AUA-dE: An Adaptive Uncertainty Guided Attention for Diffusion MRI Models Estimation

Tianshu Zheng[1], Ruicheng Ba[1], Xiaoli Wang[2], Chuyang Ye[3], and Dan Wu[1(✉)]

[1] College of Biomedical Engineering and Instrument Science, Zhejiang University, Hangzhou, China
danwu.bme@zju.edu.cn

[2] School of Medical lmaging, Weifang Medical University, Weifang, China

[3] School of Integrated Circuits and Electronics, Beijing Institute of Technology, Beijing, China

Abstract. Diffusion MRI (dMRI) is a well-established tool for probing tissue microstructure properties. However, advanced dMRI models commonly have multiple compartments that are highly nonlinear and complex, and also require dense sampling in q-space. These problems have been investigated using deep learning based techniques. In existing approaches, the labels were calculated from the fully sampled q-space as the ground truth. However, for some of the dMRI models, dense sampling is hard to achieve due to the long scan time, and the low signal-to-noise ratio could lead to noisy labels that make it hard for the network to learn the relationship between the signals and labels. A good example is the time-dependent dMRI (TD-dMRI), which captures the microstructural size and transmembrane exchange by measuring the signal at varying diffusion times but requires dense sampling in both q-space and t-space. To overcome the noisy label problem and accelerate the acquisition, in this work, we proposed an *adaptive uncertainty guided attention for diffusion MRI models estimation* (AUA-dE) to estimate the microstructural parameters in the TD-dMRI model. We evaluated our proposed method with three different downsampling strategies, including q-space downsampling, t-space downsampling, and q-t space downsampling, on two different datasets: a simulation dataset and an experimental dataset from normal and injured rat brains. Our proposed method achieved the best performance compared to the previous q-space learning methods and the conventional optimization methods in terms of accuracy and robustness.

Keywords: Diffusion MRI · Noisy Data · Parameter Estimation · Uncertainty Attention

1 Introduction

Diffusion MRI (dMRI) is a powerful medical imaging tool for probing microstructural information based on the restricted diffusion assumption of the water molecules in biological tissues [7]. The conventional dMRI uses the apparent diffusion coefficient (ADC) to measure the diffusivity change, but it's not specific to

H. Greenspan et al. (Eds.): MICCAI 2023, LNCS 14227, pp. 142–151, 2023.
https://doi.org/10.1007/978-3-031-43993-3_14

microstructural properties. Recently, a series of advanced dMRI models, such as time-dependent dMRI (TD-dMRI) models [5,8,9], have been proposed to sample diffusion at different effective diffusion-times (so-called t-space), and related biophysical models have been developed to resolve the cellular microstructures.

Advanced dMRI models are typically multi-compartment models with highly non-linear and complex formulations. Accurate parameter estimation from such models requires dense q-space and/or t-space sampling. Deep learning techniques have been proposed to improve estimation accuracy from downsampled q-space data [2,13,15], by training networks to learn the relationship between downsampled diffusion signals and microstructural metrics from fully sampled q-space. However, TD-dMRI models require dense sampling in both q-space and t-space, which is challenging for clinical usage. To our best knowledge, previous works have not investigated downsampling models in t-space or joint q-t space.

The previous q-space learning networks [2,13,15], simply learned the mapping relationship between undersampled diffusion signals in q-space and diffusion model parameters. They neglected the "noisy label" problem, and thus, may suffer from training degradation due to failure to obtain valid information. Different from the annotated label, which is called ground-truth in the natural image, in the dMRI model estimation area, we used the parameters estimated from fully sampled q-space as the "gold standard" that actually suffer from estimation error due to noise in the data acquisition and also the limited number of samples in q-space. These two problems are particularly worth noting in TD-dMRI, which is known to have low SNR, and the errors could accumulate in the joint q-t space models. This imposed a vital problem for end-to-end t-space learning in practice. To address the challenge of learning from noisy data, we proposed an *adaptive uncertainty guided attention for diffusion MRI models estimation* (AUA-dE) based upon the previous AEME network [15] for estimating general dMRI model parameters.

In this work, we proposed a reweighting strategy to reduce the negative effects of noisy label based on uncertainty. Our contributions can be summarized below:

1. We brought up an important problem of the noisy label in dMRI model estimation which was not addressed before.
2. We proposed an attention-based sparse encoder to make the network focus on the key diffusion signal out of many q-space or t-space signals.
3. We developed an uncertainty-based reweighting strategy considering the uncertainty in both dMRI channels and spatial domain for microstructural estimation.
4. We proposed an end-to-end estimation strategy in both q-space and t-space with downsampled q-t space data.

In our work, we firstly demonstrated the effectiveness of our attention-based reweighting strategy on a simulation dataset, and then we evaluated our work with three different downsampling strategies (q-space, t-space, and q-t space) on a TD-dMRI dataset of normal and injured rat brains. We tested a TD-dMRI model which estimates the transmembrane water exchange time based on time-dependent diffusion kurtosis [10], here we named tDKI.

2 Method

2.1 *q-t* Space Sparsity

In this study, we extended the q-space learning model AEME [15] into the q-t space, and the signal can be represented as follows:

$$\mathbf{Y} = \Gamma \mathbf{X} \Upsilon^{\mathrm{T}} + \mathbf{H} \tag{1}$$

where $\mathbf{Y} = (y_1 \cdots y_V) \in \mathbb{R}^{K \times V}$($K$ is the number of diffusion signals including different b-values, gradients, and diffusion times (t_ds); and V is the size of patch), $y_v (v \in \{1, \cdots, V\})$ is the diffusion signal normalized by b0 at different t_d, and $\mathbf{X} \in \mathbb{R}^{N_\Gamma \times N_\Upsilon}$ is the matrix of the mixed sparse dictionary coefficients. $\mathbf{\Gamma} \in \mathbb{R}^{K \times N_\Gamma}$ and $\mathbf{\Upsilon} \in \mathbb{R}^{V \times N_\Upsilon}$ are decomposed dictionaries that encode the information in the mixed q-t domain and the spatial domain. $\mathbf{H}$ is the noise corresponding to $\mathbf{X}$. Then the sparse encoder can be formulated using the extragradient-based method similar to [15]:

$$\mathbf{X}^{n+\frac{1}{2}} = \mathcal{H}_M \left(AUA(\mathbf{X}^n, \mathbf{\Gamma}^{\mathrm{T}} \mathbf{Y} \mathbf{\Upsilon}, \mathbf{\Gamma}^{\mathrm{T}} \mathbf{\Gamma} \mathbf{X}^n \mathbf{\Upsilon}^{\mathrm{T}} \mathbf{\Upsilon}) \right) \tag{2}$$

$$\mathbf{X}^{n+1} = \mathcal{H}_M \left(AUA(\mathbf{X}^n, \mathbf{A_1} \mathbf{\Gamma}^{\mathrm{T}} \mathbf{Y} \mathbf{\Upsilon}, \mathbf{A_1} \mathbf{\Gamma}^{\mathrm{T}} \mathbf{\Gamma} \mathbf{X}^{n+\frac{1}{2}} \mathbf{\Upsilon}^{\mathrm{T}} \mathbf{\Upsilon}) \right) \tag{3}$$

where, $\mathbf{A_1}$ denotes a scalar matrix, $AUA(\cdot)$ is the adaptive uncertainty attention function, and $\mathcal{H}_M$ denotes a nonlinear operator corresponding to the threshold layer in Fig. 1(b):

$$\mathcal{H}_M (\mathbf{X}_{ij}) = \begin{cases} 0, & \text{if } \mathbf{X}_{ij} < \lambda \\ \mathbf{X}_{ij}, & \text{if } \mathbf{X}_{ij} \geq \lambda \end{cases} \tag{4}$$

So far, we can obtain the sparse representation of the signal.

2.2 Adaptive Uncertainty Attention Modelling

Uncertainty Attention Module. Inspired by the uncertainty modeling mechanism of Bayesian neural networks [6], we defined *uncertainty attention* (UA) to address the noisy label problem. For simplicity, we used Monte Carlo dropout to obtain the posterior distribution, other uncertainty quantification methods would also work. The basic attention module is adapted from CBAM [12], which includes channel and spatial attention. To better capture feature information, we employ mean and standard deviation pooling. Then the UA module is formulated according to the gray shaded box in Fig. 1(b). The top branch is the *channel uncertainty attention* (CUA) module, and the bottom branch is the *spatial uncertainty attention* (SUA) module.

Through the CUA module, the uncertainty reweighted sparse representation of the original signal $\mathbf{X}$ can be obtained: $\mathbf{X}_{\mathrm{CUA}} = CUA(\mathbf{X})$. $CUA(\cdot)$ is the channel-wise uncertainty attention function, which is used to model the uncertainty in the diffusion channels. The CUA of $\mathbf{X}$ can be obtained after the

stochastic forward with dropout: $\mathbf{U}_{\mathrm{C}} = Var(\mathbf{X}^k_{\mathrm{CUA}})$. $\mathbf{U}_{\mathrm{C}}$ is the uncertainty of the channel-wise information, $\mathbf{X}^k_{\mathrm{CUA}}$ is a cluster of different sparse representations of $\mathbf{X}$ after dropout in the CUA module, and k is the number of stochastic forwards.

Similarly, the SUA module can be defined as: $\mathbf{X}_{\mathrm{SUA}} = SUA(\mathbf{X}_{\mathrm{CUA}})$. $SUA(\cdot)$ is the spatial-wise uncertainty attention, which is used to model the uncertainty of the spatial-wise information. The SUA of $\mathbf{X}_{\mathrm{CUA}}$ can be estimated as follows: $\mathbf{U}_{\mathrm{S}} = Var(\mathbf{X}^k_{\mathrm{SUA}})$. $\mathbf{U}_{\mathrm{S}}$ is the uncertainty of the spatial-wise information, $\mathbf{X}^k_{\mathrm{SUA}}$ is a cluster of different sparse representations of $\mathbf{X}_{\mathrm{CUA}}$ after dropout in the SUA module.

For simplicity and efficiency, in practice, we combined these two kinds of uncertainty together to reweight the whole sparse representation $\mathbf{X}$ as: $\mathbf{U} = Var(\mathbf{X}^k)$. $\mathbf{U}$ is the uncertainty of the sparse representation, $\mathbf{X}^k$ is a cluster of different sparse representations of $\mathbf{X}$ after dropout in both CUA and SUA modules.

Adaptive Reweight Mechanism. In this work, we proposed an adaptive reweighting strategy by lowering the loss weight for a patch that may be corrupted by the noise. After the uncertainty $\mathbf{U}$ is approximated, We can set a weight tenor $\mathbf{U}_{\mathrm{w}}$ as below:

$$\mathbf{U}_{\mathrm{w}} = 1 - \mathbf{U} \tag{5}$$

Then, the impact of noise can be mitigated by an adaptive weight matrix $\mathbf{R}$:

$$\mathbf{R} = \begin{cases} t, & \text{if } \mathbf{U}_{\mathrm{w}} < t \\ \mathbf{U}_{\mathrm{w}}, & \text{if } \mathbf{U}_{\mathrm{w}} \geq t \end{cases} \tag{6}$$

where, t is a trainable parameter in the network, which can be modified adaptively. Then $\mathbf{X}$ will be reweighted by the $\mathbf{R}$ in the loss function as:

$$\mathcal{L} = ||M(\mathbf{R} \odot \mathbf{X}) - \widehat{P}||^2 \tag{7}$$

where, $\odot$ denotes the element-wise multiplication, $M(\cdot)$ is a mapping function corresponding to the mapping networks, $\widehat{P}$ is the observed label of the estimated dMRI model parameters.

Network Construction. Following the q-t space sparsity analysis and the adaptive uncertainty mechanism mentioned above, we can incorporate historical information [15] into Eq. 2 and Eq. 3 to formulate the *adaptive uncertainty attention sparse encoder* (AUA-SE).

$$\tilde{\mathbf{C}}^{n+\frac{1}{2}} = \mathbf{W}^{\mathrm{m1}}\mathbf{Y}\mathbf{W}^{\mathrm{s1}} + \mathbf{X}^n - \mathbf{S}^{\mathrm{m1}}\mathbf{X}^n\mathbf{S}^{\mathrm{s1}} \tag{8}$$

$$\mathbf{C}^{n+\frac{1}{2}} = AUA(\mathbf{F}^{n+\frac{1}{2}} \odot \mathbf{C}^n + \mathbf{G}^{n+\frac{1}{2}} \odot \tilde{\mathbf{C}}^{n+\frac{1}{2}}) \tag{9}$$

$$\mathbf{X}_{\mathrm{UA}}{}^{n+\frac{1}{2}} = \mathcal{H}_M\left(\mathbf{C}^{n+\frac{1}{2}}\right) \tag{10}$$

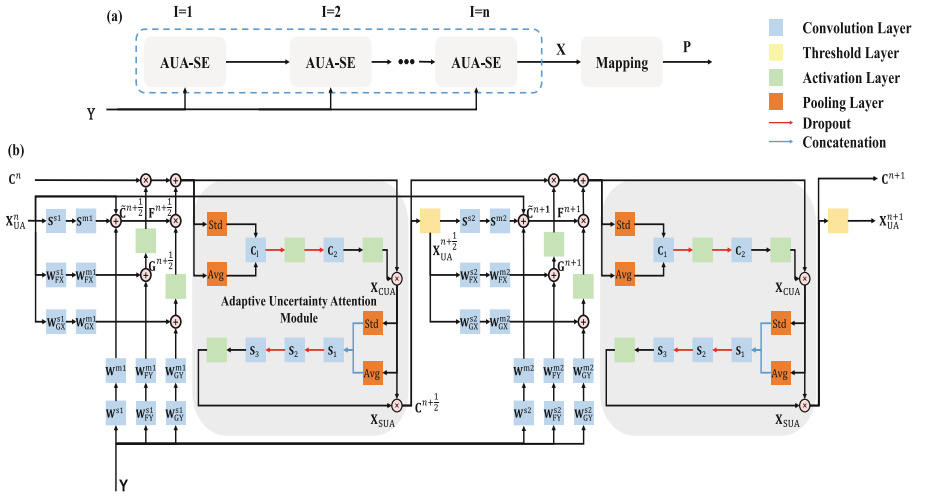

Fig. 1. (a) The overall structure of the proposed AUA-dE. The AUA-dE is made up of a number of AUA-SE units that are used to build the sparse representation $\mathbf{X}$ from the diffusion signals $\mathbf{Y}$, and then map the representation to the diffusion model parameters $\boldsymbol{P}$. **(b)** The AUA-SE unit is associated with the adaptive attention mechanism based on uncertainty (gray shaded box). $\mathbf{Y}$ is the input dMRI signal, and $\mathbf{X}_{\mathrm{UA}}^{n+1}$ is the sparse representation weighted by the uncertainty attention.

$$\tilde{\mathbf{C}}^{n+1} = \mathbf{W}^{\mathrm{m2}}\mathbf{Y}\mathbf{W}^{\mathrm{s2}} + \mathbf{X}^{n} - \mathbf{S}^{\mathrm{m2}}\mathbf{X}^{n+\frac{1}{2}}\mathbf{S}^{\mathrm{s2}} \tag{11}$$

$$\mathbf{C}^{n+1} = AUA(\mathbf{F}^{n+1} \odot \mathbf{C}^{n+\frac{1}{2}} + \mathbf{G}^{n+1} \odot \tilde{\mathbf{C}}^{n+1}) \tag{12}$$

$$\mathbf{X}_{\mathrm{UA}}{}^{n+1} = \mathcal{H}_{M}\left(\mathbf{C}^{n+1}\right) \tag{13}$$

where, $\mathbf{W}^{\mathrm{m1}} \in \mathbb{R}^{N_\Gamma \times K}$, $\mathbf{W}^{\mathrm{m2}} \in \mathbb{R}^{N_\Gamma \times K}$, $\mathbf{W}^{\mathrm{s1}} \in \mathbb{R}^{V \times N_\Upsilon}$, $\mathbf{W}^{\mathrm{s2}} \in \mathbb{R}^{V \times N_\Upsilon}$, $\mathbf{S}^{\mathrm{m1}} \in \mathbb{R}^{N_\Gamma \times N_\Gamma}$, $\mathbf{S}^{\mathrm{m2}} \in \mathbb{R}^{N_\Gamma \times N_\Gamma}$, $\mathbf{S}^{\mathrm{s1}} \in \mathbb{R}^{N_\Upsilon \times N_\Upsilon}$, and $\mathbf{S}^{\mathrm{s2}} \in \mathbb{R}^{N_\Upsilon \times N_\Upsilon}$ are the matrices that need to be learned through the neural network (convolution layers), where m and s correspond to the mixed domain and the spatial domain. $\tilde{\mathbf{C}}^{n+\frac{1}{2}}$, $\mathbf{C}^{n+\frac{1}{2}}$, $\tilde{\mathbf{C}}^{n+1}$, and $\mathbf{C}^{n+1}$ are four intermediate terms. The historical information is encoded through $\mathbf{F}^{n+\frac{1}{2}}$, $\mathbf{F}^{n+1}$, $\mathbf{G}^{n+\frac{1}{2}}$, and $\mathbf{G}^{n+1}$ and they can be defined as [15]. Further fundamental details on formulas and network architecture can be found in [14].

The overall network can be constructed by repeating the AUA-SE unit n times, and the output will be sent to the mapping networks for mapping the microstructure parameters, which consist of three fully connected layers of the feed-forward networks [15]. The overall structure is illustrated in Fig. 1 (a).

2.3 Dataset and Training

The tDKI model is defined as below [10]:

$$K(t) = \frac{2K_0\tau_m}{t} \times [1 - \frac{\tau_m}{t} \times (1 - e^{-\frac{t}{\tau_m}})] \tag{14}$$

where, $K(t)$ is the kurtosis at different t_d, K_0 is the kurtosis at $t_d = 0$, t is the t_d and τ_m is the water mixing time. $K(t)$ at individual t_d is obtained according to the DKE method [11].

Simulation Dataset was formed following the method of Barbieri et al., [1], where we plugged the varying parameters (K_0, τ_m) into Eq. 14 to obtain the kurtosis signal at different t_d (50, 100, and 200 ms). Parameter values were sampled uniformly from the following ranges according to the fitted values observed in rat brain data: K_0 between 0 and 3, τ_m between 2 and 200 ms, and a total of 409600 signals were generated. And 60% of them were used for training, 10% for validation, and 30% for testing.

In order to replicate the noisy label problem, we varied the noise level from SNR=10 to 30 in the t-space signal. In the training data, we used a Bayesian method modified from Gustafsson et al. [4] to obtain the label, and in the test data, we used "gold standard" label set from simulation.

Rat Brain Dataset was collected on a 7T Bruker scanner from 3 normal rats and 10 rats that underwent a model of ischemic injury by transient Middle Cerebral Artery Occlusion (MCAO). Diffusion gradients were applied in 18 directions per b-value at 3 b-values of 0.8, 1.5, and 2.5 ms/ μ m^2 and 5 t_d (50, 80, 100, 150, and 200 ms) with the following acquisition parameters: repetition time/echo time = 2207/18 ms, in-plane resolution = 0.3 × 0.3 mm^2, 10 slices with a slice thickness of 1 mm.

In order to get the gold standard, the DKE toolbox [11] was used to calculate the kurtosis at different t_d with the fully sampled q-space, and then used the Bayesian method mentioned above to estimate K_0 and τ_m with the fully sampled t-space. The dataset was downsampled with randomly selected 9 gradients at b = 0.8 and 1.5 ms/ μ m^2 in q-space and 3 t_d (50, 100, and 200 ms) in t-space. We mixed the 2 normal and 8 injured rats together for training (90%) and validation (10%), and 1 normal with 2 injured rats for testing.

Training. In this work, the dictionary size of our AUA-dE was set at 301, and the hidden size of the fully connected layer was 75. We used an early stopping strategy and a reducing learning rate with an initial learning rate of 1×10^{-4}. AdamW was selected as the optimizer with a batch size of 256. The dropout in the AUA-SE was 0.2 for 50 forward processes.

3 Experiments and Results

3.1 Ablation Study

We compared four different methods, including AEME [15] (baseline of the current network but without the attention or uncertainty mechanism), AEME with attention (the baseline with the attention), AEME with UA (the baseline with

UA but without adaptive mechanism), and AEME with AUA (our proposed AUA-dE) to evaluate the effectiveness of UA and AUA in mitigating the negative effect of the noisy label.

Here, we used the relative error (percentage of the gold standard) to compare different algorithms. Figure 2 showed that network structures with UA (AEME with UA and AUA-dE) achieved lower errors compared with other methods, especially in the lowest SNR environment. AUA-dE achieved the lowest estimation error because the threshold of uncertainty is a trainable parameter that does not need to be manually defined. Meanwhile, we performed paired t-tests for all comparative results, and AUA-dE showed significantly lower errors than all other methods.

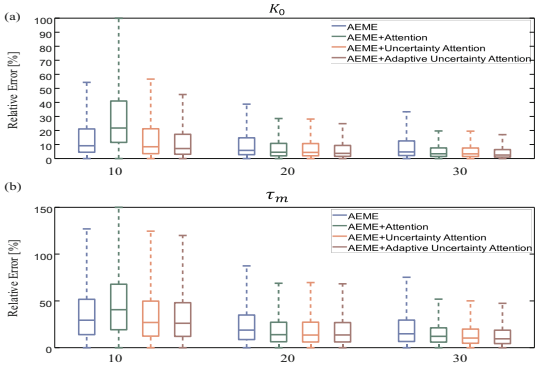

Fig. 2. Relative errors (percentage of the gold standard at the different SNR levels) of the estimated tDKI parameters at SNR levels from 10 to 30.

3.2 Performance Test

***q*-Space Downsampling.** We used our AUA-dE to estimate kurtosis at different t_d using the downsampled q-space, in comparison with an optimization based method DKE [11], a common q-space baseline q-DL [3], and the latest optimization based learning structure AEME [15]. Table 1 shows that our proposed AUA-dE achieved the lowest mean squared error (MSE) when compared to other methods for all t_d in both the normal and injured rat brain regions.

***t*-Space Downsampling.** In the t-space downsampling performance experiment, we used our AUA-dE to estimate the K_0 and τ_m based on $K(t)$ at varying t_d. We compared our method with the Bayesian method [4], q-DL, and AEME. From Table 2, it can be found our proposed AUA-dE achieved the lowest MSE compared with other methods in both normal and injured brain regions. Compared with previous methods, our error was only about 20% of the q-DL error in normal tissues and 40% in the injured regions.

Table 1. Evaluation of MSE on kurtosis using different methods on the downsampled q-space data with 9 different directions and 2 b-values (0.8 and 1.5 ms/μ m^2) under different t_d in both normal and injured rats brains. The lowest errors are in bold.

	Normal				**Injured**			
	DKE	***q*-DL**	**AEME**	**AUA-dE**	**DKE**	***q*-DL**	**AEME**	**AUA-dE**
50 ms	0.15	0.033	0.020	**0.017**	0.12	0.25	0.062	**0.052**
100 ms	0.18	0.050	0.027	**0.021**	0.15	0.65	0.078	**0.071**
200 ms	0.28	0.058	0.029	**0.027**	0.13	0.89	0.102	**0.087**

Table 2. Evaluation of MSE on tDKI parameters using different methods on the downsampled t-space data with 3 t_d (50, 100, and 200 ms) in both normal rats and injured regions. The lowest errors are in bold.

	Normal				**Injured**			
	Bayesian	***q*-DL**	**AEME**	**AUA-dE**	**Bayesian**	***q*-DL**	**AEME**	**AUA-dE**
K_0	0.122	0.078	0.062	**0.019**	0.138	0.065	0.063	**0.012**
τ_m	8567	1570	685	**303**	4569	763	537	**312**

***q-t* Space Downsampling.** In q-t space downsampling, we used our proposed AUA-dE to estimate the K_0 and τ_m with jointly downsampled q-t space data with 5 times acceleration (3 folds in q-space and 1.7 folds in t-space). Figure 3 showed the estimated K_0 and τ_m maps of an injured rat brain, using DKE+Bayesian, q-DL, AEME, and AUA-dE. Only AUA-dE was capable of capturing the abnormal rise in the injured cortex in the τ_m map (denoted by the red arrow), indicating the clinical potential of this method for diagnosis of ischemic brain injury.

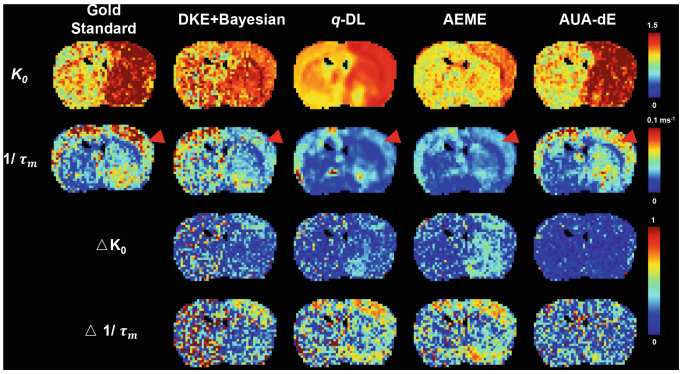

Fig. 3. The gold standard, estimated, and estimation errors of tDKI model parameters, based on DKE+Bayesian, q-DL, AEME, and AUA-dE (ours), in a injured rat brain with joint downsampled q-t space data. (Color figure online)

4 Conclusion

In this work, we proposed an *adaptive uncertainty guided attention for diffusion MRI models estimation* (AUA-dE) to address the noisy label problem in the estimation of TD-dMRI-based microstructural models. We tested our proposed network and module in a rat brain dataset and a simulation dataset. The proposed method showed the highest estimation accuracy in all of these datasets. Meanwhile, we demonstrated its performance on jointly downsampled *q*-*t* space data, for which previous algorithms did not work well with the highly accelerated setup (270/54). In the future, we will further investigate our proposed AUA module as a plug-in in different dMRI model estimation networks and also different dMRI models to test its generalizability and robustness.

References

1. Barbieri, S., Gurney-Champion, O.J., Klaassen, R., Thoeny, H.C.: Deep learning how to fit an intravoxel incoherent motion model to diffusion-weighted MRI. Magn. Reson. Med. **83**(1), 312–321 (2020)
2. Gibbons, E.K., et al.: Simultaneous NODDO and GFA parameter map generation from subsampled q-space imaging using deep learning. Magn. Reson. Med. **81**(4), 2399–2411 (2019)
3. Golkov, V., et al.: Q-space deep learning: twelve-fold shorter and model-free diffusion MRI scans. IEEE Trans. Med. Imaging **35**(5), 1344–1351 (2016)
4. Gustafsson, O., Montelius, M., Starck, G., Ljungberg, M.: Impact of prior distributions and central tendency measures on Bayesian intravoxel incoherent motion model fitting. Magn. Reson. Med. **79**(3), 1674–1683 (2018)
5. Jiang, X., et al.: In vivo imaging of cancer cell size and cellularity using temporal diffusion spectroscopy. Magn. Reson. Med. **78**(1), 156–164 (2017)
6. Kendall, A., Gal, Y.: What uncertainties do we need in Bayesian deep learning for computer vision? Adv. Neural Inform. Process. Syst. **30** (2017)
7. Mori, S., Zhang, J.: Principles of diffusion tensor imaging and its applications to basic neuroscience research. Neuron **51**(5), 527–539 (2006)
8. Panagiotaki, E., et al.: Noninvasive quantification of solid tumor microstructure using verdict MRI. Can. Res. **74**(7), 1902–1912 (2014)
9. Reynaud, O., Winters, K.V., Hoang, D.M., Wadghiri, Y.Z., Novikov, D.S., Kim, S.G.: Pulsed and oscillating gradient mri for assessment of cell size and extracellular space (pomace) in mouse gliomas. NMR Biomed. **29**(10), 1350–1363 (2016)
10. Solomon, E., et al.: Time-dependent diffusivity and kurtosis in phantoms and patients with head and neck cancer. Mag. Reson. Med. **89**(2), 522–535 (2022)
11. Tabesh, A., Jensen, J.H., Ardekani, B.A., Helpern, J.A.: Estimation of tensors and tensor-derived measures in diffusional kurtosis imaging. Magn. Reson. Med. **65**(3), 823–836 (2011)
12. Woo, S., Park, J., Lee, J.-Y., Kweon, I.S.: CBAM: convolutional block attention module. In: Ferrari, V., Hebert, M., Sminchisescu, C., Weiss, Y. (eds.) Computer Vision – ECCV 2018: 15th European Conference, Munich, Germany, September 8–14, 2018, Proceedings, Part VII, pp. 3–19. Springer, Cham (2018). https://doi.org/10.1007/978-3-030-01234-2_1

13. Ye, C.: Tissue microstructure estimation using a deep network inspired by a dictionary-based framework. Med. Image Anal. **42**, 288–299 (2017)
14. Ye, C., Li, X., Chen, J.: A deep network for tissue microstructure estimation using modified LSTM units. Med. Image Anal. **55**, 49–64 (2019)
15. Zheng, T., Zheng, W., Sun, Y., Zhang, Y., Ye, C., Wu, D.: An adaptive network with extragradient for diffusion MRI-based microstructure estimation. In: Wang, L., Dou, Q., Fletcher, P.T., Speidel, S., Li, S. (eds.) Medical Image Computing and Computer Assisted Intervention – MICCAI 2022: 25th International Conference, Singapore, September 18–22, 2022, Proceedings, Part I, pp. 153–162. Springer, Cham (2022). https://doi.org/10.1007/978-3-031-16431-6_15