Contrast Synthesis with Uncertainty Estimation using Quantile Regression and U-Nets

*****

****

Abstract. This research presents a novel approach to synthesize missing MRI contrasts with uncertainty estimation using a U-Net framework based on quantile regression. Multi-contrast brain MRI data is vital for comprehensive analyses in neurological conditions such as traumatic brain injury (TBI) and epilepsy, as different MRI contrasts provide complementary information on brain structure and pathology. However, challenges in clinical settings often lead to incomplete multi-contrast MRI datasets, which complicates downstream analyzes, particularly in deep learning applications that require consistent data across all contrasts. Our method addresses this issue by generating missing MRI contrasts from available contrasts while quantifying the uncertainty of these synthesized images. We employ a U-Net architecture with quantile regression to predict both the missing contrasts and the associated uncertainty, providing more reliable synthetic images for clinical and research applications. Experimental results demonstrate that this approach offers insights into the reliability of the generated contrasts and enhances interpretability in high-stakes applications.

Keywords: Quantile regression · Deep learning · Contrast synthesis · Uncertainty estimation · Missing data · MRI

1 Introduction

Brain MRI data often consist of multiple contrasts, such as T1-weighted, T2-weighted, FLAIR, which are generated by varying the imaging parameters to emphasize different tissue properties like proton density, relaxation times, or the presence of contrast agents[1]. Acquiring multi-contrast MRI data is crucial because different MRI contrasts capture unique and complementary information about brain structure and pathology[2][3]. For example, T1-weighted images are ideal for assessing anatomical structures[4], T2-weighted images are sensitive to water content and can detect edema[5], and FLAIR images are useful for identifying lesions adjacent to cerebrospinal fluid[6]. Together, these contrasts provide a comprehensive view of the brain, enabling more accurate diagnoses, prognosis assessments, and detailed analyses for conditions like traumatic brain injury (TBI) and epilepsy[7][8]. However, acquiring all desired MRI contrasts is often impractical due to long scanning times, patient discomfort, equipment limitations, and high costs[9][10]. This can lead to datasets where some contrasts

are missing, such as TRACK-TBI data[11], posing a challenge for multi-contrast analysis[12].

1.1 Contrast Synthesis for Handling Missing MRI Contrasts

The absence of specific MRI contrasts complicates further analysis, especially in deep learning, where models often expect consistent input data with all contrasts present[13][14]. Missing contrasts introduce variability in data, impacting the model's performance, particularly in tasks like segmentation or classification, which rely on the full set of MRI contrasts for optimal accuracy[15][16]. To address this, contrast synthesis techniques are used to synthesize the missing MRI contrasts from the available ones[17]. By learning relationships between the different contrasts, these models generate synthetic images that approximate the missing contrasts, enabling complete multi-contrast analysis[18][19].

The contrast synthesis methods to generate MRI data can be categorized into three different groups: mathematical image transformation, physics-based, and data-driven approaches[20]. In this study, we follow a data-driven approach, which is similar to mathematical transformation, except that the transformation is not hand designed but learned from a data pool using deep neural networks. The dataset used in this study contains complete contrast information, facilitating the model's ability to effectively learn the relationships between different MRI contrasts for data reconstruction.

1.2 Uncertainty Estimation for Synthesized MRI Contrasts

It is critical to estimate the uncertainty of these synthesized contrasts since they are based on predictions[21][22]. Uncertainty estimation helps quantify the model's confidence in the generated images, which is essential for downstream applications like medical diagnosis, where high stakes decisions rely on reliable data. This is particularly important when working with deep neural networks (DNNs), as these models are often poorly calibrated[23]. In addition to capturing model limitations, uncertainty measures can also provide valuable insights into the inherent variability present in the data[24]. Incorporating uncertainty into contrast synthesis provides clinicians and researchers with a better understanding of potential errors, allowing for more informed and cautious interpretations of the synthetic data[25][26].

1.3 Quantile Regression for Capturing Uncertainty

Quantile regression is a method that estimates conditional quantiles of the response variable, rather than predicting the mean as in standard regression[27][28]. It captures a fuller range of possible outcomes, making it especially valuable for uncertainty estimation[29]. In contrast synthesis methods for MRI, quantile regression can generate multiple estimates for a missing contrast, providing a distribution of possible outcomes instead of a single prediction. This enables the

model to quantify uncertainty by assessing the spread between lower and upper quantiles, which reflects the model's confidence in the synthesized image[29]. Unlike mean/variance regression, which focuses on average predictions and assumes a symmetric error distribution, quantile regression does not rely on any distributional assumptions and can capture asymmetric and more complex patterns in data. This flexibility is advantageous in medical imaging, where pixel-level uncertainties are often non-uniform and where reliable confidence intervals are crucial for accurate diagnosis and clinical interpretation[30].

2 Methods

2.1 Dataset and Preprocessing

The publicly available BraTS 2017 dataset [31] was utilized for training and evaluating our QR U-Net model as well as the downstream brain tumor segmentation task. This dataset comprises complete T1w, T2w, T1Gd, and FLAIR MRI scans from 388 subjects diagnosed with glioblastoma. Subject ID 65 was excluded from the analysis due to a significantly large tumor region that disrupted the model's ability to learn its representation effectively. All images in the dataset are pre-aligned and have an identical isotropic voxel size of 1 mm. Tumor annotations for each subject were obtained by aggregating voxel-wise labels for the enhancing tumor, peritumoral edema, and necrotic/non-enhancing tumor core regions. For both the QR U-Net model and the downstream segmentation task, the dataset was divided into a training set of 310 subjects and a test set of 77 subjects.

The following preprocessing steps were applied to both the training and test datasets. Tumor labels were first converted into binary one-hot vectors. The MRI scans were reoriented according to the RAS axcode, and non-zero intensity values were normalized to have zero mean and unit standard deviation for each channel. For the training data, random data augmentation was performed, which included cropping to a region of interest (ROI) of size $64\times64\times64$ with a randomly selected center, flipping along each axis with a 50% probability, and scaling and shifting the intensity values by factors uniformly sampled from the range 0.9 to 1.1. For the tumor segmentation model, instead of directly cropping to an ROI of $64\times64\times64$, the data was initially cropped to a larger ROI of $224\times224\times144$. This was followed by resizing to $64\times64\times64$, ensuring that the tumor region was included in most of the training samples.

2.2 Quantile Regression

To reconstruct missing MRI contrasts, our method utilizes a 3D U-Net architecture integrated with quantile regression, a robust statistical approach that models specific conditional quantiles of the target variable rather than the mean. This technique provides not only the central estimate of the data but also captures the uncertainty associated with the reconstruction process. We employed

Quantile regression to estimate the conditional median (50th percentile) and additional quantiles, specifically the upper quantile Q_H (e.g., 95th percentile) and lower quantile Q_L (e.g., 5th percentile). The median quantile represents the most likely reconstructed value, minimizing the influence of outliers, whereas the upper and lower quantiles capture the range of uncertainty.

To train the model, we employed the quantile loss function, which is defined for a given quantile τ as:

$$\mathcal{L}_\tau(y, \hat{y}) = \sum_{i=1}^{n} \max(\tau(y_i - \hat{y}_i), (\tau - 1)(y_i - \hat{y}_i)).$$

Here, y_i is the ground truth, and $\hat{y}_i$ is the predicted quantile value. The loss encourages the network to predict values that correspond to the desired quantiles by penalizing over- and underestimation appropriately.

The final loss function combines the quantile losses for all contrasts:

$$\mathcal{L}_{\text{total}} = \sum_{\text{contrast}} (\mathcal{L}_{Q_M} + \mathcal{L}_{Q_H} + \mathcal{L}_{Q_L}).$$

2.3 Implementation Details

Our method was implemented using PyTorch [32] and MONAI [33] frameworks and executed on a single NVIDIA A100 GPU. A 3D U-Net architecture [34] was designed for the reconstruction task, featuring a channel configuration of 4, 8, and 16 in the top, middle, and bottom blocks, respectively. The model employed a stride length of 2 and included two residual units, resulting in approximately 25k parameters. The network accepted 4 input channels, corresponding to the four MRI modalities, and produced 12 output channels, representing the 3 predicted quantiles for each modality. For the downstream brain tumor segmentation task, the implementation adhered to the publicly available code in the Project-MONAI GitHub repository.

Masking To simulate missing data inputs and train the model for reconstructing MRI contrasts, the dataset was systematically preprocessed to mask specific contrasts in each sample. The masking process was as follows:

Step 1. Probability Lookup Table Initialization: A predefined probability table was established to represent the approximate likelihood of a particular contrast being unavailable. Based on insights from prior studies [35] [36], the probabilities were set to 40% for FLAIR, 12% for T1w, 30% for T1Gd, and 15% for T2w. These values reflect a general approximation, acknowledging that the availability of MRI contrasts can vary across different datasets and studies.

Step 2. Mask Generation: A 2D Boolean array, representing the presence or absence of each contrast for every input sample, was generated. Each element in the array was determined via a Bernoulli trial using the corresponding probability from the lookup table. A value of `True` indicated that the contrast was masked for that sample.

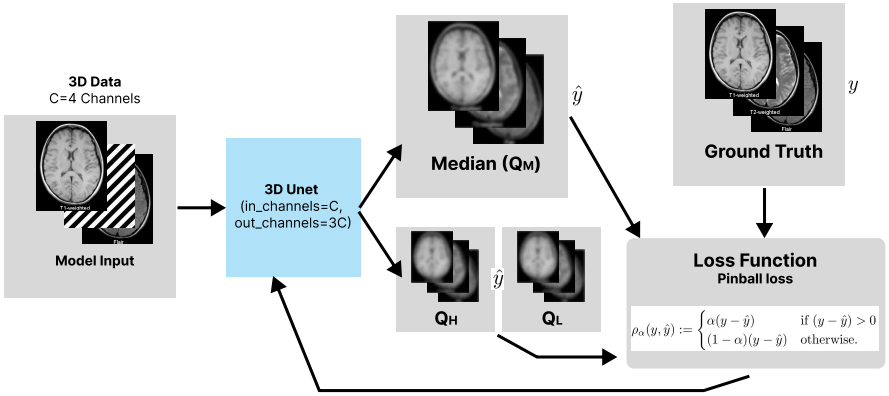

Fig. 1. Workflow illustrating the QR-UNet model: MRI contrasts with masked inputs are processed by a 3D U-Net, which outputs three quantiles (median, upper, and lower). Loss is computed for each quantile against the ground truth, guiding updates to the model parameters.

Step 3. Ensuring Validity of Samples: To avoid generating samples where all contrasts were masked (which would render the data unusable), an iterative correction process was employed. For samples where all contrasts were masked, Step 2 was reapplied until no such cases remained. Typically, this procedure converged within 2–3 iterations.

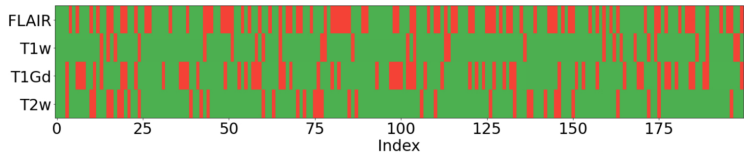

Fig. 2. Illustration of the first 200 samples of the masking matrix used during training, showing the distribution of missing MRI contrasts across the dataset. Each row represents a sample, and each column corresponds to a contrast, with masked contrasts indicated.

The resulting masking matrix, generated using this algorithm, is illustrated in Figure 2, highlighting the distribution of missing contrasts across the training dataset.

3 Experiments and Results

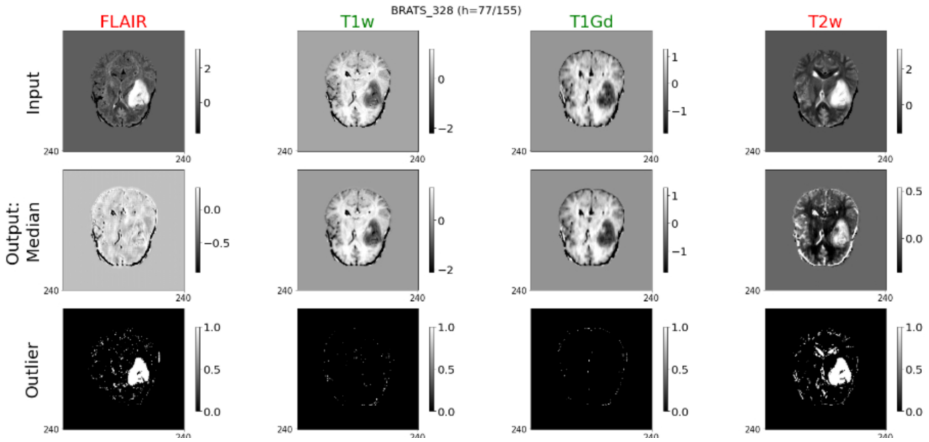

Fig. 3. The output of the QR-UNet model, evaluated on the 77th slice in the axial view of the BRATS_328 MRI image. In this case, the FLAIR and T2w contrasts were masked (shown here for reference), while the T1w and T1Gd contrasts were provided as input to the model. The second row displays the median output (τ=0.5), and the third row highlights the outlier regions, identified by filtering pixels with intensities exceeding the upper quantile (τ=0.977) or below the lower quantile (τ=0.023). The outlier filtering effectively captures the tumor region in an unsupervised manner.

Table 1. Comparison of PSNR and SSIM values for different MRI combinations.

	T1	T1+T2	T1+T2+FLAIR
PSNR	34.8	35.3	35.3
SSIM	0.898	0.910	0.909

4 Conclusion

References

1. BA Bell, DM Kean, HL MacDonald, GH Barnett, RHB Douglas, MA Smith, CNJ McGhee, JD Miller, JL Tocher, and JJK Best. Brain water measured by magnetic resonance imaging: correlation with direct estimation and changes after mannitol and dexamethasone. *The Lancet*, 329(8524):66–69, 1987.

Table 2. Evaluation scores of gT1Gd (ground-truth) with sT1Gd (synthetic), Zero (empty), T1w and mT1Gd (mean)

	sT1Gd	**Zero**	**T1w**	**mT1Gd**
MSE	**0.0008**	0.0125	0.0048	0.0021
MAE	**0.0116**	0.0425	0.0211	0.0159
PSNR	**31.7**	19.3	26.3	27.2
SSIM	0.8719	0.8198	**0.9412**	0.8843

Table 3.

	(4, 8, 16)	(64, 128, 256)
MSE	0.0009	**0.0008**
MAE	0.0121	**0.0116**
PSNR	31.3	**31.7**
SSIM	0.8396	**0.8719**

Table 4.

	MSE	MAE
MSE	**0.0009**	0.0012
MAE	**0.0121**	0.0138
PSNR	**31.7**	28.2
SSIM	**0.8396**	0.7976

Table 5. Comparison with other Medical Image Synthesis Methods

	Ours	HRC	MedGAN
MAE	**0.012**	0.029	N/A
PSNR	**31.3**	30.0	27.0
SSIM	0.872	**0.923**	0.901

Table 6. Comparison of PSNR and SSIM values for different MRI combinations.

	T1	**T1+T2**	**T1+T2+FLAIR**
PSNR	34.8	35.3	35.3
SSIM	0.898	0.910	0.909

2. Shigeki Aoki, Yasushi Sasaki, Tohru Machida, and Hisaya Tanioka. Contrast-enhanced mr images in patients with meningioma: importance of enhancement of the dura adjacent to the tumor. *American journal of neuroradiology*, 11(5):935–938, 1990.
3. Alexander M Bronstein, Michael M Bronstein, Michael Zibulevsky, and Yehoshua Y Zeevi. " unmixing" tissues: sparse component analysis in multi-contrast mri. In *IEEE International Conference on Image Processing 2005*, volume 2, pages II–

1282. IEEE, 2005.
4. F Feliner, K Holl, P Held, C Fellner, R Schmitt, and H Böhm-Jurkovic. A t1-weighted rapid three-dimensional gradient-echo technique (mp-rage) in preoperative mri of intracranial tumours. *Neuroradiology*, 38:199–206, 1996.
5. Shoji Naruse and Kimiyoshi Hirakawa. Brain edema studied by magnetic resonance. In *Seminars in Neurology*, volume 6, pages 53–64. © 1986 by Thieme Medical Publishers, Inc., 1986.
6. R Ashikaga, Y Araki, and O Ishida. Mri of head injury using flair. *Neuroradiology*, 39:239–242, 1997.
7. Ramon Diaz-Arrastia, Mark A Agostini, Alan B Frol, Bruce Mickey, James Fleckenstein, Eileen Bigio, and Paul C Van Ness. Neurophysiologic and neuroradiologic features of intractable epilepsy after traumatic brain injury in adults. *Archives of neurology*, 57(11):1611–1616, 2000.
8. Clifford R Jack Jr. Magnetic resonance imaging in epilepsy. In *Mayo Clinic Proceedings*, volume 71, pages 695–711. Elsevier, 1996.
9. P Van der Meulen, JP Groen, AMC Tinus, and G Bruntink. Fast field echo imaging: an overview and contrast calculations. *Magnetic resonance imaging*, 6(4):355–368, 1988.
10. Jürgen Hennig, A Nauerth, and HRARE Friedburg. Rare imaging: a fast imaging method for clinical mr. *Magnetic resonance in medicine*, 3(6):823–833, 1986.
11. John K Yue, Mary J Vassar, Hester F Lingsma, Shelly R Cooper, David O Okonkwo, Alex B Valadka, Wayne A Gordon, Andrew IR Maas, Pratik Mukherjee, Esther L Yuh, et al. Transforming research and clinical knowledge in traumatic brain injury pilot: multicenter implementation of the common data elements for traumatic brain injury. *Journal of neurotrauma*, 30(22):1831–1844, 2013.
12. Yongyue Zhang, Michael Brady, and Stephen Smith. Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm. *IEEE transactions on medical imaging*, 20(1):45–57, 2001.
13. Endre Grøvik, Darvin Yi, Michael Iv, Elizabeth Tong, Line Brennhaug Nilsen, Anna Latysheva, Cathrine Saxhaug, Kari Dolven Jacobsen, Åslaug Helland, Kyrre Eeg Emblem, et al. Handling missing mri sequences in deep learning segmentation of brain metastases: a multicenter study. *NPJ digital medicine*, 4(1):33, 2021.
14. Collins Achepsah Leke and Tshilidzi Marwala. *Deep learning and missing data in engineering systems*. Springer, 2019.
15. Hugh G Pemberton, Jiaming Wu, Ivar Kommers, Domenique MJ Müller, Yipeng Hu, Olivia Goodkin, Sjoerd B Vos, Sotirios Bisdas, Pierre A Robe, Hilko Ardon, et al. Multi-class glioma segmentation on real-world data with missing mri sequences: Comparison of three deep learning algorithms. *Scientific reports*, 13(1):18911, 2023.
16. Kenneth I Vaden Jr, Mulugeta Gebregziabher, Stefanie E Kuchinsky, and Mark A Eckert. Multiple imputation of missing fmri data in whole brain analysis. *Neuroimage*, 60(3):1843–1855, 2012.
17. Haoyu Lan, Alzheimer Disease Neuroimaging Initiative, Arthur W Toga, and Farshid Sepehrband. Sc-gan: 3d self-attention conditional gan with spectral normalization for multi-modal neuroimaging synthesis. *BioRxiv*, pages 2020–06, 2020.
18. Juan E Iglesias, Benjamin Billot, Yaël Balbastre, Colin Magdamo, Steven E Arnold, Sudeshna Das, Brian L Edlow, Daniel C Alexander, Polina Golland, and Bruce Fischl. Synthsr: A public ai tool to turn heterogeneous clinical brain scans into high-resolution t1-weighted images for 3d morphometry. *Science advances*, 9(5):eadd3607, 2023.

19. Yulin Wang, Honglin Xiong, Kaicong Sun, Shuwei Bai, Ling Dai, Zhongxiang Ding, Jiameng Liu, Qian Wang, Qian Liu, and Dinggang Shen. Towards general text-guided image synthesis for customized multimodal brain mri generation. *arXiv preprint arXiv:2409.16818*, 2024.
20. Tom Hilbert, Patrick Omoumi, Marcus Raudner, and Tobias Kober. Synthetic contrasts in musculoskeletal mri: a review. *Investigative Radiology*, 58(1):111–119, 2023.
21. Han Liu, Can Cui, Dario J Englot, and Benoit M Dawant. Uncertainty estimation in medical image localization: Towards robust anterior thalamus targeting for deep brain stimulation. In *Interpretable and Annotation-Efficient Learning for Medical Image Computing: Third International Workshop, iMIMIC 2020, Second International Workshop, MIL3ID 2020, and 5th International Workshop, LABELS 2020, Held in Conjunction with MICCAI 2020, Lima, Peru, October 4–8, 2020, Proceedings 3*, pages 130–137. Springer, 2020.
22. Alex Kendall and Yarin Gal. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems*, 30, 2017.
23. Chuan Guo, Geoff Pleiss, Yu Sun, and Kilian Q Weinberger. On calibration of modern neural networks. In *International conference on machine learning*, pages 1321–1330. PMLR, 2017.
24. Armen Der Kiureghian and Ove Ditlevsen. Aleatory or epistemic? does it matter? *Structural safety*, 31(2):105–112, 2009.
25. Jacob C Reinhold, Yufan He, Shizhong Han, Yunqiang Chen, Dashan Gao, Junghoon Lee, Jerry L Prince, and Aaron Carass. Validating uncertainty in medical image translation. In *2020 IEEE 17th International Symposium on Biomedical Imaging (ISBI)*, pages 95–98. IEEE, 2020.
26. Felix JS Bragman, Ryutaro Tanno, Zach Eaton-Rosen, Wenqi Li, David J Hawkes, Sebastien Ourselin, Daniel C Alexander, Jamie R McClelland, and M Jorge Cardoso. Uncertainty in multitask learning: joint representations for probabilistic mr-only radiotherapy planning. In *Medical Image Computing and Computer Assisted Intervention–MICCAI 2018: 21st International Conference, Granada, Spain, September 16-20, 2018, Proceedings, Part IV 11*, pages 3–11. Springer, 2018.
27. Quantile Regression. *Handbook of quantile regression*. CRC Press: Boca Raton, FL, USA, 2017.
28. Roger Koenker and Stephen Portnoy. *Quantile regression*. ABE, 1996.
29. Haleh Akrami, Anand A Joshi, Sergül Aydöre, and Richard M Leahy. Deep quantile regression for uncertainty estimation in unsupervised and supervised lesion detection. *The journal of machine learning for biomedical imaging*, 1, 2022.
30. KA Cauley, Y Hu, and SW Fielden. Pediatric head ct: automated quantitative analysis with quantile regression. *American Journal of Neuroradiology*, 42(2):382–388, 2021.
31. Bjoern H Menze, Andras Jakab, Stefan Bauer, Jayashree Kalpathy-Cramer, Keyvan Farahani, Justin Kirby, Yuliya Burren, Nicole Porz, Johannes Slotboom, Roland Wiest, et al. The multimodal brain tumor image segmentation benchmark (brats). *IEEE transactions on medical imaging*, 34(10):1993–2024, 2014.
32. Adam Paszke, Sam Gross, Francisco Massa, Adam Lerer, James Bradbury, Gregory Chanan, Trevor Killeen, Zeming Lin, Natalia Gimelshein, Luca Antiga, et al. Pytorch: An imperative style, high-performance deep learning library. *Advances in neural information processing systems*, 32, 2019.
33. M MONAI Consortium et al. Monai: Medical open network for ai. *Online at https://doi. org/10.5281/zenodo*, 5525502, 2020.

34. Olaf Ronneberger, Philipp Fischer, and Thomas Brox. U-net: Convolutional networks for biomedical image segmentation. In *Medical image computing and computer-assisted intervention–MICCAI 2015: 18th international conference, Munich, Germany, October 5-9, 2015, proceedings, part III 18*, pages 234–241. Springer, 2015.
35. Yael H Moshe, Yuval Buchsweiler, Mina Teicher, and Moran Artzi. Handling missing mri data in brain tumors classification tasks: Usage of synthetic images vs. duplicate images and empty images. *Journal of Magnetic Resonance Imaging*, 60(2):561–573, 2024.
36. Reza Azad, Nika Khosravi, Mohammad Dehghanmanshadi, Julien Cohen-Adad, and Dorit Merhof. Medical image segmentation on mri images with missing modalities: A review. *arXiv preprint arXiv:2203.06217*, 2022.