Posterior approximation for simultaneous DCE-MRI pharmaco-kinetic parameter and uncertainty estimation

Yannick Bliesener and Krishna S. Nayak
University of Southern California, Los Angeles, CA, United States

Target Audience: Radiologists and MRI physicists who use dynamic contrast-enhanced (DCE) MRI.

Purpose: To explore the potential of machine learning (ML) to simultaneously estimate DCE-MRI pharmacokinetic (PK) parameters and uncertainty. Dynamic contrast enhanced (DCE) MRI aims to estimate sub-voxel parameters of pathology pharmaco-kinetics through fitting of pharmaco-kinetic models to contrast agent concentration-time curves [1]. Many of the involved cost functions from non-linear models are not strongly convex or not even convex which introduces ambiguity in the parameters and deteriorates accuracy and precision due to susceptibility to noise and initializations. Development of such estimators is furthermore challenged by lack of reference methods and ground truth, and measures of precision can only be obtained through time-consuming Monte-Carlo simulations (MCS) with multiple noise realizations and initializations, or variance estimation through linear error propagation [2]. The present work explores the use of re-enforcement learning of neural networks to estimate posterior distributions [3] from which ranges of possible pharmaco-kinetic parameter as well as various metrics of certainty of the estimation can be derived.

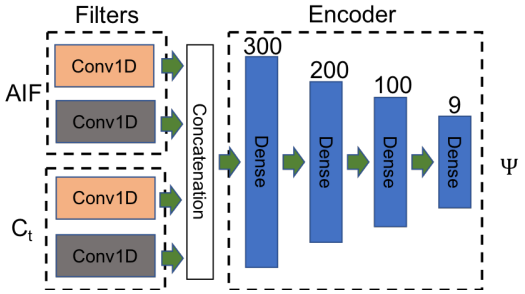

Figure 1: Network architecture: Input filters with temporal convolutions of different kernel lengths, encoder with dense layers.

Methods: The neural network (**Figure 1**) consists of two separate input filter stages for measured concentration-time curves C_t and arterial input functions (AIF) whose output is concatenated and fed into an encoder network of four dense layers. Each layer except for the last is followed by a LeakyReLU layer; the last is activated by sigmoids. The output of the last layer are 9 parameters Ψ per pixel for location, scale, and rotation of a uniform ellipsoidal distribution q_Ψ which serves as approximation to the posterior distribution. The cost function for training is given by:

$$L = \min_{\Psi} \mathrm{E_q}\|C_t - \Phi(\theta; AIF)\|_2^2 + \lambda\, \mathrm{E_q}[q_\Psi(\theta; C_t, AIF)]$$

The first term enforces data consistency of the PK parameters θ drawn from the posterior q_Ψ. The second term, a negative entropy, enforces spread of the posterior to cover all possible PK parameters that could explain the data. Training data consisted of 1572 patches of 20x20 pixels and 50 time points whose noisy concentration time curves are generated by random maps of pathologically realistic parameters [4] for the extended Tofts model with 61 AIFs with synchronized bolus arrival measured from clinical exams at our institution. Test data was taken from a pathologically and anatomically realistic digital reference object [4].

Results: As shown in **Figure 2**, the proposed method is able to estimate PK parameter maps that are comparable to conventional model fitting with multiple initializations. Scatter plots of the standard deviations of Kt predicted by the proposed method and MCS of conventional methods show good correlation between the two predictions, yet no absolute agreement. **Figure 3** shows concentration time curves, and contour plots of the data consistency cost function for various projections into the PK parameter plane. The true value (red) is well contained inside the ellipsoid (blue cloud) and samples from the ellipsoid (blue cloud) result in very similar concentration time curves (blue shaded area).

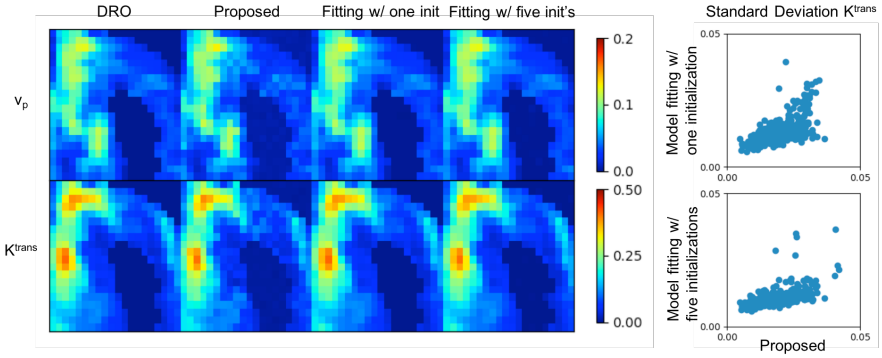

Figure 2: Evaluation in a brain tumor digital reference object. Left: Comparison of PK parameter maps. Right: Correlation of standard deviation as predicted by proposed method and conventional Monte Carlo method.

Discussion: The results indicate that the proposed method can simultaneously predict PK parameters in tumor tissue as well as provide metrics for uncertainty of the estimation due to noise as well as the general structure of the cost function landscape. Since minimum variance unbiased estimation is impossible for non-linear models in the Gaussian noise case a posterior distribution was chosen that does not give preference to a single value (like the mode). Convex posterior distributions like the uniform ellipsoid break down at low values of Kt when ve is impossible to estimate and the cost function is non-convex. Possible solutions are the use of two super-imposed posterior distributions, or sparsity constraints on Kt that enforce the posterior to expand into the desired trench of the cost function landscape.
It remains future work to show how the derived metrics of uncertainty can be used to inform tumor classification, and ultimately impact diagnostics. Limitations of the current work are the almost exclusive use of artificial data instead of real clinical data and the synchronization of the bolus arrival. Training procedure and network architecture were set up however to be readily extensible to clinical data for which no ground truth exists that could be used during training. The effects of changing levels of noise and the correct choice of hyper-parameters also need to be investigated.

Conclusion: Neural networks can be used to simultaneously provide PK parameter estimates and measures of uncertainty without relying on sampling methods, linearizations, or ground truth labels during training.

Acknowledgements: NIH Grant #R33-CA225400
References: [1] Sourbron, et al., NMR (2013) [2] Garpebring, et al., MRM (2013) [3] Dalca, et al., arXiv (2018) [4] Bosca, et al., Phys. Med. Bio (2016)

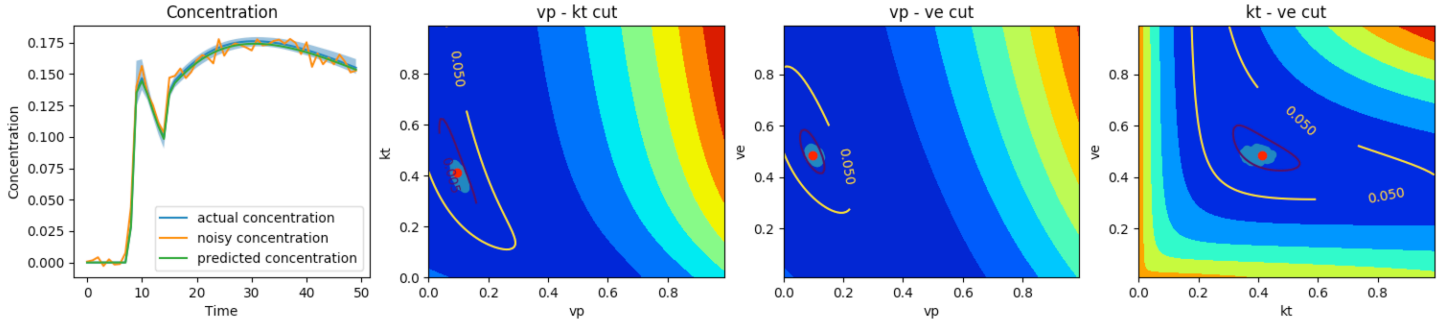

Figure 3: Contour plots of cost functions for PK parameter planes. Red dot is the true value. Blue cloud shows samples from posterior distribution which are used to generate concentration time curves in the blue area.