

SYNTHETIC Q-SPACE LEARNING WITH DEEP REGRESSION NETWORKS FOR PROSTATE CANCER CHARACTERISATION WITH VERDICT

Vanya Valindria¹, Marco Palombo¹, Eleni Chiou¹,
Saurabh Singh², Shonit Punwani², Eleftheria Panagiotaki¹.

¹ Centre for Medical Image Computing, Department of Computer Science,
University College London, London (UK)

² Centre for Medical Imaging, University College London, London (UK)

ABSTRACT

Traditional quantitative MRI (qMRI) signal model fitting to diffusion-weighted MRI (DW-MRI) is slow and requires long computational time per patient. Recently, q-space learning utilises machine learning methods to overcome these issues and to infer diffusion metrics. However, most of q-space learning studies use simple multi layer perceptron (MLP) for model fitting, which might be sub-optimal when estimating more complex diffusion models with many free parameters. Previous works only investigate the application of q-space learning on diffusion models in the brain. In this work, we explore q-space learning for prostate cancer characterization. Our results show that while simple MLP is adequate to estimate parametric maps on simple models like classic VERDICT, deep residual regression networks are needed for more complex models such as VERDICT with compensated relaxation (R-VERDICT).

Index Terms— diffusion MRI, regression, residual, deep networks, model fitting, qMRI, prostate cancer microstructure

1. INTRODUCTION

Prostate cancer (PCa) is the second most commonly occurring cancer among men in the world. However, PCa diagnosis still relies on biopsy, which is invasive and unpleasant for the patients [1]. Thus, development of reliable non-invasive diagnostic is needed. VERDICT (Vascular, Extracellular and Restricted Diffusion for Cytometry in TumoUrs) [2] is an advanced microstructural imaging technique for cancer characterization, which has been proposed as an additional, non-invasive PCa diagnostic tool. VERDICT combines an optimized DW-MRI acquisition protocol with a mathematical model to estimate microstructural features to characterize cancer.

In this work, we compare two VERDICT models, classic VERDICT [2][1] and VERDICT with compensated relaxation (R-VERDICT) [3]. R-VERDICT model [3] incorporates compartment-specific relaxation effects that decouple relaxation and diffusion properties of the tissue. Hence, we

have additional maps in R-VERDICT; 9 quantitative maps of prostate diffusion and relaxation properties (see Figure 1) instead of 5 maps in classic VERDICT.

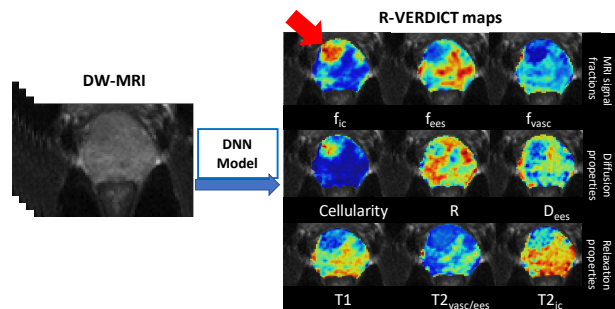


Fig. 1. The nine R-VERDICT parametric maps obtained from DW-MRI images (with different b-values), generated by Deep Neural Networks (DNN) model.

The task in quantitative DW-MRI is to find a mapping from a limited number of noisy signal samples to quantify microstructural tissue properties [4]. However, computation of numerical modelling can be time consuming. Machine learning has become an important tool to efficiently solve computational problems in many applications in diffusion DW-MRI applications [5]. Machine learning approaches called q-space learning can infer the parameters of diffusion MRI using deep neural networks [6].

Pioneer work on q-space learning by [7] estimates model parameters using MLP as regression neural network. In [7], training datasets are prepared using the fitting results of real MRI data (*in-vivo*), which may introduce more variation in training data. Synthetic (*in-silico*) q-space learning proposed by [8] only uses synthetic training datasets generated from equations derived from a model. Use of synthetic training data has been proven to have the same quality as using the real data (*in-vivo*) in deep learning model fitting[9]. Additionally, synthetic q-space learning has fully controllable quality of training data, clearly defined gold standard parameters [6] and shown to be useful when the training signals are not avail-

able.

Most previous works in q-space learning relied on simple MLP and only investigated diffusion models applied on brain [10][7][8][9][6]. Here, we aim to do synthetic q-space learning on prostate cancer characterization. In this paper, we also investigate the performance of deep regression models in noise level matching (NLM) synthetic training data.

2. MATERIALS AND METHOD

2.1. Datasets

This study is part of the INNOVATE clinical trial [11] which involved 44 patients (median age, 64.8 years; range, 49.5–79.6 years). VERDICT MRI data for all patients were acquired using an optimised imaging protocol for VERDICT prostate characterization with 5 b-values (90, 500, 1500, 2000, 3000 s/mm²) in 3 orthogonal directions [12]. The DWI sequence was acquired with a voxel size of 1.25×1.25×5 mm³, and the images were reconstructed to a 176×176×14 matrix size. The regions of interest (ROIs) corresponding to prostate lesions were contoured by an experienced radiologist. The pre-processing pipeline includes denoising of the raw DW-MRI data then correction for Gibbs ringing. Then, we applied mutual-information rigid and affine registration.

2.2. Q-space learning for R-VERDICT

Machine learning techniques can explore data-driven computational methods that express relationships to the DW-MRI signal [5]. Q-space learning infers diffusion model parameters using regression neural networks [7]. This work aims to generate the 9 parametric maps of R-VERDICT (see Figure 1) using q-space learning. As proposed in [3], the mathematical equation for R-VERDICT is:

$$S(b, TE, TR) = S_0 \left(1 - e^{-\frac{TR}{T_1}}\right) \left[f_{vasc} e^{-\frac{TE}{T_2^{vasc/ees}}} S_{vasc}(D_{vasc}, b) + f_{ic} e^{-\frac{TE}{T_2^{ic}}} S_{ic}(D_{ic}, R, b) + f_{ees} e^{-\frac{TE}{T_2^{vasc/ees}}} S_{ees}(D_{ees}, b) \right] \quad (1)$$

Please refer to [3] for further explanation of the Eq.1. We generate 100,000 synthetic DW-MRI signals (split into 80% for training and 20% for validation) using Eq.1 with different values for the model parameters randomly chosen between biophysical range: $S_0 = [0, \max b=0 \text{ intensity}]$, $T_1 = [10, 4000] \text{ ms}$, $T_2^{ic} = [1, 150] \text{ ms}$, $T_2^{vasc/ees} = [150, 800] \text{ ms}$, $f_{ees} = [0.01, 0.99]$, $f_{ic} = [0.01, 0.99]$, $R = [0.01, 15] \mu\text{m}$ and $D_{ees} = [0.5, 3] \text{ m}^2/\text{ms}$. We also add Rician noise corresponding to SNR = 35 to the training data [3]. The number of samples, type and level of noise for synthetic DW-MRI signals are chosen experimentally and based on real MRI data.

In this q-space learning, we treat each image voxel individually as a data sample. We investigate different models to predict the parametric maps from the protocols, as illustrated

MSE	Model 1	Model 2	Model 3
f _{vasc}	0.0023	0.0024	0.0026
f _{ic}	0.0015	0.0017	0.0018
f _{ees}	0.0018	0.0018	0.0018
R	0.4007	0.4044	0.4100
Cellularity	5.6×10 ⁻⁶	5.4×10 ⁻⁶	5.4×10⁻⁶

Table 1. MSE of the 5 classic VERDICT parametric maps generated by deep regression models, compared to NLSS. Bold numbers show the least MSE obtained for the estimated maps.

in Figure 2. For all models, we train the networks by back-propagating the mean squared error (MSE) between the real model parameters and networks predictions, as the objective we try to minimize. We perform the optimisation with the adaptive moment estimation (ADAM) for 100 epochs (early stopping to prevent overfitting) and mini-batch of 100. We normalize the input data to [0,1] and rescale the prediction back from the networks. The number of input matched with the dimension of input signal generated by DWI-MRI protocol and output neurons equal to the number of tissue parameters to be estimated. We consider three different models with different architectures, as described in the following section.

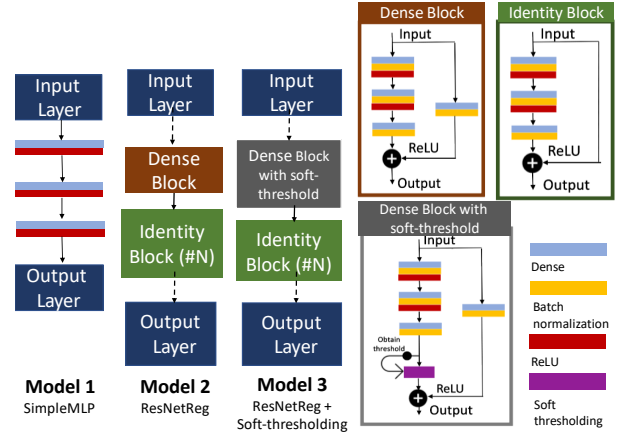


Fig. 2. Deep regression networks in our study: Model 1: Three-layer MLP. Model 2: stacked residual regression networks. Model 3: stacked residual regression networks with soft-thresholding.

2.2.1. Model 1: Simple MLP

For the first model, we use a multi-layer perceptron (MLP) with three hidden layers, each consisting of 150 hidden units with a nonlinearity function rectified linear unit (ReLU) and a final regression layer. This model has been applied and worked well on brain microstructure model fitting (DKI, NODDI, etc.) in [7][6].

2.2.2. Model 2: ResNetReg

Previous trend on deep learning is to go deeper in the networks to achieve a better accuracy. However, much deeper networks face vanishing gradient problem and fail to perform better than the shallow counterpart [13]. This problem in training very deep networks has been alleviated with the residual neural network (ResNet) [13]. ResNet introduced shortcut connection to train very deep convolutional models, with no extra parameters and no added computation complexity. For this nonlinear regression implementation, we replace the convolution layers by fully connected layers, as shown to work in [14].

There are two building blocks in this residual regression network, identity block and dense block, as illustrated in Figure 2 for Model 2. In identity block, input and output have the same dimension, whereas in dense block they have different dimension. The network structure in Model 2 begins with an input block, then every dense block is followed by two identity blocks, so forth, and the output block in the end. We adapted this model to our R-VERDICT q-space learning with the optimized network at depth 31 and width fixed at 128.

2.2.3. Model 3: ResNetReg + Soft-thresholding

In Model 3, we propose to add the stacked residual regression network on Model 2 with soft thresholding, as it is shown to work effectively to remove noise-related features in [15]. MRI signals are formed as magnitudes of complex values and the noise distribution appeared as Rician. In DW-MRI VERDICT protocol, raw data is also noisy hence may not be an exact noise level matching with our experimental synthetic training data. Soft thresholding has often been used in signal denoising methods where the unimportant features are enforced to be zeros, so that the learned high-level features can become more discriminative [15]. Using deep architecture, we assume that the value of soft threshold can be learned in the dense block unit, as shown in Figure 2, where it is inserted as nonlinear transformation layer before addition with residual connection.

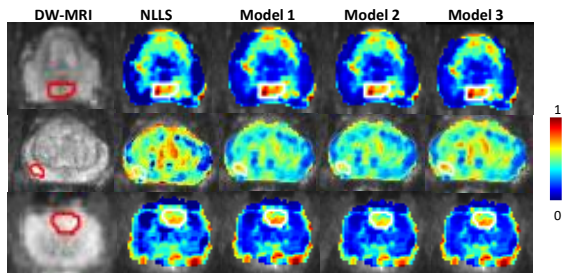


Fig. 3. VERDICT f_{ic} maps generated by deep regression networks compared to NLLS as reference. Lesions are contoured on DW-MRI (red) and on the maps (white).

3. RESULTS

We evaluated the different models of deep regression network on real DW-MRI data from 44 patients. As benchmark, we compare the performance of our models to the conventional non-linear least square (NLLS) minimisation. Previous literature has focused extensively on comparisons between NLLS and deep learning based model fitting [7][10] demonstrating that deep networks is a valid alternative to NLLS.

3.1. Deep regression networks for classic VERDICT

VERDICT model is the sum of three major tissue compartments that mostly contribute to the measured DW-MRI signal: intra-cellular, intra-vascular and extra-cellular/extra-vascular [2][1]. For classic VERDICT, we estimate 5 parametric maps and evaluate the quality of the models in terms of MSE. We observe that for Model 2 and 3, setting narrower width (16) and lower depth (28) could help improving the performance of residual regression networks for classic VERDICT. From Table 1, we can see that Model 1 best estimates f_{ic} maps and is slightly better than Model 2 and 3 in estimating R maps. Although for f_{ic} Model 1 gives the most similar maps to NLLS, but visually in Figure 3, all models could show the lesion ROIs with high values in f_{ic} maps.

3.2. Deep regression networks for R-VERDICT

We compare the performance of all the models in terms of MSE for all the 9 parametric maps of R-VERDICT. This represents a quantitative evaluation of the results presented. As shown in Table 2, mostly residual regression based networks have lowest errors in predicting the maps compared to simple MLP. Especially for the parametric map that is the most prominent for cancer characterization f_{ic} [16], residual regression networks (Model 2 and 3) give better estimation compared to Model 1.

Since MSE has different range for each of R-VERDICT maps, we also computed the MAPE (Mean Absolute Percentage Error) so that we can have overall error comparison for all maps. With MAPE, we calculated the errors with respect to each map range. From Table 3, Model 3 gives the least error for all R-VERDICT parametric maps.

Visually, we also show the results of the parametric maps generated by different models, with respect to the ground truth (standard NLLS), as depicted in Figure 4. Parametric map f_{ic} is the most useful for lesion characterization, with Cellularity, $T2_{ic}$, and $T1$ maps from R-VERDICT are complementary. In Figure 4, all of the models basically could show lesion in the f_{ic} maps but Model 3 can discriminate it better (high value in lesion ROI). Superiority of deep regression networks in detecting cancer lesion is shown in Figure 4 for Cellularity maps where the lesion can be barely seen in the maps generated by NLLS. For relaxation properties: $T2_{ic}$

Table 2. MSE of all of the R-VERDICT maps generated from deep regression models, compared to NLLS as the reference. Bold shows the least MSE obtained for the estimated maps.

MSE	Model 1	Model 2	Model3
fvasc	0.0018	0.0016	0.0013
fic	2.51×10^{-4}	2.24×10^{-4}	1.53×10^{-4}
fees	1.1×10^{-3}	1.04×10^{-3}	1.00×10^{-3}
R	0.1775	0.2227	0.1918
Cellularity	3.4×10^{-6}	4.5×10^{-6}	2.5×10^{-6}
Dees	0.0077	0.0062	0.0061
T2vasc_ees	290	329	354
T2ic	18	19	17
T1	2007	2390	2161

Table 3. MAPE for all R-VERDICT maps.

MAPE	Model 1	Model2	Model 3
All maps	0.44	0.42	0.39

and T1 maps, deep regression networks can spot the lesion region with low values.

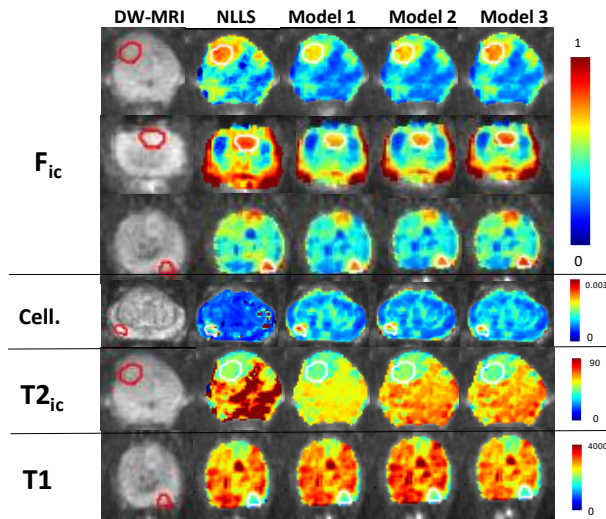


Fig. 4. Example of R-VERDICT parametric maps generated by deep regression networks compared to NLLS. Lesions are contoured on DW-MRI (red) and on the maps (white).

4. DISCUSSION

We studied deep regression network for synthetic q-space learning, to estimate VERDICT parameters from DW-MRI data. Our results suggest that while Model 1 is already adequate to generate classic VERDICT parametric maps, it is outperformed by residual regression networks in R-VERDICT, which has more parametric maps. In our experiments, we also have tried another advanced model: auto-

encoder ResNetReg [17]. For classic VERDICT, we found that auto-encoder ResNetReg gives similarly good prediction for all maps like other models. However, its performance was way below the MLP for R-VERDICT so that we only present the stacked residual architecture models in this work.

In general, deep regression networks give much lower MSE for all basic 5 parametric maps (Table 2) compared to the ones predicted with classic VERDICT (Table 1). For R-VERDICT, f_{ic} maps are best estimated by Model 3 (Table 2) with significantly reduced MSE compared to Model 1.

Residual regression networks work better than simple MLP because R-VERDICT model is more complex, with 9 maps to be estimated compared to other diffusion MRI models with less number of parametric maps estimated in [6], such as DKI (2 parameters), DTI (4 parameters), and NODDI (3 parameters). Stacked residual regression architectures work well because deeper networks perform better at approximating more complex nonlinear functions like R-VERDICT equation (Eq.1). The addition of soft-thresholding module in Model 3 improves the estimation of R-VERDICT maps, especially on the volume fraction maps. It is because soft-thresholding could handle noise in real MRI data better, which may not be an exact match with synthetic training data [6]. We also show that R-VERDICT maps generated by deep residual regression networks retain cancer conspicuity.

There are several limitations that we would like to address. Firstly, we considered only VERDICT-based qMRI models for prostate. Secondly, we train the model on synthetic training data with noise matching level, which is an ideal condition. In the future, we will assess our findings by considering additional diffusion models and more noise variation in synthetic training data.

5. SUMMARY

Deep regression network is a viable approach for q-space learning on R-VERDICT model fitting. The use of synthetic training data requires less time and yields good quality parameter estimation without the need of paired parametric maps from real MRI data. While most q-space learning studies utilise simple MLP, we show that deep residual regression networks are more robust for complex models with many parameters, such as R-VERDICT, and can improve lesion conspicuity. Hence, the maps can be used for improving PCA diagnosis by better targeting biopsies and avoid unnecessary biopsies.

6. COMPLIANCE WITH ETHICAL STANDARDS

The project was conducted using human subject data from ClinicalTrials.gov identifier NCT02689271 for INNOVATE.

7. ACKNOWLEDGMENTS

This work has been funded by EPSRC grand EP/N021967/1. Marco Palombo is supported by UKRI Future Leaders Fellowship MR/T020296/1. This study was supported by the National Institute for Health Research University College London Hospitals Biomedical Research Centre.

8. REFERENCES

- [1] Eleftheria Panagiotaki, Rachel W Chan, Nikolaos Dikaios, Hashim U Ahmed, James O’Callaghan, Alex Freeman, David Atkinson, Shonit Punwani, David J Hawkes, and Daniel C Alexander, “Microstructural characterization of normal and malignant human prostate tissue with vascular, extracellular, and restricted diffusion for cytometry in tumours magnetic resonance imaging,” *Investigative radiology*, vol. 50, no. 4, pp. 218–227, 2015.
- [2] Eleftheria Panagiotaki, Simon Walker-Samuel, Bernard Siow, S Peter Johnson, Vineeth Rajkumar, R Barbara Pedley, Mark F Lythgoe, and Daniel C Alexander, “Noninvasive quantification of solid tumor microstructure using verdict mri,” *Cancer research*, vol. 74, no. 7, pp. 1902–1912, 2014.
- [3] Marco Palombo, Saurabh Singh, Hayley Whitaker, Shonit Punwani, Daniel C Alexander, and Eleftheria Panagiotaki, “Relaxed-verdict: decoupling relaxation and diffusion for comprehensive microstructure characterization of prostate cancer,” *Proc of ISMRM*, 2020.
- [4] Elizabeth B Hutchinson, Alexandru V Avram, M Okan Irfanoglu, C Guan Koay, Alan S Barnett, Michal E Komlosch, Evren Özarlan, Susan C Schwerin, Sharon L Juliano, and Carlo Pierpaoli, “Analysis of the effects of noise, dwi sampling, and value of assumed parameters in diffusion mri models,” *Magnetic resonance in medicine*, vol. 78, no. 5, pp. 1767–1780, 2017.
- [5] Daniele Ravi, Nooshin Ghavami, Daniel C Alexander, and Andrada Ianus, “Current applications and future promises of machine learning in diffusion mri,” in *CDMRI, MICCAI*. Springer, 2019, pp. 105–121.
- [6] Yoshitaka Masutani, “Noise level matching improves robustness of diffusion mri parameter inference by synthetic q-space learning,” in *IEEE ISBI*. IEEE, 2019, pp. 139–142.
- [7] Vladimir Golkov, Alexey Dosovitskiy, Jonathan I Sperl, Marion I Menzel, Michael Czisch, Philipp Sämann, Thomas Brox, and Daniel Cremers, “Q-space deep learning: twelve-fold shorter and model-free diffusion mri scans,” *IEEE TMI*, vol. 35, no. 5, pp. 1344–1351, 2016.
- [8] Chuyang Ye, Yue Cui, and Xiuli Li, “Q-space learning with synthesized training data,” in *CDMRI-MICCAI*. Springer, 2019, pp. 123–132.
- [9] Francesco Grussu, Marco Battiston, Marco Palombo, Torben Schneider, Claudia Wheeler-Kingshott, and Daniel Alexander, “Deep learning model fitting for diffusion-relaxometry: a comparative study,” in *CDMRI, MICCAI*. Springer, 2020, pp. 1–14.
- [10] Sebastiano Barbieri, Oliver J Gurney-Champion, Remy Klaassen, and Harriet C Thoeny, “Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted mri,” *Magnetic resonance in medicine*, vol. 83, no. 1, pp. 312–321, 2020.
- [11] Edward W. Johnston, “Innovate: A prospective cohort study combining serum and urinary biomarkers with novel diffusion-weighted magnetic resonance imaging for the prediction and characterization of prostate cancer,” *BMC Cancer*, 2016.
- [12] E Panagiotaki, A Ianus, E Johnston, RW Chan, D Atkinson, D Alexander, et al., “Optimised verdict mri protocol for prostate cancer characterisation,” 2015.
- [13] Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun, “Deep residual learning for image recognition,” in *Proceedings of the IEEE CVPR*, 2016, pp. 770–778.
- [14] Dongwei Chen, Fei Hu, Guokui Nian, and Tiantian Yang, “Deep residual learning for nonlinear regression,” *Entropy*, vol. 22, no. 2, pp. 193, 2020.
- [15] Minghang Zhao, Shisheng Zhong, Xuyun Fu, Baoping Tang, and Michael Pecht, “Deep residual shrinkage networks for fault diagnosis,” *IEEE Transactions on Industrial Informatics*, vol. 16, no. 7, pp. 4681–4690, 2019.
- [16] Edward W Johnston, Elisenda Bonet-Carne, Uran Ferizi, Ben Yvernault, Hayley Pye, Dominic Patel, Joey Clemente, Wivijin Piga, Susan Heavey, Harbir S Sidhu, et al., “Verdict mri for prostate cancer: intracellular volume fraction versus apparent diffusion coefficient,” *Radiology*, vol. 291, no. 2, pp. 391–397, 2019.
- [17] Lianfa Li, Ying Fang, Jun Wu, and Jinfeng Wang, “Autoencoder based residual deep networks for robust regression prediction and spatiotemporal estimation,” *arXiv preprint arXiv:1812.11262*, 2018.