Uncertainty-Supervised Super-Resolution Deep Learning Network in Diffusion MRI

Chun Wang, Jiquan Ma*

Heilongjiang University, Haerbin 150080, China

* Corresponding Author Email: majiquan@hlju.edu.cn

Abstract. The research of uncertainty has shown great potential in the field of medical image processing. However, most of the research in the field of medical image is aimed at quantifying uncertainty. In this paper, we introduce an uncertainty supervised learning method. Specifically, we integrate the dropout variable reference and heterostatic noise model to estimate uncertainty and then guide super-resolution processing. Finally, we evaluate the enhancement effect of uncertainty supervised learning on super-resolution processing under the demonstration of DIQT model.

Keywords: Uncertainty, Diffusion MRI, Super-resolution.

1. Introduction

In recent years, deep learning has been successfully applied in the field of medical image super-resolution processing (such as [1, 2]). However, the existing research [3, 4] is rarely use the uncertainty of the results to further guide the super-resolution processing. In medical diagnosis, it is important to form a risk mechanism of quantitative uncertainty for the clinical application of dMRI image super-resolution. Uncertainty estimation is mainly divided into two categories. The first is aleatoric uncertainty, and the second is epistemic uncertainty. In this work, we proposed to an uncertainty supervised learning method combined with the uncertainty estimation to guide super-resolution processing. We use heteroscedastic model and dropout variational inference to estimate uncertainty and predict high-resolution image respectively. Finally, we validate the ability of uncertainty estimation to improve the super-resolution results.

2. Uncertainty-supervised super-resolution learning model

In this section, we first introduce the DIQT [1] and 3D-ESPCN [5] SR model. Then we introduce a two-branch network for uncertainty estimation to output epistemic uncertainty and aleatoric uncertainty. Finally, the output variance of the uncertainty estimation is used to reconstruct a new loss function to further improve the super-resolution accuracy.

2.1. DIQT baseline model

IQT transfers the high-quality image information obtained from the high-performance scanner to the low-quality standard image that is easily obtained in the clinic. The DIQT model is improved based on the IQT, and applies checkpoint technology and RevNets reversible network to optimize the training speed of the network. By adding multiple RevNet blocks in front of each ESPCN layer to expand the depth of the network structure, the accuracy of the prediction model is effectively improved with negligible memory growth.

2.2. 3D-ESPCN SR Network

The ESPCN model proposed by Shi [5] is friendly for super-resolution processing of 3D image as far as its improved computational performance and reduced computational cost are concerned. Specifically, the special feature of the full convolutional network ESPCN is the shuffle operation in which the individual channel dimensions of the feature are recombined to output high-resolution 3D images correspondingly. Figure 1 shows the 3D illustration of ESPCN, where the network consists of a combination of three convolutional layers and Relu.

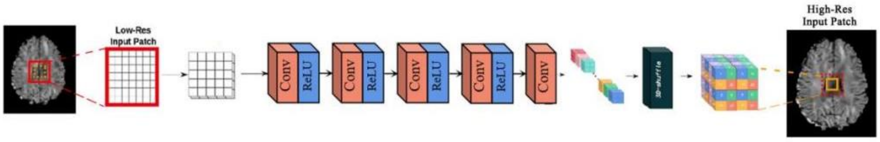

Fig. 1 3D-ESPCN SR Network

2.3. Uncertainty estimation model

Traditional uncertainty estimation models [4, 6] tend to weaken the impact of pixels with large variance in the setting of loss function. However, in medical images, the reconstruction of cerebral cortex folds and edges is more worthy of our attention. Therefore, we amplify the influence of areas with large variance, fuse the epistemic and aleatoric uncertainty model, and use the variance to further guide the super-resolution processing.

2.3.1 Estimation of epistemic uncertainty

In order to quantify the cognitive uncertainty in Bayesian networks, we set up a Gaussian distribution W~N (0, 1) for its existing weights. The difference from deterministic networks is that it optimizes the distribution of parameter weights. However, the inference process of Bayesian network is complex, and the evaluation of posterior distribution is difficult. Dropout variational inference, a commonly used approximate inference method applied to large models, can be a good solution to the above problem. It uses the posterior distribution to learn the dropout rate of each weight layer to approximate the posterior sample, which can be understood as the variational Bayesian approximation. In the regression task, a mean network is usually trained to predict high-resolution images, so as to quantify data uncertainty and capture prediction variance. The first term is the noise amount, and the second term is the prediction uncertainty. The negative log-likelihood function can be expressed as

$$-\log p(y_i \mid f(x_i)) \propto \frac{1}{2\sigma^2} \| y_i - f(x_i) \|^2 + \frac{1}{2} \log \sigma^2, \tag{1}$$

2.3.2 Estimation of aleatoric uncertainty

The calculation of cognitive uncertainty is mainly performed by approximating the distribution, while the calculation of arbitrary uncertainty is also performed by adjusting the noise parameter. The difference between the homoskedastic model and the heteroskedastic model is whether the noise varies with the input. The output is mainly modeled by the heteroscedastic model, which is more effective in regions with relatively high noise, and the noise parameter in Bayesian networks are more highly regarded and can be treated as a function of the data.

Variational inference differs from approximate distribution, which deals in weight parameters. While variational inference is a MAP inference that determines unique model parameter values, this approach is destined to lack the estimation of cognitive uncertainty. So, the calculation of both uncertainties needs to be combined and model decay is applied to the heteroskedastic model.

2.3.3 Epistemic uncertainty and aleatoric uncertainty fusion model

In this section, we build the prediction model between the low-resolution image and the high-resolution image, x_i denoting the low-resolution image, y_i denoting the high-resolution image, $f(\cdot)$ denotes the prediction SR network, and $f(x_i)$ is the mean value of the estimated SR image output, and the uncertainty is represented by the σ_{i} variance. To ensure the stability of the results, we construct a Gaussian likelihood function and derive the loss output by minimizing the objective function, as follow

$$L_{BNN(\theta)} = \frac{1}{N} \sum_i \sigma_i^{-2} \| y_i - f(x_i) \|_1 + \log \sigma_i, \tag{2}$$

N represents the number of output pixels and σ_i is the output of BNN network variance. The loss function is determined by uncertainty and regularization terms respectively. To ensure the stable and effective output of the variance and avoid the effect that the denominator returns to zero, the loss function is

$$L_{BNN_1} = \frac{1}{N} exp(-s_i) \| y_i - f(x_i) \|_1 + s_i, \tag{3}$$

In order to make the uncertainty guide the precision of super-resolution image output, the dual network structure is shown in Figure 2. We modify the weight of the uncertainty in the loss function to improve $\exp(-s_i)$ to a monotone increasing function $\theta_t = s_i - \min(s_i)$, network 1 is responsible for obtaining the mean value $f(x_i)$ and uncertainty σ_i under the training of loss function $\theta_t = s_i - \min(s_i)$, and input σ_i to network 2 to output high-resolution images under the training of loss function L_{BNN_2}, then the new loss function is

$$L_{BNN_2} = \frac{1}{N} \sum_{i=1}^{N} \theta_t \| y_i - f(x_i) \|_1, \tag{4}$$

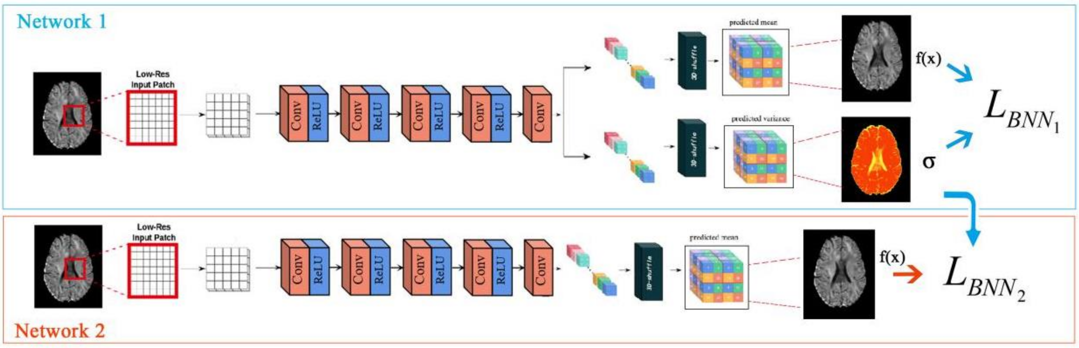

Fig. 2 Uncertainty-supervised super-resolution learning model

3. Experiments and results

We use 40 subjects in the Human Con-nectome Project (HCP) cohort and each subject includes 90 DWIs of voxel size $1.25^3 mm^3$ with b=1000 s/mm^2. Our dataset is divided training and validation sets into ratio of 4:1 and the input patch are $11*11*11$. The training uses the pytorch model, ADAM optimizer and RMSE loss. The learning rate is 10^{-5}, batch size is 16, and the training is finished after 200 epochs.

In order to highlight the image enhancement effect of uncertainty supervised learning, we evaluate the RMSE in the brain and out of the brain respectively of the high-resolution images output by Network I and Network II. And to validate the improved image accuracy, we compare the final super-resolution results with 3D-ESPCN (baseline) network and DIQT network in Table 1.

Table 1. Comparison of interior and exterior RMSE on 4 models

Models	RMSE Interior	RMSE Exterior
3D-ESPCN (baseline)	6.324 ± 0.19	13.750 ± 0.96
DIQT	5.64 ± 0.31	12.51 ± 1.37
Network 1	5.34 ± 0.23	11.96 ± 1.28
Network 2	5.03 ± 0.18	10.89 ± 1.19

In addition, we compare the visualization results before and after uncertainty supervised learning in Figure 3. The high-resolution results output by network 2 are significantly better than the image quality in network1. And we estimate the uncertainty in the final output HR results.

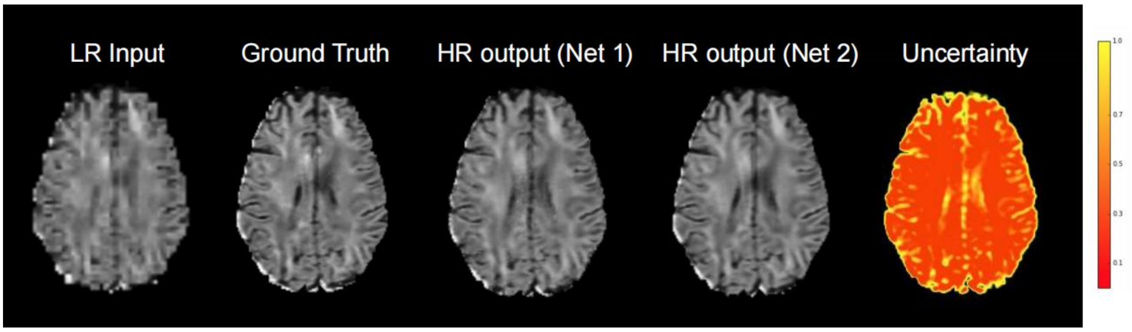

Fig. 3 A visualisation of mean diffusivity maps on an axial slice.

4. Conclusion

In this work, we use 3D-ESPCN architecture combined with heteroscedasticity model and dropout to estimate uncertainty and predict high-resolution image respectively, and use uncertainty quantification to improve the precision of super-resolution. However, the focus of our work is not the accuracy of super-resolution results. We aim to study the effect of uncertainty estimation in the super-resolution of diffusion magnetic resonance images, and hope to extend uncertainty modeling to more medical image processing tasks in the future.

Acknowledgements

The authors gratefully acknowledge the financial support from Natural Science Foundation of Heilongjiang Province (LH2021F046).

References

[1] Blumberg S B, Tanno R, Kokkinos I, et al. Deeper image quality transfer: Training low-memory neural networks for 3d images [C]//Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part I. Springer International Publishing, 2018: 118-125.

[2] Chatterjee S, Sciarra A, Dünnwald M, et al. ShuffleUNet: Super resolution of diffusion-weighted MRIs using deep learning [C]//2021 29th European Signal Processing Conference (EUSIPCO). IEEE, 2021: 940-944.

[3] Tanno R, Worrall D E, Ghosh A, et al. Bayesian image quality transfer with CNNs: exploring uncertainty in dMRI super-resolution [C]//Medical Image Computing and Computer Assisted Intervention− MICCAI 2017: 20th International Conference, Quebec City, QC, Canada, September 11-13, 2017, Proceedings, Part I 20. Springer International Publishing, 2017: 611-619.

[4] Tanno R, Worrall D E, Kaden E, et al. Uncertainty modelling in deep learning for safer neuroimage enhancement: Demonstration in diffusion MRI [J]. NeuroImage, 2021, 225: 117366.

[5] Shi W, Caballero J, Huszár F, et al. Real-time single image and video super-resolution using an efficient sub-pixel convolutional neural network [C]//Proceedings of the IEEE conference on computer vision and pattern recognition. 2016: 1874-1883.

[6] Ye C, Li Y, Zeng X. An improved deep network for tissue microstructure estimation with uncertainty quantification [J]. Medical image analysis, 2020, 61: 101650.