Arterial input function and tracer kinetic model-driven network for rapid inference of kinetic maps in Dynamic Contrast-Enhanced MRI (AIF-TK-net)

Joseph Kettelkamp[1], Sajan Goud Lingala[1,2]
[1]Roy J Carver Department of Biomedical Engineering
[2]Department of Radiology
The University of Iowa, Iowa city

ABSTRACT

We propose a patient-specific arterial input function (AIF) and tracer kinetic (TK) model-driven network to rapidly estimate the extended Tofts-Kety kinetic model parameters in DCE-MRI. We term our network as AIF-TK-net, which maps an input comprising of an image patch of the DCE-time series and the patient-specific AIF to the output image patch of the TK parameters. We leverage the open-source NEURO-RIDER database of brain tumor DCE-MRI scans to train our network. Once trained, our model rapidly infers the TK maps of unseen DCE-MRI images on the order of a 0.34 sec/slice for a 256x256x65 time series data on a NVIDIA GeForce GTX 1080 Ti GPU. We show its utility on high time resolution DCE-MRI datasets where significant variability in AIFs across patients exists. We demonstrate that the proposed AIF-TK net considerably improves the TK parameter estimation accuracy in comparison to a network, which does not utilize the patient AIF.

1. INTRODUCTION

Dynamic contrast-enhanced magnetic resonance imaging (DCE-MRI) is a powerful quantitative technique to non-invasively characterize tumor/cancer biology and its response to treatment. In the brain, DCE-MRI characterizes the blood-brain barrier leakiness, which characterizes brain tumors [1]. DCE-MRI involves the intravenous administration of a contrast agent (e.g., Gadolinium) and continuous acquisition of images to track the passage of the contrast through the volume. Tracer-kinetic (TK) modeling of the enhancement kinetics of the contrast agent enables quantification of TK parameters such as K^{trans}, K^{ep} which are respectively the forward and backward volume transfer coefficient across the capillary endothelium, v_p: plasma volume, and F_p: plasma flow [2].

Various TK models have been proposed in DCE-MRI. These include linear models such as the 2-parameter Patlak model, non-linear models such as the 3-parameter extended Tofts-Kety (ETK) model, and the 4-parameter two-compartment exchange model [2]. The application at hand usually motivates the choice of model selection. For example, since the ETK model accounts for back-flux of contrast from the extravascular space to the plasma, it has been used in several brain tumor DCE-MRI studies (e.g., [3], [4]).

In this paper, we focus on the problem of improving the speed of non-linear ETK modeling in DCE-MRI. Classical approaches to estimate TK parameters in the ETK model employ iterative solvers to perform non-linear voxel-wise fitting. These include optimization methods such as the Newtonian solver [5], quasi Newtonian limited memory Broyden-Fletcher-Goldfarb-Shanno (L-BFGS) solver [6], or the non-linear conjugate gradient solver. While these are routinely used, they are challenged by long processing times (e.g., on the order of few hours for a single slice 256x256x65 DCE-time series data on a modern 32-core CPU). Recently, a convolutional neural network (CNN) based approach has shown to significantly improve the estimation time of ETK modeling (e.g., on the order of few seconds for a single DCE-slice) [7]. However, this work did not account for the arterial input function (AIF) in the network, partly due to the use of low time resolution (73 secs) and long scan time (~24 mins) DCE-MRI data. In this work, we improve the accuracy of TK model fitting of this network by incorporating patient-specific AIFs, which can be extracted from large vessels in the data (e.g., sagittal sinus). We term our network as AIF-TK-net, which maps an input comprising of an image patch of the DCE-time series and the patient-specific AIF to the output image patch of the TK parameters. We train the network using high time resolution (4.8 secs), short scan time (~5 minutes) brain tumor DCE-MRI scans from the open-source NEURO-RIDER database [8]. Once trained, our model rapidly infers the TK maps of unseen DCE-MRI images on the order of 0.34 secs/slice for a 256x256x65 time series data on a NVIDIA GeForce GTX 1080 Ti GPU. We demonstrate that there exists significant variability in AIFs across patients in the 4.8 sec time resolution DCE data. We finally show that the proposed AIF-TK net considerably improves the TK parameter estimation accuracy in comparison to the network, which does not utilize the patient AIF.

2. METHODS

2.1 DCE-MRI Dataset specifications:

We used the NEURO-RIDER database, which contained multi-parametric MRI scans from 19 brain tumor patients, including DCE-MRI, diffusion tensor imaging (DTI), and post-contrast 3D FLAIR, all acquired on a 1.5 T scanner [8]. DCE-MRI was acquired using a T1 weighted FLASH sequence: with flip angle of 25°; spatial resolution of 1 x 1 x 5 mm^3 voxels; TR: 3.8 ms; TE: 1.8 ms; temporal resolution of 4.8 sec; scan time of 5.2 minutes. The images were

acquired while an intravenous injection of Magnevist 0.1 mmol/kg at 3ccs/sec, was given to the subject 24 secs after the start of the scan. The database also included multi-flip 3D FLASH scans at flip angles of 5°, 10°, 15°, 20°, 25°, and 30° with a TE of 2.1 ms, and TR of 4.43 ms. We used the DESPOT 1 T1-mapping sequence to estimate the T1 maps from the multi-flip angle images. We then converted the DCE image intensity vs. time profiles to the concentration vs. time profiles using the steady-state spoiled gradient-echo equation [9]. The AIF of each patient was found by averaging the concentrations in a region of interest containing the sagittal sinus. The averaged AIF was then fitted to a 10-parameter AIF model to smoothen any remaining noise [10]. Brain tumors were localized in only a few slices of the entire volume. We, therefore, considered only those slices that contained brain tumors from all the patients. We included cases that had enhancing tumors of at least 2 cm (as determined by standard bi-directional assessment). We did not include tumors that were adjacent to the sagittal sinus, as this would result in an atypical AIF. With this criterion, a total of 25 slices were selected. Of these, 20 were used for training, and 5 for testing. Slices were segregated at the patient level to ensure accurate validation. Figure 1 shows exemplar post-contrast images (last frame of the DCE-time series), and patient-specific AIFs from the test set. As seen here, there exists a range of variability in the AIFs across patients with regards to the dispersion and the maximum height of the AIF curves.

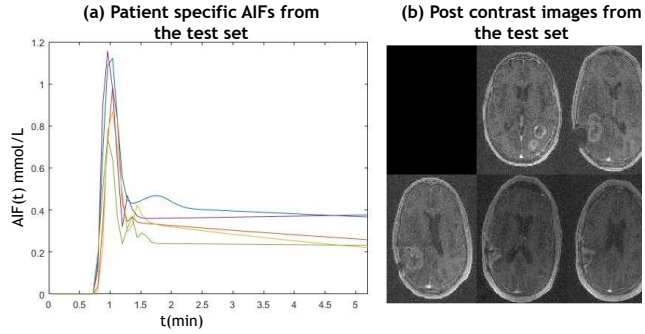

Figure 1: The patient-specific AIFs (a) and post-contrast images (b) from the test set. As seen in (a), the AIFs depict a range of variability across patients.

2.2 AIF-TK network architecture and training:

AIF-TK network architecture is shown in figure 2. We modified the recently proposed architecture for kinetic parameter estimation [7] to account for the patient-specific AIF. The input to our network is a 24x24x65 patch time series from the DCE image time series, and the 1x1x65 AIF time series repeated as a 24x24x65 patch time series. The AIF-TK net maps the above input to a 24x24x3 patch of TK maps (K^{trans}, K^{ep}, v_p) at the output. The patches were extracted from the DCE image time series training data with a stride length of 6 pixels. Similar to [7], the architecture had two paths, a global and local pathway. While the local pathway extracts local neighborhood features, the global pathway extracts contextual global characteristics by using three dilated convolution layers with dilation factors of 2,4,8. The Rectified Linear Unit (ReLU) activation function was used after each convolution layer. The number of features in each layer was 64. The output of the pathways is then merged and fed into three pseudo-fully connected layers with 256 and 128 features. The third layer maps the input patch with 128 features to an output patch containing three features, each representing the K^{trans}, K^{ep}, v_p parameters. Convolutions in these three pseudo-fully connected layers employed a 1x1 kernel size, so the spatial dimension (24x24) and relationships of the propagating patch are maintained.

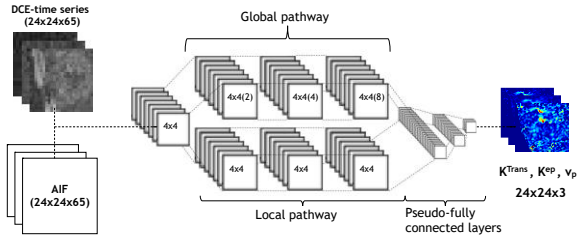

Figure 2: The proposed AIF-TK-net architecture. The network maps the input patch time series from the DCE images and the patient-specific (AIF broadcast to the dimension of the patch time series) to the output TK parameters. The global pathway has dilation factors of 2,4,8. Both the pathways consist of 4x4 convolution blocks, followed by ReLU activation with 64 features. The output of these pathways is fed to three pseudo-fully connected layers being a 1x1 convolution.

For training, we used reference TK maps that were obtained by solving the extended Tofts-Kety model with the L-BFGS iterative algorithm. We trained the AIF-TK-net to minimize the root mean square error between the predicted TK maps, and the above reference TK maps. We used the following training parameters: total patches in the training set = 17000; batch size = 1000 about one slice; learning rate = 0.0001; epsilon = 10^{-5}; steps = 4400; an adaptive moment estimation (ADAM) optimizer. The stopping criterion was determined empirically by evaluating the training loss and the validation loss to avoid over-fitting. The architecture was implemented in Tensor Flow. The total training time for the 20 slices of DCE-time series was approximately 8 hours on a NVIDIA GeForce GTX 1080 Ti GPU. We also implemented the network, which did not account for the patient-specific AIF. This is termed as (TK-net without AIF), and was trained in the same manner as the AIF-TK net, but did not include the AIF patch time series as an input.

2.3 Evaluation:

To evaluate the effect of including patient-specific AIFs in the network, we compared the performance of the proposed AIF-TK net against the TK-net without AIF. The reference TK maps were obtained from non-linear ETK fitting via the L-BFGS iterative solver. The local pathway was designed to extract smaller features and, therefore, is not dilated. The global pathway extracts features throughout the image. We used 4x4 convolution. A comparison was made on the test set containing five slices of DCE concentration vs. time series from different patients. Quantitative comparisons were performed in terms of Bland-Altman plots of the difference in the estimated TK maps and the reference TK maps. All the comparisons were made on manually selected regions of interest (ROI) that contained tumors within each slice.

On the test data, we compared the average computation execution time to estimate the TK parameters with AIF-TK-net and contrasted it against the execution time with the conventional L-BFGS solver. The forward pass implementation of the AIF-TK-net was both on the 1080 Ti GPU and a modern 32 core CPU. The L-BFGS solver was implemented in MATLAB on the same 32 core CPU machine.

3. RESULTS

Figure 3 shows an exemplar result from the test set on a patient with a large glioblastoma tumor. This compares the K^{trans}, K^{ep}, and v_p maps from the estimated AIF-TK net, and the TK-net without AIF against the reference. A significant underestimation in the K^{trans} and K^{ep} maps were observed with the net without the AIF (see orange arrows in figure 3). In contrast, the AIF-TK net provided an improvement in the accuracy of estimating the K^{trans} and K^{ep} maps. The v_p maps with the AIF-TK net showed only a subtle qualitative improvement over the net without AIF. With the AIF-TK-net, we qualitatively observed subtle smearing of v_p maps at the boundaries of vessels and tumor tissue (see white arrow in figure 3). This is attributed to partial volume averaging as the network encourages similarity of TK parameters in local spatial neighborhoods.

Figure 4 shows the Bland-Altman plots of the difference between estimated TK maps and the reference TK maps on all the tumor ROIs combined from the 5 test cases. The individual insets also show the mean $\pm$ 1.96 (standard deviation) ($\mu \pm 1.96\ (\sigma)$) of the difference in the estimated and reference maps. The AIF-TK-net showed a significant reduction in the systemic bias (μ) in estimating K^{trans}, compared to the net without AIF (0.0016 min^{-1} vs. 0.0107 min^{-1}). Similar significant improvements were observed with bias reduction in estimating K^{ep} with AIF-TK-net compared to the net without AIF (-0.00008 min^{-1} vs. 0.0057 min^{-1}). Also, as depicted in figure 4, a reduction in uncertainty (σ) in estimating the K^{trans}, K^{ep} maps were observed with the AIF-TK net compared to the net without AIF. Finally, the error statistics in the v_p maps did not demonstrate significant differences amongst the AIF-TK-net and the net without AIF.

As the AIF-TK-net accounts for the patient-specific AIFs, the kinetic mapping in the vessels is expected to improve. However, due to implicit partial voluming with the net, we observe a subtle increase in the bias of v_p with AIF-TK-net, as evident in the Bland-Altman plot of v_p.

Once trained, the AIF-TK net demonstrated a fast execution time to estimate the TK maps on the test DCE-time series data. The average mean processing times of the forward pass of the AIF-TK-net were 0.34 sec/slice on the GPU and less than a minute on the CPU. In contrast, the conventional L-BFGS solver took ~5 hours/slice on the CPU.

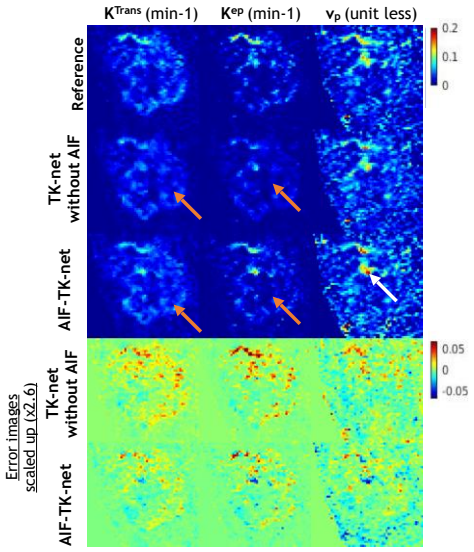

Figure 3: Exemplar results from a glioblastoma tumor. Shown are the K^{trans}, K^{ep}, and v_p maps from the reference (first row), the TK-net without AIF (second row), and the proposed AIF-TK-net (third row). The fourth and fifth rows show the error in the estimates scaled up by a factor of 2.6 for better visualization. AIF-TK-net is robust to the significant under-estimation observed in the TK-net without AIF (see orange arrows in the K^{trans}, K^{ep} maps). Differences in the v_p maps were not significant. Subtle partial volume averaging is observed in the AIF-TK-net on pixels interfacing tissue and vessels (e.g., see white arrow).

4. DISCUSSION

We have proposed, a patient-specific AIF and kinetic model-driven network to rapidly estimate the extended Tofts-Kety kinetic model parameters in DCE-MRI. On brain tumor DCE datasets, we have demonstrated that by accounting for patient-specific AIF variations, the accuracy in estimating the K^{trans} and K^{ep} maps significantly improves. These AIF variations exist in DCE-MRI datasets particularly captured with high time resolution (<5 secs). Once trained, the proposed AIF-TK-net rapidly infers the TK maps **of** the order of 0.34 seconds/slice (on a GPU), or less than a minute (on a CPU) compared to traditional voxel-wise non-linear fitting (e.g., 5 hours/slice with L-BFGS solver on a CPU). The significant gains in the computational speed make the AIF-TK-net an attractive option for use in recently proposed TK model-based reconstruction schemes from under-sampled k-

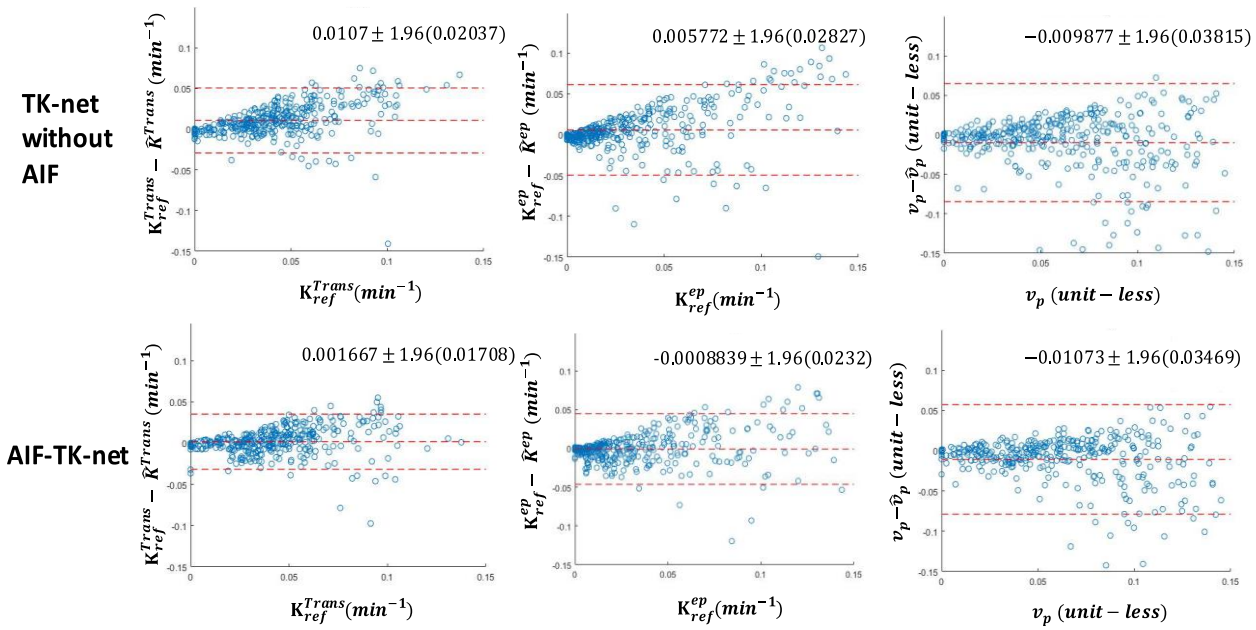

Figure 4: Bland-Altman plots of the difference between the reference TK maps and the estimated TK maps on all the tumor ROIs combined from the 5 test cases. The top row shows the plots for the TK-net without the AIF, and the bottom row shows the plots from the proposed AIF-TK-net. The insets show the mean ± 1.96 (standard deviation) (μ ± 1.96 (σ)) of the difference in the estimated and reference maps.

space data (e.g., [6]), where the optimization iterates between data consistency, and TK estimation.

The proposed AIF-TK net can be improved in several ways. It can be extended to account for nested TK models, where different spatial regions can be modeled by TK models of varying complexity. This can account for the partial volume effects we observe across vessel and tissue pixels, particularly in the v_p maps. Second, the 2D convolutions can be extended to 3D convolutions to exploit similarities along the third dimension of the DCE time series. These extensions are scopes of our future work.

REFERENCES

[1] A. K. Heye, R. D. Culling, M. D. C. Vald??s Hern??ndez, M. J. Thrippleton, and J. M. Wardlaw, "Assessment of blood-brain barrier disruption using dynamic contrast-enhanced MRI. A systematic review," *NeuroImage: Clinical*, vol. 6. pp. 262–274, 2014.

[2] S. P. Sourbron and D. L. Buckley, "Classic models for dynamic contrast-enhanced MRI," *NMR Biomed.*, 2013.

[3] T. Abe, Y. Mizobuchi, K. Nakajima, Y. Otomi, S. Irahara, Y. Obama, M. Majigsuren, D. Khashbat, T. Kageji, S. Nagahiro, and M. Harada, "Diagnosis of brain tumors using dynamic contrast-enhanced perfusion imaging with a short acquisition time," *Springerplus*, 2015.

[4] X. Li, Y. Zhu, H. Kang, Y. Zhang, H. Liang, S. Wang, and W. Zhang, "Glioma grading by microvascular permeability parameters derived from dynamic contrast-enhanced MRI and intratumoral susceptibility signal on susceptibility weighted imaging Head & neck imaging," *Cancer Imaging*, vol. 15, no. 1, 2015.

[5] S. R. Barnes, T. S. C. Ng, N. Santa-Maria, A. Montagne, B. V Zlokovic, and R. E. Jacobs, "ROCKETSHIP: a flexible and modular software tool for the planning, processing and analysis of dynamic MRI studies.," *BMC Med. Imaging*, vol. 15, p. 19, 2015.

[6] Y. Guo, S. G. Lingala, Y. Bliesener, R. M. Lebel, Y. Zhu, and K. S. Nayak, "Joint arterial input function and tracer kinetic parameter estimation from undersampled dynamic contrast-enhanced MRI using a model consistency constraint," *Magn. Reson. Med.*, 2018.

[7] C. Ulas, D. Das, M. J. Thrippleton, M. del C. Valdés Hernández, P. A. Armitage, S. D. Makin, J. M. Wardlaw, and B. H. Menze, "Convolutional Neural Networks for Direct Inference of Pharmacokinetic Parameters: Application to Stroke Dynamic Contrast-Enhanced MRI," *Front. Neurol.*, 2019.

[8] D. Barboriak, "Data From RIDER_NEURO_MRI," *The Cancer Imaging Archive*, 2015. [Online]. Available: http://doi.org/10.7937/K9/TCIA.2015.VOSN3HN1.

[9] K. L. Li, X. P. Zhu, J. Waterton, and A. Jackson, "Improved 3D quantitative mapping of blood volume and endothelial permeability in brain tumors," *J. Magn. Reson. Imaging*, vol. 12, no. 2, pp. 347–357, 2000.

[10] G. J. M. Parker, C. Roberts, A. Macdonald, G. A. Buonaccorsi, S. Cheung, D. L. Buckley, A. Jackson, Y. Watson, K. Davies, and G. C. Jayson, "Experimentally-derived functional form for a population-averaged high-temporal-resolution arterial input function for dynamic contrast-enhanced MRI," *Magn. Reson. Med.*, vol. 56, no. 5, pp. 993–1000, 2006.