Deep-Learning Estimation of Perfusion Kinetic Parameters in Contrast-Enhanced Ultrasound Imaging

Enrico Grisan¹, Sevan Harput¹, Bernd Raffeiner³, U. Fiocco², Roberto Stramare²

¹ School of Engineering, London South Bank University, London, UK
  enrico.grisan@lsbu.ac.uk

² Department of Medicine, University of Padova, Padova (Italy)

³ Rheumatology Unit, Bolzano Hospital, Bolzano (Italy)

ABSTRACT

Contrast-enhanced ultrasound (CEUS) is a sensitive imaging technique to evaluate blood perfusion and tissue vascularity, whose quantification can assist in characterizing different perfusion patterns, e.g. in cancer or in arthritis. The perfusion parameters are estimated by fitting non-linear parametric models to experimental data, usually through the optimization of non-linear least squares, maximum likelihood, free energy or other methods that evaluate the adherence of a model adherence to the data. However, low signal-to-noise ratio and the nonlinearity of the model make the parameter estimation difficult.

We investigate the possibility of providing estimates for the model parameters by directly analyzing the available data, without any fitting procedure, by using a deep convolutional neural network (CNN) that is trained on simulated ultrasound datasets of the model to be used.

We demonstrated the feasibility of the proposed method both on simulated data and experimental CEUS data. In the simulations, the trained deep CNN performs better than constrained non-linear least squares in terms of accuracy of the parameter estimates, and is equivalent in term of sum of squared residuals (goodness of fit to the data). In the experimental CEUS data, the deep CNN trained on simulated data performs better than non-linear least squares in term of sum of squared residuals.

Index Terms—parameter estimation, parametric modelling, deep learning, non-linear-least squares, perfusion estimation, contrast enhanced ultrasound

1. INTRODUCTION

At medical ultrasound imaging frequencies (1-10 MHz), blood signal is weak. Contrast-enhanced ultrasound (CEUS) is an imaging modality that exploits microbubbles as contrast agent to visualize blood flow. Since microbubbles are constrained into the vascular bed and have sizes comparable to red blood cells, they are an ideal agent to non-invasively assess vascularization and perfusion e.g. in arthritis [1-5]. CEUS data are generally quantified within a user-defined region of interest (ROI) by analyzing the time intensity curve (TIC) which is the average of all ROI pixels in each frame. However, the heterogeneity of the tissues of interest can only be assessed through a pixel-wise quantification, which involves fitting a parametric model of tissue perfusion to the TICs extracted from individual pixels [1-5].

The procedure to obtain the value of the parameters that best adapt a model to the available observations usually involve the minimization of a cost function measuring the distance of the measurements from the model estimates, and it is performed through classical approaches as non-linear least squares (NLLS), maximum likelihood (ML) or maximum a posteriori (MAP).

Model fitting may yield parameter estimates with large variance as well as considerable bias, since the measured CEUS data are generally very noisy and involve only a small number of sampling points. Moreover, all methods may converge to erroneous solutions since the objective function can have multiple local minima, making them sensitive to initial conditions. Finally, model fitting is computationally demanding, and may become a relevant bottleneck when a model should be fitted to large datasets, with a large number of pixels (and thus TICs) for every patient.

A. Related works

Few attempts have been made in parametric model fitting using machine learning in general and deep neural networks in particular, as the bulk of research in the applications of deep neural networks to dynamical system concentrates in the prediction of future system states, using AI as a black-box or gray-box (for a review see [6]), and only retrospectively linking the network weights to some physical meaning.

Additionally, many state-of-the-art classification records have been achieved by deep convolutional neural networks (CNNs) [7-9]. However, parametric mapping in imaging applications using deep learning has been attempted only recently, especially in medical magnetic resonance dynamic imaging data [10-13], where CNNs were used to directly estimate perfusion parameters from temporal sequences. The major drawback of these methods is that no ground-truth data (couples of time-curves and true model parameters) is
available for training the proposed neural networks: all proposed methods rely on using the results from other methods as target values for the supervised learning. The trained CNNs hence incorporate all errors possibly present in the reference method.

B. Contributions

Similarly to other methods, the present work aims at employing CNNs for direct estimation of model parameter values when modelling a temporal sequence. The key novelty is that, at variance with all currently available methods, we want to investigate the possibility of training a CNN on simulated data, allowing a controlled setting and an accurate monitoring of the network performance, and applying it on real data maintaining state-of-the-art performance. The work will compare the performance of the proposed approach with classical nonlinear-least squares (NNLS) [1,5,14] both on data simulated from a non-linear model, and on real data from CEUS images.

2. METHODS

A. Training data definition

In the following sections each training example \(d_i\) is assumed to be a one dimensional vector \(1 \times N_s\), where \(N_s\) is the number of time points \(t_1, \ldots, t_{N_s}\), obtained by a sampling of a curve described by a parametric model \(d(p, t)\) with additional noise \(n(t)\).

B. Convolutional neural network architecture for each model parameter

The proposed architecture is represented in Fig. 1 and is composed by three blocks, each composed by a sequence of two convolutional layers with non-linear activation (leaky ReLU) followed by a max-pooling layer. The convolutional filters are 1x3 to reflect the 1-dimensional nature of the problem, and the max pooling has dimension 2 (halving the size of the input data). The convolutional blocks are then followed by a flattening and three fully connected layers performing the parameter regression. The loss function minimizes the mean relative error, so to be able to reliable estimates across all spectrum of possible parameter values.

Being \(N_p\) the number of available training samples, \(p\) the true value of a scalar parameter and \(\hat{p}\) its current estimate, the relative error loss is computed as:

\[
L(p, \hat{p}) = \sum_{j=1}^{N_p} \frac{|p_j - \hat{p}_j|}{p_j} \quad (1)
\]

for which the derivative to be used in the backpropagation becomes:

\[
\frac{\partial L(p, \hat{p})}{\partial \hat{p}} = \sum_{j=1}^{N_p} \text{sign}(p_j - \hat{p}_j) \frac{p_j}{p_j} \quad (2)
\]

3. DATA

A. Simulated data

Following the perfusion model proposed in [15], we assume that a time-activity curve can be formalized as a nested model composed by a principal component described by a Gamma-variate curve \(c_{\text{gamma}}\) and a slow recirculation component \(c_{\text{SCR}}\):

\[
c_{\text{gamma}}(p, t) = \frac{A}{c_{\text{gamma, max}}} \cdot (t - t_0)^{a} \cdot e^{-\frac{t-t_0}{\beta}} \quad t \geq t_0 \quad (5)
\]

\[
c_{\text{SCR}}(p, t) = c_{\text{gamma}}(p, t) + \frac{\alpha_1}{AUC_g} \int_{t_0}^{t} c_{\text{gamma}}(r)dr \quad t \geq t_0 \quad (6)
\]

Figure 1 Deep-convolutional network architecture for the estimation of a single parameter. For each computational block the dimension of its output is shown, considering an input curve containing \(N_s\) time points. Computational layers are color-coded by type: input and output in blue, convolutional in yellow, nonlinear activation in orange, pooling in gray and fully connected layers in green.
With \( \mathbf{p} = [A, \alpha, t_0, \beta, a_1] \) being the parameter vector defining the model. In order to better control the generation of physiologically-meaningful curves and avoid numerical difficulties when using fitting methods, a different parametrization has been chosen: \( \mathbf{p}' = [A, \alpha, t_0, t_{\text{max}}, b] \) where:

\[
\begin{align*}
\beta &= \frac{t_{\text{max}} - t_0}{\alpha} \\
b &= 1 - a_1
\end{align*}
\]

With this parametrization we can generate \( N_p \) set of parameters, and the corresponding curves assuming a time range \( t \in [0,100] \) s; a white noise and a proportional noise has been added to each generated curve to obtained the simulated data curve \( d(\mathbf{p}, t) \):

\[
\begin{align*}
d(\mathbf{p}, t) &= c_{\text{SIR}}(\mathbf{p}, t) + c_{\text{gamma}}(\mathbf{p}, t) + n(t) \\
n(t) &= \mathcal{N}(0,0.01) + \mathcal{N}(0,0.25 \cdot d(\mathbf{p}, t))
\end{align*}
\]

B. Contrast enhanced ultrasound data
From previously reported studies [15, 16] of subjects affected by inflammatory arthritis, a random patient was used in the current study. From the original cohort, all patients showed signs of inflammatory finger joint involvement: the joint with the highest disease activity was chosen for CEUS examination. Each subject underwent a 2-minute CEUS scan as described in [15]. CEUS images were motion corrected and co-registered to the corresponding B-mode anatomical image, and the time-intensity curves were normalized by the maximum value. Only pixels within the synovial region (manually or semi-automatically outlined [17]) that showed a significant enhancement were evaluated, resulting in 7238 TICs extracted from the single patient considered. All curves were resampled and trimmed to 100 seconds to meet the input size of the CNN. Some representative time-intensity curves are shown in Fig. 2.

4. RESULTS

A. Parameter estimation on simulated data
A set of \( N_p = 10000 \) data curves were generated, and split using 80% for training, and 20% for testing. The CNN was implemented and trained in Matlab (The Mathworks Inc.) using MatConvNet [18]. The training was performed for 200 epochs with a decreasing learning rate (from \( 10^{-3} \) to \( 10^{-6} \)) with the Adam [19] update and a minibatch size of 100. The results were compared with those obtained using a constrained non-linear least squares method; to provide a fair comparison, the constrained non-linear least squares estimation method was provided as as parameter limits those used in the generation of the curves, that might be an additional knowledge that leaks in the CNN during the learning phase.

The absolute relative error (ARE) of the \( j \)th parameter was defined as in Eq. 12, and results are reported in Tab. II:

\[
ARE_j = \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{|\hat{p}_i(j) - p_i(j)|}{\hat{p}_i(j)}
\]

However, since the mean absolute relative error is heavily affected by relatively small errors on parameters with values close to zero, to mitigate this effect, we also reported the mean weighted absolute error (WAE) (Tab. III), that is:

\[
WAE_j = \frac{1}{N_p} \sum_{i=1}^{N_p} \frac{|\hat{p}_i(j) - p_i(j)|}{\hat{p}_i(j)}
\]

where \( \hat{p}_i(j) \) is the average of the real parameter values.

Surprisingly, despite the difference in the values of parameter estimates, the average mean squared error (mean of sum of squared residuals - MSSR) between the estimated curves and the data was similar for CNN (6.2) and NLLS (6.0), with no statistically significant difference.

### Table I Simulated Data: Weighted Relative Error Between Estimated Parameter and Ground Truth (Test Set)

| Parameter | \( A \) | \( \alpha \) | \( t_0 \) | \( t_{\text{max}} \) | \( b \) |
|-----------|--------|--------|--------|----------------|------|
| ARE mean (std) | CNN: (2.5, 3.3) | NLLS: (2.4, 2.8) | CNN: (9.5, 28.4, 10.2, 29.2) | NLLS: (1.2, 1.6, 1.2, 1.7) | CNN: (2.9, 4.5, 2.7, 3.9) |
| WAE mean (std) | CNN: (2.2, 6.6) | NLLS: (2.5, 4.4) | CNN: (8.7, 161.8, 11.9, 82.1) | NLLS: (3.9, 39.5, 2.4, 4.7) | CNN: (5.2, 9.1, 5.0, 7.6) |
The Deep-CNN trained on simulated data was applied on real data for estimating the parameters of the chosen model. In order to assess the performance of the proposed method, the parameters were estimated on the same data using a classical constrained NLLS curve fitting. At variance with the results on simulated data, the MSSR obtained using the parameter obtained from CNN and from the constrained NLLS methods vary significantly (Tab. II). Moreover, the mean estimated values for the different parameters are reported in Tab. III: for all parameters, the difference between CNN and NLLS estimates is significant, and it can qualitatively appreciated by visually inspecting the parametric maps shown in Fig. 3.

### TABLE II CEUS DATA: MODEL FITTING PERFORMANCE (MEAN SQUARED ERROR) WITH THE PARAMETERS

|        | CNN mean (std) | Constrained NLLS mean (std) |
|--------|----------------|-----------------------------|
| MSSR   | 3.44 (1.1)     | 8.24 (7.7); \( p < 10^{-10} \) |

**DISCUSSION & CONCLUSIONS**

Fit of non-linear parametric models to CEUS data is usually carried out through iterative procedures that are sensitive to initial condition and local minima of the objective function used as fit criterion. This makes the estimation and interpretation a very sensitive issue, as the estimated parameters are used to obtain insight into the inner behavior of physiological, pathological or cognitive mechanism, so that modellers often need to check and adjust the estimation procedure until convergence is ensured. Some attempt into using deep-learning regression have been tried, using as ground truth the values of the model parameters obtained by alternative methods.

At variance with these approaches, the proposed method is trained on simulated data, and the results show that it outperforms classical methods in term of estimation accuracy and goodness of fit results, maintaining the performance when applied to real CEUS data, providing parameters that fit the data better than NLLS, in a fraction of the time.

The main current limitations of the proposed method is the need to be able to generate a reasonable simulation of the data, and, more importantly, the need to have a constant sampling grid, as the input of the CNN is a curve sampled at \( N_T \) time points.

**ACKNOWLEDGMENTS**

The work has been partially supported by University of Padova intramural research project BIRD160889/16. The authors have no relevant financial or non-financial interests to disclose.
COMPLIANCE WITH ETHICAL STANDARDS

The study was approved by the local Ethic Committee of the University Hospital of Padova (Italy) (number 52723; October 11, 2010), and written informed consent was obtained from each participant in accordance with the principles outlined in the Declaration of Helsinki, after being informed about the intent and the methodology of the study.

REFERENCES

[1] Klausner A, Demharter J, De Marchi A, Sureda D, Barile A, Masciocchi C, et al. Contrast enhanced gray-scale sonography in assessment of joint vascularity in rheumatoid arthritis: results from the IACUS study group. Eur Radiol. 2005;15(12):2404–10.

[2] De Zordo T, Mlekusch SP, Feuchter GM, Mur E, Schirmer M, Klausner AS. Value of contrast-enhanced ultrasound in rheumatoid arthritis. Eur J Radiol. 2007;64(2):222–30.

[3] U. Fiocco, R. Stramare, V. Martini, et al.: “Pixel-based method of contrast-enhanced ultrasound demonstrates the linear relationship between synovial vascular perfusion and the recruitment of pathogenic IL-17A-F+ IL-23+ CD161+ CD4+ T helper cells in the psoriatic arthritis joints”, Clinical Rheumatology, February 2017, Volume 36, Issue 2, pp 391–399

[4] Mouterde G, Aegerter P, Correes J-M, Breban M, D’Agostino M-A. Value of contrast-enhanced ultrasonography for the detection and quantification of enthesitis vascularization in patients with spondyloarthritis. Arthritis Care Res (Hoboken). 2014;66(1):131–8.

[5] Rizzo G, Raffeiner B, Coran A, Cipriani L, Fiocco U, Botuo C, et al. Pixel-based approach to assess contrast enhanced ultrasound kinetics parameters for differential diagnosis of rheumatoid arthritis. J Med Imaging.

[6] A. Chiuso and G. Pillonetto, “System Identification: A Machine Learning Perspective”, Annual Review of Control, Robotics, and Autonomous Systems, Vol. 2:281-304, 2019

[7] Y. LeCun, Y. Bengio, G. Hinton, “Deep Learning”, Nature. 28;521(7553):436-44, 2015

[8] A. Karpathy, T. Leung, “Large-scale Video Classification with Convolutional Neural Networks”, ICCV, pp. 1725-1732, 2014

[9] S. Lathuilière, P. Mesejo, X. Alameda-Pineda, R. Horaud, “A Comprehensive Analysis of Deep Regression”, arXiv:1803.08450

[10] K.C. Ho, F. Scalzo, K.V. Sarma, S. Es-Saden, and C.W. Arnold. “A temporal deep learning approach for MR perfusion parameter estimation in stroke.” 23rd International Conference on Pattern Recognition (ICPR), pages 1315–1320, 2016.

[11] C.M. Scannell, P. van den Bosch, A. Chiribiri, J. Lee, M. Breeuwer, M. Veta, “Deep learning-based prediction of kinetic parameters from myocardial perfusion MRI”, Medical Imaging with Deep Learning, MIDL 2019

[12] C. Ulus, D. Das D, M.J. Thrippleton, M. Valdés Hernández, P.A. Armitage, S.D. Makin, J.M. Wardlaw, B.H. Menze, “Convolutional Neural Networks for Direct Inference of Pharmacokinetic Parameters: Application to Stroke Dynamic Contrast-Enhanced MRI”, Front Neurrol. Jan 8;9:1147, 2019

[13] F. Glang, A. Deshmane, S. Prokudin, F. Martin, K. Herz, T. Lindig, B. Bender, K. Scheflller, M. Zais, “DeepCEST 3T: Robust MRI parameter determination and uncertainty quantification with neural networks—application to CEST imaging of the human brain at 3T”, Magnetic Resonance in Medicine, Volume94, Issue1, Pages 450-466, 2020

[14] Turco S, Frinking P, Wildeboer R, Arditi M, Wijkstra H, Lindner JR, Mischi M. Contrast-Enhanced Ultrasound Quantification: From Kinetic Modeling to Machine Learning. Ultrasound Med Biol. 2020 Mar;46(3):518-545

[15] G. Rizzo, M. Tonietto, M. Castellaro, B. Raffeiner, A. Coran, U. Fiocco, R. Stramare and E. Grisan, “Bayesian quantification of contrast-enhanced ultrasound images with adaptive inclusion of an irreversible component”, IEEE Trans Med Im, Apr;36(4):1027-1036, 2017

[16] Stramare R, Raffeiner B, Cipriani L, Scagliori E, Coran A, Perissinotto E, et al. Evaluation of finger joint synovial vascularity in patients with rheumatoid arthritis using contrast-enhanced ultrasound with water immersion and a stabilized probe. J Clin Ultrasound. 2012;40(3):147–54.

[17] E. Veronesi, R. Stramare, A. Campion, et al: ‘Improved detection of synovial boundaries in ultrasound examination by using a cascade of active-contours’, Medical Engineering & Physics 2013, February, 35, 188–19

[18] A. Vedaldi, K. Lenc, “MatConvNet - Convolutional Neural Networks for MATLAB”, MM’15: Proceedings of the 23rd ACM international conference on Multimedia, 2015

[19] D.P. Kingma and J.L. Ba. “Adam : A method for stochastic optimization.” 2014. arXiv:1412.6980v9