15
	Good (3)	1	3	24

6. Discussion

6.1. Comparison with other methods

In this paper, we have proposed a novel unpaired deep learning method, based on physics-driven CycleGAN, for pharmacokinetic estimation. Our findings demonstrate that the proposed CycleGAN method accurately captures tumor characteristics and provides estimation results comparable to the LLS method while reducing computational complexity. However, supervised learning methods, even when incorporating the physics model, tend to yield exaggerated values for K^{trans} and v_e and exhibit limitations in accurately estimating v_p in the tumor region. Furthermore, our CycleGAN approach shows consistent and reliable results when applied to the Patlak model. These results highlight the robustness and applicability of our CycleGAN method for accurate pharmacokinetic parameter estimation.

Additionally, our evaluation of AIF quality demonstrates the superiority of our proposed method. In cases where the software generates AIFs of poor quality, our approach consistently delivers higher-quality estimations. This underscores the effectiveness of our method in overcoming limitations associated with suboptimal AIF data. Moreover, the AIF comparison in Fig. 10 provides further support, clearly indicating that our proposed method outperforms the software by providing more accurate AIF estimations. These results emphasize the value of our approach in obtaining reliable and high-quality AIFs, which contribute to improved pharmacokinetic parameter estimation and enhance the overall performance of the model.

Overall, our proposed method presents a promising approach for pharmacokinetic parameter estimation, offering improved accuracy, reduced computational complexity, and the ability to overcome limitations associated with suboptimal AIF data. Further advancements in training data quality and AIF estimation techniques will enhance our method's performance and applicability.

6.2. Limitations & future works

Despite successfully demonstrating accurate PK parameter estimation with reduced runtime and comparable performance to existing methods, our proposed method has certain limitations. For example, it is important to note that our method still requires PK parameter maps estimated by NLLS or LLS methods for training, although the training data need not be paired. Therefore, the performance of our proposed method may be affected if the quality of the training data is not optimal. Furthermore, while our method generally demonstrated good or fair AIF estimation results, there were infrequent instances where the proposed method did not yield optimal AIF. In our future works, we aspire to enhance the quality of AIF estimated by our method.

In Figs. 7 and 8, the limits of agreement appear large. The observed largeness in limits of agreement is influenced by the fact that standard deviations of the differences tend to increase with the mean of two measures. In the context of PK parameters, which typically have small values, the model may more easily predict compared to large ones, contributing to the observed discrepancy. Additionally, the discrepancy in parameter value range between the training and test sets may contribute to the observed trends in Figs. 7 and 8. To address this, future work will focus on training the model with a more representative dataset with a wider range of PK parameter values.

In Fig. 10, it is observed that the CA concentration remains high after a few minutes. While previous works (Nguyen et al., 2013; Keil et al., 2017) have indicated that CA concentration can exceed 1 mM after a few minutes depending on the data source, this could also originated from inherent errors in commercial AIF measurement software. Given that our proposed method is designed and trained to reproduce AIF data from the software, any inaccuracies in the software could influence the performance of our method.

Moreover, it is essential to re-train our model whenever there is a modification in the imaging sequence of the data or the TK model. This necessity arises due to the inherent limitation of CNNs, which can only handle data with a fixed number of channels. Consequently, existing models become incompatible if alterations in the imaging sequence lead to changes in the number of time scans. While one potential strategy for adapting to changes in the number of channels involves resampling the input DCE-MR image, this approach may not yield accurate results, particularly when there are differences in physics parameters such as the time interval. Additionally, changes in the TK model accompany the variation in the number of PK parameters. To address this, the network requires explicit information on which TK model to apply, enabling its versatility across multiple TK models. In our future work, we aim to develop an integrated model capable of accommodating multiple datasets, providing a more robust and adaptable solution.

In our upcoming studies, we plan to clinically evaluate our method by incorporating the following evaluation methods and conducting additional studies. Firstly, a promising approach involves assessing the diagnostic accuracy of Ktrans values through external validation by categorizing brain tumors into low and high grades. Additionally, for conditions such as dementia, the method could be tested with additional data. Drawing upon existing research demonstrating changes in blood–brain barrier permeability in the hippocampus or alterations in choroid plexus permeability in Alzheimer's disease compared to mild cognitive impairment or normal conditions, diagnostic accuracy could be compared to assess if the proposed model enhances diagnostic capabilities. Moreover, considering the improved K^{trans} map demonstrated by the proposed method, particularly in Fig. 9, the potential application for mapping in stereotactic radiosurgery is noteworthy.

7. Conclusion

This paper presents a novel unpaired deep learning method for PK parameter estimation utilizing a physics-driven CycleGAN, based on the physical model of PK parameters. Our CycleGAN architecture, derived from optimal transport theory, is specifically designed for the known physics model, resulting in a single generator and discriminator. Despite being trained with unpaired datasets, our method consistently produces reliable PK parameter maps compared to the LLS method and other supervised methods. Moreover, our proposed method demonstrates superior performance in quantitative evaluations when compared to alternative approaches. In addition, our proposed method also demonstrates superior performance in estimating the arterial input function, providing more accurate and reliable AIF estimations when the quality of AIFs estimated by existing methods is poor. We believe that our method can provide valuable insights and accurate PK parameter estimation even in scenarios where paired data is unavailable or missing.

CRediT authorship contribution statement

Gyutaek Oh: Writing – review & editing, Writing – original draft, Visualization, Methodology, Investigation, Formal analysis, Conceptualization. **Yeonsil Moon:** Data curation. **Won-Jin Moon:** Writing – review & editing, Validation, Supervision, Project administration, Formal analysis, Data curation, Conceptualization. **Jong Chul Ye:** Writing – review & editing, Supervision, Project administration, Methodology, Conceptualization.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

Armitage, P.A., Farrall, A.J., Carpenter, T.K., Doubal, F.N., Wardlaw, J.M., 2011. Use of dynamic contrast-enhanced MRI to measure subtle blood–brain barrier abnormalities. Magn. Reson. Imaging 29 (3), 305–314.

Bazyar, S., Ramalho, J., Eldeniz, C., An, H., Lee, Y.Z., 2016. Comparison of cerebral blood volume and plasma volume in untreated intracranial tumors. PLoS One 11 (9), e0161807.

Branch, M.A., Coleman, T.F., Li, Y., 1999. A subspace, interior, and conjugate gradient method for large-scale bound-constrained minimization problems. SIAM J. Sci. Comput. 21 (1), 1–23.

Cha, S., Yang, L., Johnson, G., Lai, A., Chen, M.-H., Tihan, T., Wendland, M., Dillon, W., 2006. Comparison of microvascular permeability measurements, Ktrans, determined with conventional steady-state T1-weighted and first-pass T2*-weighted MR imaging methods in gliomas and meningiomas. Am. J. Neuroradiol. 27 (2), 409–417.

Chao, S.-L., Metens, T., Lemort, M., 2017. TumourMetrics: a comprehensive clinical solution for the standardization of DCE-MRI analysis in research and routine use. Quant. Imag. Med. Surgery 7 (5), 496.

Choi, K.S., You, S.-H., Han, Y., Ye, J.C., Jeong, B., Choi, S.H., 2020. Improving the reliability of pharmacokinetic parameters at dynamic contrast-enhanced MRI in astrocytomas: A deep learning approach. Radiology 297 (1), 178–188.

Chung, C.Y., Hu, R., Peterson, R.B., Allen, J.W., 2021. Automated processing of head CT perfusion imaging for ischemic stroke triage: A practical guide to quality assurance and interpretation. Am. J. Roentgenol. 217 (6), 1401–1416.

Di Giovanni, P., Azlan, C., Ahearn, T.S., Semple, S., Gilbert, F.J., Redpath, T.W., 2009. The accuracy of pharmacokinetic parameter measurement in DCE-MRI of the breast at 3 T. Phys. Med. Biol. 55 (1), 121.

Fang, K., Wang, Z., Li, Z., Wang, B., Han, G., Cheng, Z., Chen, Z., Lan, C., Zhang, Y., Zhao, P., 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. Imag. 53 (6), 1898–1910.

Flouri, D., Lesnic, D., Sourbron, S.P., 2016. Fitting the two-compartment model in DCE-MRI by linear inversion. Magn. Reson. Med. 76 (3), 998–1006.

Heye, A.K., Thrippleton, M.J., Armitage, P.A., Hernández, M.d.C.V., Makin, S.D., Glatz, A., Sakka, E., Wardlaw, J.M., 2016. Tracer kinetic modelling for DCE-MRI quantification of subtle blood–brain barrier permeability. Neuroimage 125, 446–455.

Kang, K.M., Choi, S.H., Chul-Kee, P., Kim, T.M., Park, S.-H., Lee, J.H., Lee, S.-T., Hwang, I., Yoo, R.-E., Yun, T.J., et al., 2021. Differentiation between glioblastoma and primary CNS lymphoma: application of DCE-MRI parameters based on arterial input function obtained from DSC-MRI. Eur. Radiol. 31 (12), 9098–9109.

Keil, V.C., Mädler, B., Gieseke, J., Fimmers, R., Hattingen, E., Schild, H.H., Hadizadeh, D.R., 2017. Effects of arterial input function selection on kinetic parameters in brain dynamic contrast-enhanced MRI. Magn. Reson. Imaging 40, 83–90.

Lim, S., Park, H., Lee, S.-E., Chang, S., Sim, B., Ye, J.C., 2020. Cyclegan with a blur kernel for deconvolution microscopy: Optimal transport geometry. IEEE Trans. Comput. Imag. 6, 1127–1138.

Mao, X., Li, Q., Xie, H., Lau, R.Y., Wang, Z., Paul Smolley, S., 2017. Least squares generative adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2794–2802.

Martin Arjovsky, S., Bottou, L., 2017. Wasserstein generative adversarial networks. In: Proceedings of the 34 Th International Conference on Machine Learning, Sydney, Australia.

Mouridsen, K., Christensen, S., Gyldensted, L., Østergaard, L., 2006. Automatic selection of arterial input function using cluster analysis. Magn. Reson. Med. Official J. Int. Soc. Magn. Reson. Med. 55 (3), 524–531.

Nguyen, V.L., Kooi, M.E., Backes, W.H., van Hoof, R.H., Saris, A.E., Wishaupt, M.C., Hellenthal, F.A., van der Geest, R.J., Kessels, A.G., Schurink, G.W.H., et al., 2013. Suitability of pharmacokinetic models for dynamic contrast-enhanced mri of abdominal aortic aneurysm vessel wall: a comparison. PLoS One 8 (10), e75173.

Padhani, A.R., 2002. Dynamic contrast-enhanced MRI in clinical oncology: current status and future directions. Magn. Reson. Med. Official J. Int. Soc. Magn. Reson. Med. 16 (4), 407–422.

Patlak, C.S., Blasberg, R.G., Fenstermacher, J.D., 1983. Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. J. Cerebral Blood Flow Metabol. 3 (1), 1–7.

Sim, B., Oh, G., Kim, J., Jung, C., Ye, J.C., 2020. Optimal transport driven cyclegan for unsupervised learning in inverse problems. SIAM J. Imaging Sci. 13 (4), 2281–2306.

Smith, S.M., 2002. Fast robust automated brain extraction. Hum. Brain Mapp. 17 (3), 143–155.

Sourbron, S.P., Buckley, D.L., 2013. Classic models for dynamic contrast-enhanced MRI. NMR Biomed. 26 (8), 1004–1027.

Szomolanyi, P., Rohrer, M., Frenzel, T., Noebauer-Huhmann, I.M., Jost, G., Endrikat, J., Trattnig, S., Pietsch, H., 2019. Comparison of the relaxivities of macrocyclic gadolinium-based contrast agents in human plasma at 1.5, 3, and 7 T, and blood at 3 T. Investigat. Radiol. 54 (9), 559.

Thrippleton, M.J., Backes, W.H., Sourbron, S., Ingrisch, M., van Osch, M.J., Dichgans, M., Fazekas, F., Ropele, S., Frayne, R., van Oostenbrugge, R.J., et al., 2019. Quantifying blood-brain barrier leakage in small vessel disease: review and consensus recommendations. Alzheimer's Dementia 15 (6), 840–858.

Tofts, P.S., Brix, G., Buckley, D.L., Evelhoch, J.L., Henderson, E., Knopp, M.V., Larsson, H.B., Lee, T.-Y., Mayr, N.A., Parker, G.J., et al., 1999. Estimating kinetic parameters from dynamic contrast-enhanced T1-weighted MRI of a diffusable tracer: standardized quantities and symbols. Magn. Reson. Med. Official J. Int. Soc. Magn. Reson. Med. 10 (3), 223–232.

Ulas, C., Das, D., Thrippleton, M.J., Valdes Hernandez, M.d.C., Armitage, P.A., Makin, S.D., Wardlaw, J.M., Menze, B.H., 2019. Convolutional neural networks for direct inference of pharmacokinetic parameters: Application to stroke dynamic contrast-enhanced MRI. Front. Neurol. 9, 1147.

Ulyanov, D., Vedaldi, A., Lempitsky, V., 2016. Instance normalization: The missing ingredient for fast stylization. arXiv preprint arXiv:1607.08022.

Van De Haar, H.J., Burgmans, S., Jansen, J.F., Van Osch, M.J., Van Buchem, M.A., Muller, M., Hofman, P.A., Verhey, F.R., Backes, W.H., 2016. Blood-brain barrier leakage in patients with early Alzheimer disease. Radiology 281 (2), 527–535.

Verma, S., Turkbey, B., Muradyan, N., Rajesh, A., Cornud, F., Haider, M.A., Choyke, P.L., Harisinghani, M., 2012. Overview of dynamic contrast-enhanced MRI in prostate cancer diagnosis and management. Am. J. Roentgenol. 198 (6), 1277–1288.

Villani, C., 2008. Optimal Transport: Old and New, vol. 338, Springer Science & Business Media.

Zhu, J.-Y., Park, T., Isola, P., Efros, A.A., 2017. Unpaired image-to-image translation using cycle-consistent adversarial networks. In: Proceedings of the IEEE International Conference on Computer Vision. pp. 2223–2232.