

High resolution diffusion MRI (dMRI) data is often constrained by limited scanning time in clinical settings, thus restricting the use of downstream analysis techniques that would otherwise be available. In this work we develop a 3D recurrent convolutional neural network (RCNN) capable of super-resolving dMRI volumes in the angular (q-space) domain. Our approach formulates the task of angular super-resolution as a patch-wise regression using a 3D autoencoder conditioned on target b-vectors. Within the network we use a convolutional long short term memory (ConvLSTM) cell to model the relationship between q-space samples. Code for this project is available at <https://github.com/m-lyon/dMRI-RCNN>

Data

Human Connectome Project (HCP) dMRI data is used for both training and evaluation, and is initially processed with the standard HCP pre-processing pipeline [1]. Each 4D dMRI volume within each subject in the HCP dataset contains three shells of b-values 1000, 2000, and 3000. Each shell is processed independently and contains 90 diffusion directions, of which the low angular resolution dataset are subsampled from. Data are denoised using the patch2self algorithm [2]. Training, validation, and test datasets comprised of data from 27, 3, and 8 HCP subjects, respectively.

Model Architecture

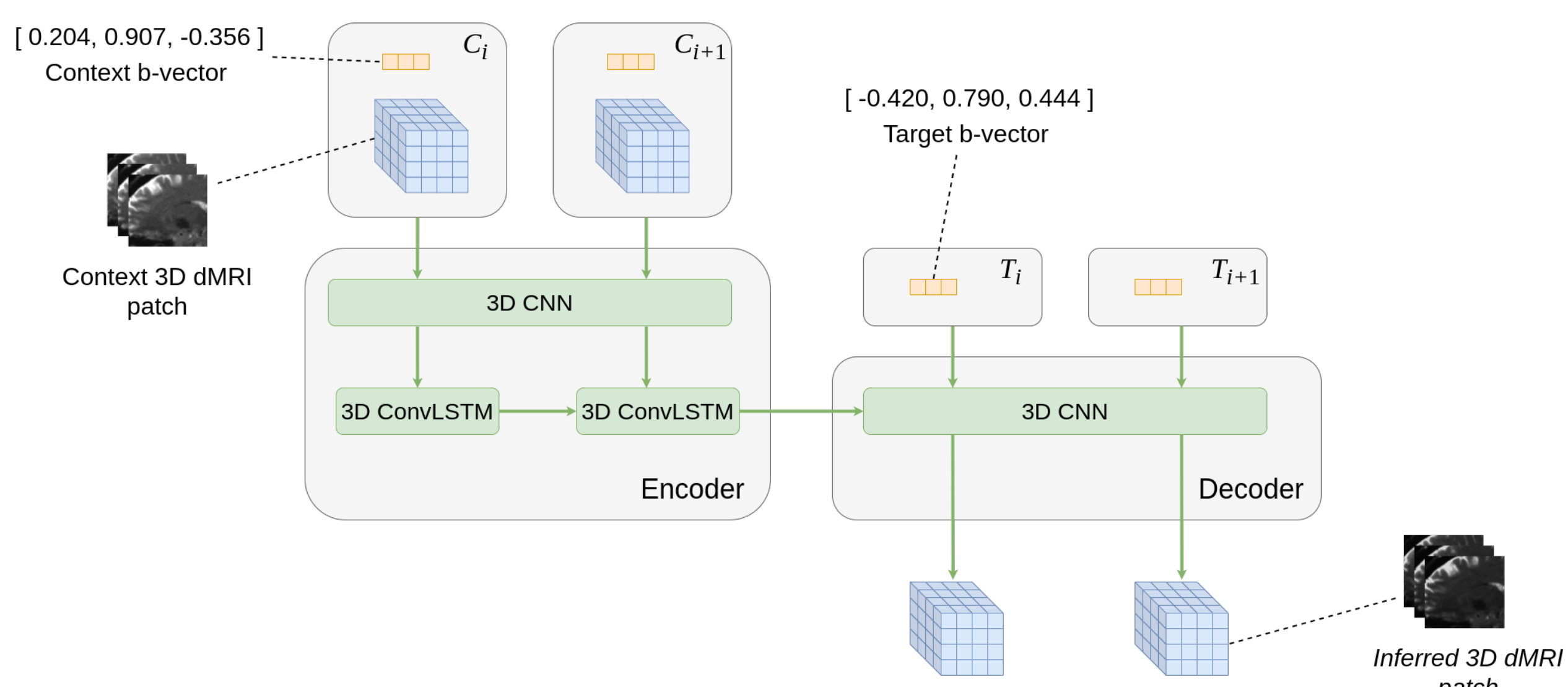


Figure 1: RCNN model design. Q-space context data C_i are given to the encoder sequentially until all context examples $C_{q_{in}}$ are seen. The internal hidden state of the ConvLSTM is passed to the 3D CNN decoder along with target data T_i to infer 3D dMRI patches along the given diffusion direction.

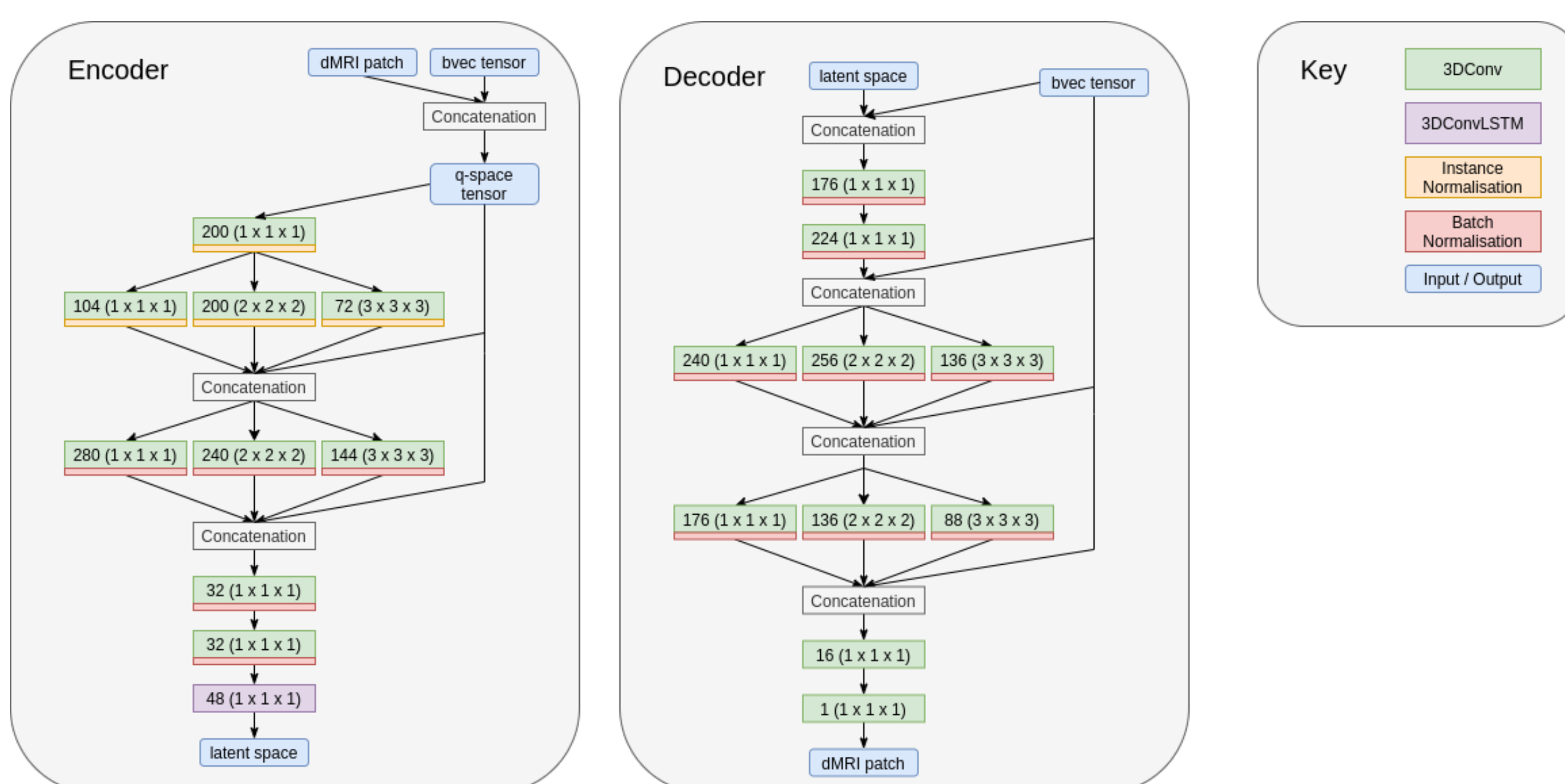


Figure 2: 3D RCNN model diagram with convolutional filter sizes and channel dimensions. Each convolution node specifies the number of filters used (left) and filter size (right).

Parallel convolution blocks are used within the encoder and decoder that apply a convolution operation to the same input in parallel. The non-pointwise kernels have padding applied prior to the convolution operation, such that the resultant shape is equal to the unpadded input tensor. These blocks are inspired by work done in [3]. All convolutional layers use a stride of $(1 \times 1 \times 1)$. Hyperparameters used for each layer were obtained by a hyperparameter search using the Hyperband algorithm [4]. Models were trained for 120 epochs using the optimizer Adam [5] with MAE loss function and a learning rate of 0.001.

References

- [1] Matthew F Glasser et al. “The minimal preprocessing pipelines for the Human Connectome Project”. In: *Neuroimage* 80 (2013), pp. 105–124.
- [2] Shreyas Fadnavis, Joshua Batson, and Eleftherios Garyfallidis. “Patch2Self: Denoising Diffusion MRI with Self-Supervised Learning”. In: *arXiv preprint arXiv:2011.01355* (2020).
- [3] Christian Szegedy et al. “Rethinking the inception architecture for computer vision”. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*. 2016, pp. 2818–2826.
- [4] Lisha Li et al. “Hyperband: A novel bandit-based approach to hyperparameter optimization”. In: *The Journal of Machine Learning Research* 18.1 (2017), pp. 6765–6816.
- [5] Diederik P Kingma and Jimmy Ba. “Adam: A method for stochastic optimization”. In: *arXiv preprint arXiv:1412.6980* (2014).

Model Performance

3D RCNN angular super-resolution was compared against a spherical harmonic (SH) interpolation baseline, a 1D variant of the same architecture, and a non-recurrent (CNN) version of the network.

Model	$q_{in} = 6, q_{out} = 84$		$q_{in} = 10, q_{out} = 80$		$q_{in} = 30, q_{out} = 60$	
	RMSE	MSSIM	RMSE	MSSIM	RMSE	MSSIM
SH Interpolation	119.0 ± 50.3	0.9460 ± 0.0419	65.1 ± 12.7	0.9854 ± 0.0044	63.8 ± 10.4	0.9867 ± 0.0033
1D RCNN	102.5 ± 31.6	0.9639 ± 0.0225	70.0 ± 15.1	0.9852 ± 0.0054	64.1 ± 10.6	0.9875 ± 0.0033
3D CNN	84.8 ± 15.9	0.9758 ± 0.0085	68.3 ± 12.6	0.9855 ± 0.0043	66.4 ± 10.4	0.9873 ± 0.0032
3D RCNN	78.4 ± 13.9	0.9787 ± 0.0071	63.4 ± 12.5	0.9870 ± 0.0040	63.4 ± 10.2	0.9876 ± 0.0032

Table 1: Average performance of angular super-resolution in eight subjects with $b = 1000$ across different models. Best results are highlighted in bold.

Model	$b = 1000$		$b = 2000$		$b = 3000$	
	RMSE	MSSIM	RMSE	MSSIM	RMSE	MSSIM
SH Interpolation	65.1 ± 12.7	0.9854 ± 0.0044	64.5 ± 9.1	0.9659 ± 0.0088	66.7 ± 13.5	0.9292 ± 0.0242
1D RCNN	70.0 ± 15.1	0.9852 ± 0.0054	51.3 ± 6.7	0.9766 ± 0.0056	48.1 ± 8.0	0.9566 ± 0.0133
3D CNN	68.3 ± 12.6	0.9855 ± 0.0043	50.0 ± 6.3	0.9779 ± 0.0048	70.7 ± 12.7	0.9350 ± 0.0157
3D RCNN	63.4 ± 12.5	0.9870 ± 0.0040	48.1 ± 6.2	0.9796 ± 0.0045	42.6 ± 5.5	0.9633 ± 0.0088

Table 2: Average performance of angular super-resolution in eight subjects with $q_{in} = 10$ and $q_{out} = 80$. Best results are highlighted in bold.

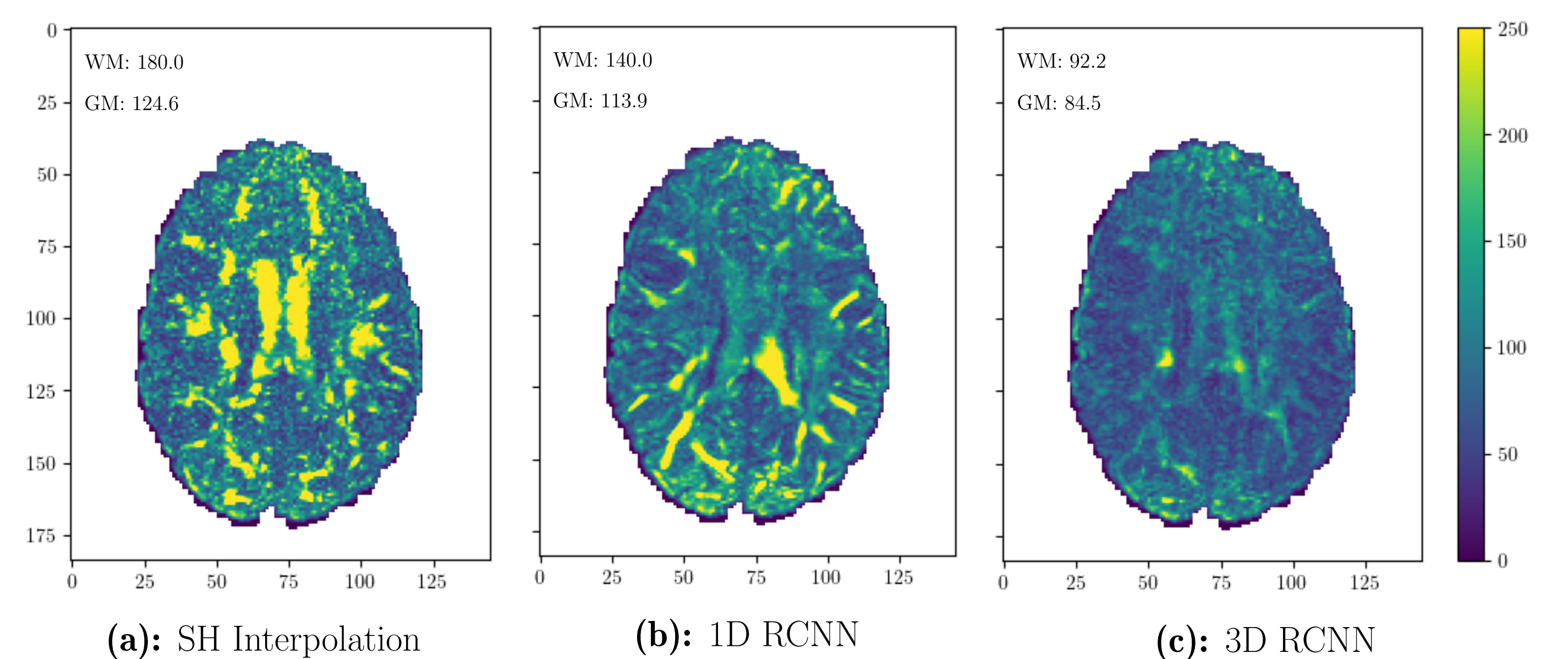


Figure 3: Axial slice of RMSE in one subject from the test dataset. Angular super-resolution is performed with $q_{in} = 6, q_{out} = 84$. Each pixel is the RMSE averaged across q_{out} directions. WM and GM values are the RMSE averaged across spatial voxels and q_{out} directions.

DTI Analysis

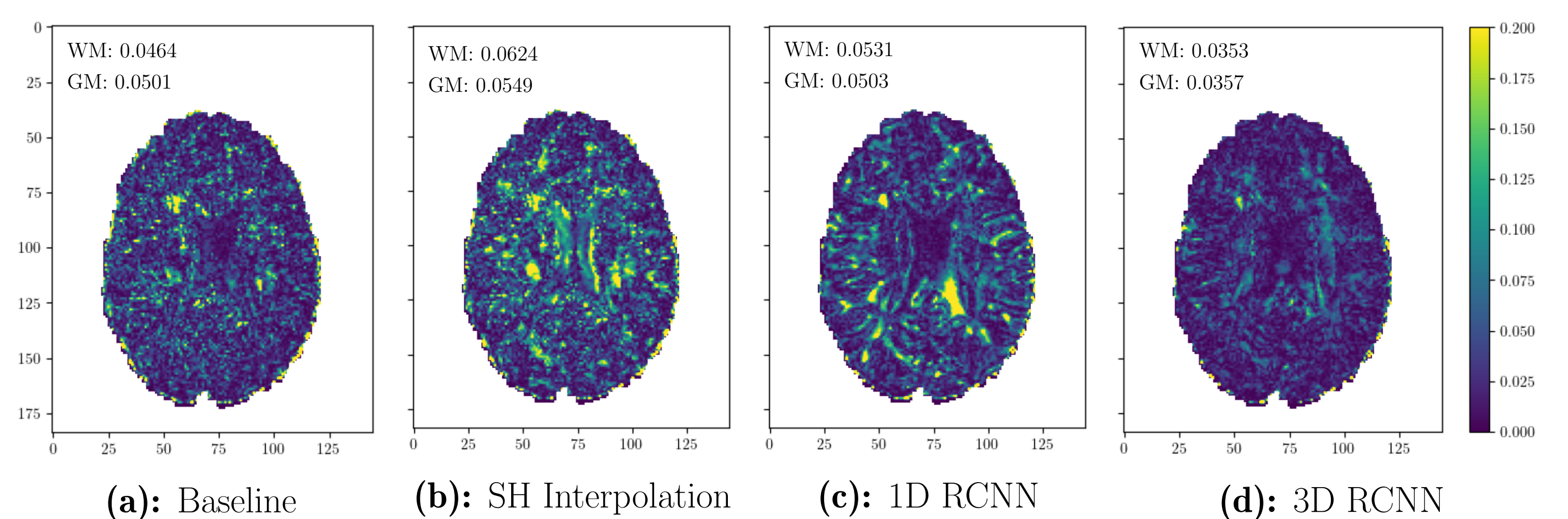


Figure 4: Axial slice of FA average error in one subject from the test dataset. Angular super-resolution is performed with $q_{in} = 6, q_{out} = 84$. WM and GM values are averaged across voxels only within the WM and GM mask, respectively. The baseline FA map is calculated from q_{in} volumes whilst other FA maps are derived from both q_{in} and q_{out} data.

Conclusions and Further Work

1. 3D RCNN architecture has the lowest error rates across various subsampling ratios and b-values when compared against other super angular resolution techniques.
2. 3D RCNN architecture can be used at very low angular resolution for DTI analysis.
3. Further work is needed to quantify robustness of this methodology in out-of-distribution datasets, as well as multi-shell angular super-resolution.