

Direct Estimation of Pharmacokinetic Parameters from DCE-MRI using Deep CNN with Forward Physical Model Loss

Cagdas Ulas¹, Giles Tetteh¹, Michael J. Thrippleton², Paul A. Armitage⁴, Stephen D. Makin², Joanna M. Wardlaw², Mike E. Davies³, Bjoern H. Menze¹

¹ Department of Computer Science, Technische Universität München, Germany

² Department of Neuroimaging Sciences, University of Edinburgh, UK

³ Institute for Digital Communications, University of Edinburgh, UK

⁴ Department of Cardiovascular Sciences, University of Sheffield, UK

Abstract. Dynamic contrast-enhanced (DCE) MRI is an evolving imaging technique that provides a quantitative measure of pharmacokinetic (PK) parameters in body tissues, in which series of T_1 -weighted images are collected following the administration of a paramagnetic contrast agent. Unfortunately, in many applications, conventional clinical DCE-MRI suffers from low spatiotemporal resolution and insufficient volume coverage. In this paper, we propose a novel deep learning based approach to directly estimate the PK parameters from undersampled DCE-MRI data. Specifically, we design a custom loss function where we incorporate a forward physical model that relates the PK parameters to corrupted image-time series obtained due to subsampling in k-space. This allows the network to directly exploit the knowledge of true contrast agent kinetics in the training phase, and hence provide more accurate restoration of PK parameters. Experiments on clinical brain DCE datasets demonstrate the efficacy of our approach in terms of fidelity of PK parameter reconstruction and significantly faster parameter inference compared to a model-based iterative reconstruction method.

1 Introduction

Dynamic contrast-enhanced (DCE) MRI involves the administration of a T_1 -shortening Gadolinium-based contrast agent (CA), followed by the acquisition of successive T_1 -weighted images as the contrast bolus enters and subsequently leaves the organ [9]. In DCE-MRI, changes in CA concentration are derived from changes in signal intensity over time, then regressed to estimate pharmacokinetic (PK) parameters related to vascular permeability and tissue perfusion [6]. Since perfusion and permeability are typically affected in the presence of vascular and cellular irregularities, DCE imaging has been considered as a promising tool for clinical diagnostics of brain tumours, multiple sclerosis lesions, and neurological disorders where disruption of blood-brain barrier (BBB) occurs. [7,4].

Despite its effectiveness in quantitative assessment of microvascular properties, conventional DCE-MRI is challenged by suboptimal image acquisition that

severely restricts the spatiotemporal resolution and volume coverage [2,3]. The shortest possible scanning time often leads to limited spatial resolution hampering detection of small image features and accurate tumor boundaries. Low temporal resolution hinders accurate fitting of PK parameters. Furthermore, volume coverage is usually inadequate to cover the known pathology, for instance in the case multiple metastatic lesions [3]. Facing such severe constraints, DCE imaging can significantly benefit from undersampled acquisitions.

So far, existing works in [6,2,11] have proposed compressed sensing and parallel imaging based reconstruction schemes to accelerate DCE-MRI acquisitions, mainly targeting to achieve better spatial resolution and volume coverage while retaining the same temporal resolution. These methods are referred to as indirect methods [3] because they are based on the reconstruction of dynamic DCE image series first, followed by a separate step for fitting the PK parameters on a voxel-by-voxel level using a tracer kinetic model [9]. More recently, a model-based direct reconstruction model [3] has been proposed to directly estimate PK parameters from undersampled (k,t) space data. The direct reconstruction method generally poses the estimation of PK maps as an error minimization problem. This approach has been shown to produce superior PK parameter maps and allows for higher acceleration compared to indirect methods. However, the main drawback of this method is that parameter reconstruction of an entire volume requires considerably high computation time.

Motivated by the recent advances of deep learning in medical imaging, in this paper, we present a novel deep learning based approach to directly estimate PK parameters from undersampled DCE-MRI data. First, our proposed network takes the corrupted image-time series as input and *residual* parameter maps, which represent deviations from a kinetic model fitting on fully-sampled image-time series, as output, and aims at learning a nonlinear mapping between them. Our motivation for learning the *residual* PK maps is based on the observation that residual maps are more sparse and topologically less complex compared to target parameter maps. Second, we propose the *forward physical model loss*, a custom loss function in which we exploit the physical relation between true contrast agent kinetics and measured time-resolved DCE signals when training our network. Third, we validate our method experimentally on human *in vivo* brain DCE-MRI dataset. We demonstrate the superior performance of our method in terms of parameter reconstruction accuracy and significantly faster estimation of parameters during testing, taking approximately 1.5 seconds on an entire 3D test volume. To the best of our knowledge, we present the first work leveraging the machine learning algorithms – specifically deep learning – to directly estimate PK parameters from undersampled DCE-MRI time-series.

2 Methods

We treat the parameter inference from undersampled data in DCE imaging as a mapping problem between the corrupted intensity-time series and *residual* parameter maps where the underlying mapping is learned using deep convolutional neural networks (CNNs). We provide a summary of general tracer kinetic

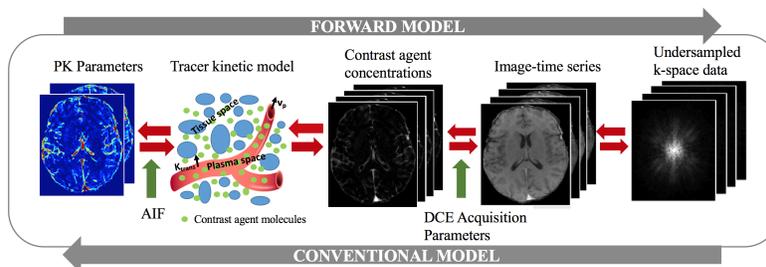


Fig. 1. Computational steps in the forward model and the conventional pipeline of PK parameter estimation in DCE-MRI.

models applied in DCE-MRI in Sec. 2.1, formulate the forward physical model relating the PK parameters to undersampled data in Sec. 2.2, finally describe our proposed deep learning methodology for PK parameter inference in Sec. 2.3.

2.1 Tracer Kinetic Modeling in DCE-MRI

Tracer kinetic modeling aims at providing a link between the tissue signal enhancement and the physiological or so-called pharmacokinetic parameters, including the fractional plasma volume (v_p), the fractional interstitial volume (v_e), and the volume transfer rate (K^{trans}) at which contrast agent (CA) is delivered to the extravascular extracellular space (EES). One of the well-established tracer kinetic models is known as Patlak model [8]. This model describes a highly perfused two compartment tissue, ignoring backflux from the EES into the blood plasma compartment. The CA concentration in the tissues is determined by,

$$C(\mathbf{r}, t) = v_p(\mathbf{r})C_p(t) + K^{\text{trans}}(\mathbf{r}) \int_0^t C_p(\tau) d\tau, \quad (1)$$

where $\mathbf{r} \in (x, y, z)$ represent image domain spatial coordinates, $C(\mathbf{r}, t)$ is the CA concentration over time, and $C_p(t)$ denotes the arterial input function (AIF) which is usually measured from voxels in a feeding artery.

In this work, we specifically employ the Patlak model for tracer pharmacokinetic modeling and estimation of ground truth tissue parameters. This model is a perfect match for our DCE dataset because it is often applied when the temporal resolution is too low to measure the cerebral blood flow, and it has been commonly used to measure the BBB leakage with DCE-MRI in acute brain stroke and dementia [4,9]. An attractive feature of Patlak model is that the model equation in (1) can be linearized and fitted using linear least squares which has a closed-form solution, hence parameter estimation is fast [9].

2.2 Forward Physical Model: From PK Parameters to Undersampled Data

Figure 1 depicts the conventional and forward model approaches relating the PK parameter estimation to undersampled or fully-sampled k-space data, and vice versa. For direct estimation of PK parameters from the measured k-space data, as proposed in [1,3], a forward model can be formulated by inverting the steps in the conventional model as follows:

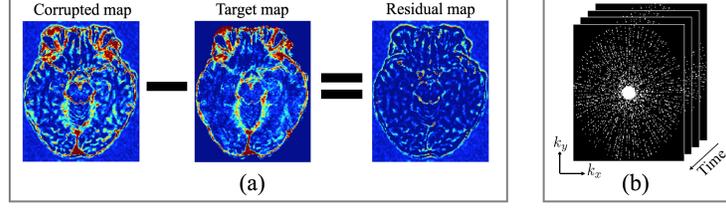


Fig. 2. (a) The relation between a corrupted (θ_u), target (θ_t) and residual (θ_r) PK maps, (b) Exemplary golden-angle sampling scheme in the k_x - k_y plane through time.

1. Given the sets of PK parameter pairs ($K^{\text{trans}}(\mathbf{r}), v_p(\mathbf{r})$) and arterial input function $C_p(t)$, CA concentration curves over time $C(\mathbf{r}, t)$ are estimated using the Patlak model equation in (1).
2. Dynamic DCE image series $S(\mathbf{r}, t)$ are converted to $C(\mathbf{r}, t)$ through the steady-state spoiled gradient echo (SGPR) signal equation [3], given by

$$S(\mathbf{r}, t) = \frac{M_0(\mathbf{r})\sin\alpha(1 - e^{-(K+L)})}{1 - \cos\alpha e^{-(K+L)}} + \left(S(\mathbf{r}, 0) - \frac{M_0(\mathbf{r})\sin\alpha(1 - e^{-K})}{1 - \cos\alpha e^{-K}} \right) \quad (2)$$

where $K = T_R/T_{10}(\mathbf{r})$, $L = r_1 C(\mathbf{r}, t)T_R$, T_R is the repetition time, α is the flip angle, r_1 is the contrast agent relaxivity taken as $4.2 \text{ s}^{-1}\text{mM}^{-1}$, $S(\mathbf{r}, 0)$ is the baseline (pre-contrast) image intensity, and $T_{10}(\mathbf{r})$ and $M_0(\mathbf{r})$ are respectively the T_1 relaxation and equilibrium longitudinal magnetization that are calculated from a pre-contrast T_1 mapping acquisition.

3. The undersampled raw (k,t)-space data $S(\mathbf{k}, t)$ can be related to $S(\mathbf{r}, t)$ for a single-coil data by an undersampling fast Fourier transform (FFT), F_u ,

$$S(\mathbf{k}, t) = F_u S(\mathbf{r}, t), \quad (3)$$

where $\mathbf{k} \in (k_x, k_y, k_z)$ represents k-space coordinates.

By simply integrating the three computation steps in (1-3), we can form a single function f_m modeling the signal evolution in (k-t) space given the PK maps $\theta = \{K^{\text{trans}}(\mathbf{r}), v_p(\mathbf{r})\}$, as $S(\mathbf{k}, t) = f_m(\theta; \xi)$, where ξ denotes all the predetermined acquisition parameters as mentioned above.

Given the undersampled (k,t)-space data $S(\mathbf{k}, t)$, the corrupted image series $S_u(\mathbf{r}, t)$ can be obtained by applying IFFT to $S(\mathbf{k}, t)$, i.e. $S_u(\mathbf{r}, t) = F_u^T S(\mathbf{k}, t)$. We further define a new function \tilde{f}_m that integrates only the first two computation steps (1-2) to compute the dynamic DCE image series. We will incorporate \tilde{f}_m in our custom loss function that will be explained in the following section.

2.3 PK Parameter Inference via Forward Physical Model Loss

Formulation. We hypothesize that a direct inversion between corrupted PK parameter maps θ_u and $S_u(\mathbf{r}, t)$ is available through forward model, i.e., $S_u(\mathbf{r}, t) = \tilde{f}_m(\theta_u)$. However, this cannot provide yet sufficiently accurate estimate of target parameter maps θ_t obtained from fully-sampled data $S(\mathbf{r}, t)$. To this end, we estimate a correction or residual map θ_r from the available signal $S_u(\mathbf{r}, t)$ satisfying $\theta_r = \theta_u - \theta_t$. As shown in Fig. 2-(a), we observe that *residual* PK maps involve more sparse representations and exhibit spatially less varying structures inside

the brain. The task of learning a residual mapping was shown to be much easier and effective than the original mapping [10]. Following the same approach, we adopt the residual learning strategy using deep CNNs. Our CNN is trained to learn a mapping between $S_u(\mathbf{r}, t)$ and θ_r to output an estimate of residual maps $\tilde{\theta}_r$; $\tilde{\theta}_r = \mathcal{R}(S_u(\mathbf{r}, t) | \mathbf{W})$, where \mathcal{R} represents the forward mapping of the CNN parameterised by \mathbf{W} . The final parameter estimate is obtained via $\tilde{\theta}_t = \theta_u - \tilde{\theta}_r$.

Loss Function. We simultaneously seek the signal belonging to the corrected model estimates to be sufficiently close to true signal, i.e., $\mathbf{f}_m(\tilde{\theta}_t) \approx S(\mathbf{r}, t)$. Therefore, we design a custom loss function which requires solving the forward model in every iteration of the network training. We refer the resulting loss as *forward physical model loss*. Given a set of training samples \mathcal{D} of input-output pairs $(S_u(\mathbf{r}, t), \theta_r)$, we train a CNN model that minimizes the following loss,

$$\mathcal{L}(\mathbf{W}) = \sum_{(S_u(\mathbf{r}, t), \theta_r) \in \mathcal{D}} \lambda \|\theta_r - \tilde{\theta}_r\|_2^2 + (1 - \lambda) \|S(\mathbf{r}, t) - \tilde{\mathbf{f}}_m(\theta_u - \tilde{\theta}_r; \boldsymbol{\xi})\|_2^2, \quad (4)$$

where λ is a regularization parameter balancing the trade-off between the fidelity of the parameter and signal reconstruction. We emphasize that the second term in (4) allows the network to intrinsically exploit the underlying contrast agent kinetics in training phase.

Network Architecture. Figure 3 illustrates our network architecture. The network takes a 4D image-time series as input, where time frames are stacked as input channels. The first convolutional layer applies 3D filters to each channel individually to extract low-level temporal features which are aggregated over frames via learned filter weights to produce a single output per voxel. Following the first layer, inspired by the work on brain segmentation [5], our network consists of parallel dual pathways to efficiently capture multi-scale information. The local pathway at the top focuses on extracting details from the local vicinity while the global pathway at the bottom is designed to incorporate more contextual global information. The global pathway consists of 4 dilated convolutional layers with dilation factors of 2, 4, 8, 16, implying increased receptive field sizes. The filter size of each convolutional layer including dilated convolutions is $3 \times 3 \times 3$, and the rectified linear units (ReLU) activation is applied after each convolution. Local and global pathways are then concatenated to form a multi-scale feature set. Following this, 2 fully-connected layers are used to determine the best possible feature combination that can accurately map the input to output of the network. Finally, the last layer outputs the estimated residual maps.

3 Experiments and Results

Datasets. We perform experiments on fully-sampled DCE-MRI datasets acquired from three mild ischaemic stroke patients. DCE image series were acquired using a 1.5T clinical scanner with a 3D T1W spoiled gradient echo sequence (TR/TE = 8.24/3.1 ms, flip angle = 12°, FOV = 24 × 24 cm, matrix = 256 × 192, slice thickness = 4 mm, 73 sec temporal resolution, 21 dynamics). An intravenous bolus injection of 0.1 mmol/kg of gadoterate meglumine (Gd-DOTA) was administered simultaneously. The total acquisition time for DCE-MRI was

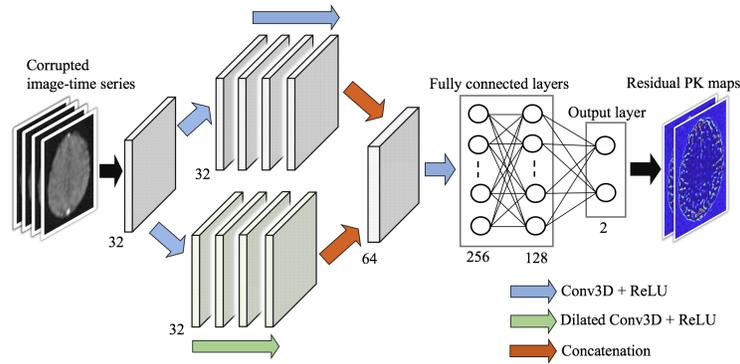


Fig. 3. The network architecture used for the estimation of residual PK maps. The number of filters and output nodes are provided at the bottom of each layer.

approximately 24 minutes. Two pre-contrast acquisitions were carried out at flip angles of 2° and 12° to calculate pre-contrast longitudinal relaxation times.

Preprocessing. Undersampling was retrospectively applied to the fully-sampled data in the k_x - k_y plane using a randomized golden-angle sampling pattern [12] over time (see Fig. 2-(b)) with a 10-fold undersampling factor. The pre-contrast first frame was fully sampled. Due to the low temporal resolution of our data, we estimated subject-specific vascular input functions (VIFs) extracted by averaging a few voxels located on the superior sagittal sinus where the inflow artefact was reduced compared to a feeding artery [4]. Data augmentation was employed by applying rigid transformations on image slices. We generated random 2D+t undersampling masks to be applied on the images of different orientations. This allows the network to learn diverse patterns of aliasing artifacts. All the subject’s data required for network training/testing were divided into non-overlapping 3D blocks of size $52 \times 52 \times 33$, resulting in 64 blocks per subject.

Experimental setup. All experiments were performed in a leave-one-subject-out fashion. The networks were trained using the Adam optimizer with a learning rate of 10^{-3} (using a decay rate of 10^{-4}) for 300 epochs and mini-batch size of 4. To demonstrate the advantage of the proposed method, we compare it with the state-of-the-art model-based iterative parameter reconstruction method using the MATLAB implementation provided by the authors [3]. We use the concordance correlation coefficient (CCC) and structured similarity metric (SSIM) metrics to quantitatively assess the PK parameter reconstruction, and peak signal-to-noise ratio (PSNR) metric to assess the image reconstruction. Experiments were run on a NVIDIA GeForce Titan Xp GPU with 12 GB RAM.

Results. Figure 4 shows the qualitative PK parameter reconstructions obtained from different methods using 10-fold undersampling. The results indicate that CNN- $\lambda = 0.5$ incorporating two loss terms simultaneously produces better maps and considerably higher SSIM score calculated with respect to fully-sampled PK maps. The model-based iterative reconstruction yields the PK maps where the artifacts caused by undersampling are still observable. In Fig. 5 we present the exemplary reconstructed images obtained by applying the operation \hat{f}_m to the

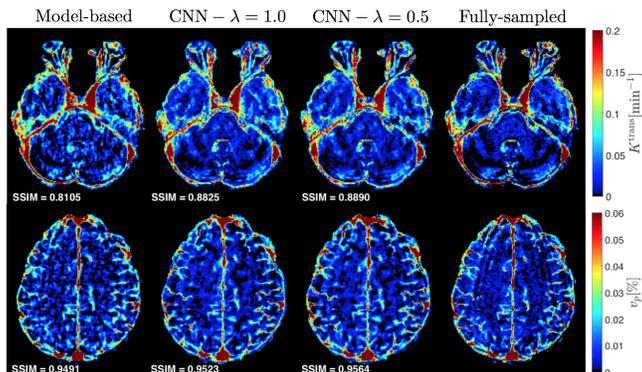


Fig. 4. Reconstructed PK parameter maps of two exemplary slices of a test subject with a 10-fold undersampling. Brain masks are applied to estimated maps. Our CNN model incorporating both loss terms ($\lambda = 0.5$) achieves the best parameter estimates. The resulting SSIM values are provided at the bottom-left corner of each map.

estimated PK maps. All the reconstruction approaches result in high quality images, however, the model-based reconstruction can better preserve the finer details. Unfortunately, our fully-sampled data suffer from Gibbs artifacts appearing as multiple parallel lines throughout the image. As marked by white arrows, our CNN method can significantly suppress these artifacts whereas they still appear in the image obtained by model-based iterative reconstruction. Finally, Fig. 6 demonstrates the quantitative results of parameter estimation and image reconstruction. The highest CCC and SSIM values for parameter estimation are achieved by our CNN model when both loss terms are incorporated with $\lambda = 0.3$ and $\lambda = 0.5$, yielding an average score of 0.88 and 0.92, respectively. The difference is statistically significant for both CCC ($p = 0.017$) and SSIM ($p = 0.0086$) when compared against model-based reconstruction. The model-based reconstruction performs the highest PSNR for image reconstruction, where it is followed by the proposed CNN with $\lambda = 0.3$. The difference between them is statistically significant with $p \ll 0.05$. The PSNR also shows a decreasing trend with increasing λ as expected.

We emphasize that the parameter inference of our method on a 3D test volume takes around 1.5 seconds while the model-based method requires around 95 minutes to reconstruct the same volume, enabling $\approx 4 \times 10^3$ faster computation.

4 Conclusion

We present a novel deep learning based framework for direct estimation of PK parameter maps from undersampled DCE image-time series. Specifically, we design a *forward physical model loss* function through which we exploit the physical model relating the contrast agent kinetics to the time-resolved DCE signals. Moreover, we utilize the residual learning strategy in our problem formulation. The experiments demonstrate that our proposed method can outperform the state-of-the-art model-based reconstruction method, and allow almost instantaneous inference of the PK parameters in the clinical workflow of DCE-MRI.

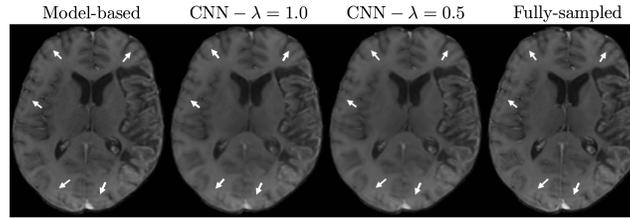


Fig. 5. Visual comparison of the image reconstruction results of an exemplary DCE slice. White arrows indicate a few regions where the Gibbs artifacts are observable. Our CNN model with both $\lambda = 0.5$ and 1.0 can significantly suppress the artifacts appearing in fully-sampled image and model-based reconstruction as well.

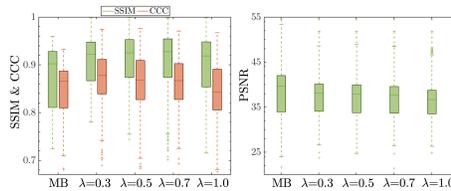


Fig. 6. Parameter estimation (SSIM & CCC) and image reconstruction (PSNR) performances calculated on all test slices for model-based (MB) reconstruction method and our proposed CNN model with different λ settings.

Acknowledgements. The research leading to these results has received funding from the European Unions H2020 Framework Programme (H2020-MSCA-ITN-2014) under grant agreement no 642685 MacSeNet.

References

1. Fang, R., et al.: Direct estimation of permeability maps for low-dose CT perfusion. In: IEEE ISBI. pp. 739–742 (April 2016)
2. Guo, Y., et al.: High-resolution whole-brain DCE-MRI using constrained reconstruction. *Med. Phys.* 43(5), 2013–2023 (2016)
3. Guo, Y., et al.: Direct estimation of tracer-kinetic parameter maps from highly undersampled brain dynamic contrast enhanced MRI. *MRM* 78(4) (2017)
4. Heye, A.K., et al.: Tracer kinetic modelling for DCE-MRI quantification of subtle bloodbrain barrier permeability. *NeuroImage* 125, 446 – 455 (2016)
5. Kamnitsas, K., et al.: Efficient multi-scale 3d CNN with fully connected CRF for accurate brain lesion segmentation. *Medical Image Analysis* 36, 61–78 (2017)
6. Lebel, R.M., et al.: Highly accelerated dynamic contrast enhanced imaging. *MRM* 71(2), 635–644 (2014)
7. O’Connor, J.P.B., et al.: Dynamic contrast-enhanced MRI in clinical trials of anti-vascular therapies. *Nat. Rev. Clin. Oncol.* 9(3), 167–77 (Mar 2012)
8. Patlak, C.S., et al.: Graphical evaluation of blood-to-brain transfer constants from multiple-time uptake data. *J. Cereb. Blood Flow Metab.* 3(1), 1–7 (1983)
9. Sourbron, S.P., Buckley, D.L.: Classic models for dynamic contrast-enhanced MRI. *NMR in Biomedicine* 26(8), 1004–1027 (2013)
10. Zhang, K., et al.: Beyond a gaussian denoiser: Residual learning of deep CNN for image denoising. *IEEE Trans. Image Process.* 26(7), 3142–3155 (July 2017)
11. Zhang, T., et al.: Fast pediatric 3d free-breathing abdominal dynamic contrast enhanced MRI with high spatiotemporal resolution. *JMRI* 41(2), 460–473 (2015)
12. Zhu, Y., et al.: GOCART: Golden-angle Cartesian randomized time-resolved 3D MRI. *Magn. Reson. Imag.* 34(7), 940 – 950 (2016)