
RS-Net: Regression-Segmentation 3D CNN for Synthesis of Full Resolution Missing Brain MRI in the Presence of Tumour

Raghav Mehta
Centre for Intelligent Machines
McGill University
Montreal, QC
raghav@cim.mcgill.ca

Tal Arbel
Centre for Intelligent Machines
McGill University
Montreal, QC
arbel@cim.mcgill.ca

Abstract

Accurate synthesis of a full 3D MR image containing pathologies from other available MRI (e.g. to replace an image that is currently unavailable or corrupted) would provide a clinician, as well as downstream inference methods, with important complementary information for disease analysis. In this paper, we present an end-to-end 3D convolution neural network that takes a set of acquired MR image sequences (e.g. T1, T2, T1ce) as input and concurrently performs (1) regression of the missing full resolution 3D MRI (e.g. FLAIR) and (2) segmentation of the tumour into sub-types. Experiments on the BraTS 2017 [1] datasets show that: (1) replacing real MR volumes with the synthesized MRI does not lead to significant degradation in tumour and sub-structure segmentation accuracy and (2) performing synthesis and segmentation together allows the network to focus on attaining superior synthesis quality in the region of the tumour. The system further provides uncertainty estimates based on Monte Carlo (MC) dropout [9] for the synthesized volume at each voxel.

1 Introduction

Access to a variety of different Magnetic Resonance (MR) sequences (e.g. T1, T2, Fluid Attenuated Inverse Recovery (FLAIR)) improves medical image analysis results, in the context of neurological diseases for example, because different sequences provide complementary information [2]. However, in real clinical practice, not all MR image sequences are always available for each patient for a variety of reasons, including cost or time constraints or corruption due to noise or patient motion. As such, both clinical practice and automatic segmentation techniques would benefit greatly from the synthesis of one or more of the missing 3D MR image sequences based on the others provided [3]. Recently, modality synthesis has gained some attention from the medical image analysis community [4,5,6]. Several approaches, ranging from patch-based random forest [4] to 2D Convolutional Neural Network (CNN) [7], have been explored.

In this paper, an end-to-end 3D CNN is developed that takes as input a set of acquired MRI sequences of patients with tumours and simultaneously performs (1) regression to generate a full resolution missing 3D MR modality and (2) segmentation of the tumours into sub-classes. The hypothesis is that by performing regression and segmentation concurrently, the network should produce high quality 3D MR images, particularly the area of the tumour. The network is trained and tested on BraTS 2017 datasets. Most papers [4,6,7] on modality synthesis assess their synthesis results based on evaluation metrics such as Mean Squared Error (MSE) and Peak Signal-to-Noise Ratio (PSNR), which assess global synthesis quality. In this paper, we choose instead to evaluate synthesis quality at the tumour location by examining its performance on the subsequent independent downstream task of tumour segmentation. The proposed network also quantifies the uncertainty of the regressed synthetic

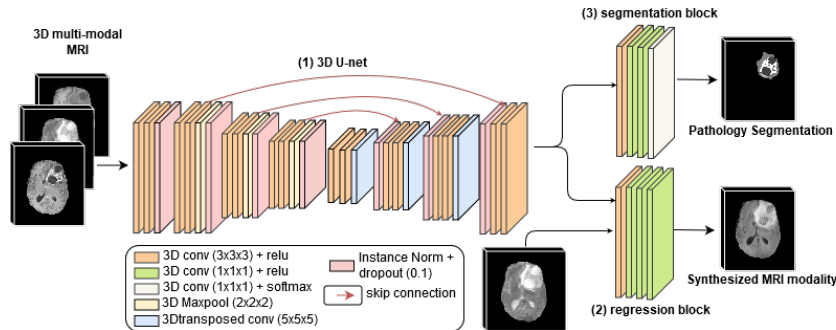


Figure 1: Proposed Regression-Segmentation CNN architecture (RS-Net): (1) A 3D U-net, (2) Regression and (3) Segmentation convolution blocks. The model takes as input several full 3D MR image sequences, synthesizes the missing 3D MRI, while concurrently generating the segmentation of the pathology, in this case a tumour.

volumes through Monte Carlo dropout [9]. This permits the confidence in the synthesis results to be conveyed to radiologists and clinicians and to automatic downstream methods that would use the synthesized volumes as inputs.

2 Regression-Segmentation CNN Architecture

A flowchart of the proposed Regression-Segmentation CNN architecture (**RS-Net**) can be seen in Figure 1. The network consists of three main components: (1) a modified 3D U-net [8], (2) regression convolution block for synthesizing the image sequence, and (3) segmentation convolution block for multi-class tumour segmentation. RS-Net takes as input full 3D volumes of all available sequences of a patient. The U-net generates an intermediate latent representation of the inputs which is provided to the regression and the segmentation convolution blocks. These then synthesize the missing 3D MR image sequences while also providing a segmentation of the tumour, both results produced at the same resolution. The U-net learns the latent representation common to both tumour segmentation and synthesis, focusing on high accuracy in the area containing tumour structures. In addition to the U-net output, the regression block is also provided with one of the input MRIs, which will provide the necessary brain MR contextual information to the regression block. The weighted mean Squared Error (MSE) loss is used for the synthesis task, and the weighted Categorical Cross Entropy (CCE) loss for segmentation. Here, the weights are defined such that the weight increases whenever there are fewer voxels in a particular tumour class.

3 Experiments and Results

The 2017 MICCAI BraTS [1] datasets were used for all the experiments. The BraTS training dataset was used to train the networks. This dataset is comprised of 210 High-Grade Glioma and 75 Low-Grade Glioma patients with T1, T1 post contrast (T1ce), T2, and FLAIR MRI for each patient, along with expert tumor labels for each of 3 classes: edema, necrotic/non-enhancing core, and enhancing tumor core. 228 volumes were randomly selected for training the network and another remaining 57 for network validation. A separate BraTS 2017 validation dataset, held out during training, was used to test the synthesis and segmentation performance. This dataset contains 46 patient multi-channel MRI (with no labels provided). We synthesize each four MR modalities in a leave one out approach where three modalities (e.g. T1c, FLAIR, T1) are used to synthesize the fourth image (e.g. T2).

Qualitative results (Figure 2) indicate that the network is able to produce high-quality, high-resolution, 3D synthesized MR volumes. The system also indicates locations where the network is uncertain about the regressed output, using Monte-Carlo dropout [9] based uncertainty estimates. Uncertainties can then be communicated to a clinician or radiologist to indicate trustworthy regions of the synthesized images, and automatic downstream methods using the synthesized volumes can focus computations on the areas of high confidence.

We train a separate segmentation CNN (**S-Net**) for the specific task of multi-class tumor segmentation based on 4 real MRI input volumes. This network is similar to RS-Net but modified such that the

	T1	T2	T1ce	FLR	Dice Enhance	Dice Tumour	Dice Core
Real	✓	✓	✓	✓	68.2	87.9	75.7
T1 Synthesis	⊙	✓	✓	✓	67.6	87.9	75.5
	•	✓	✓	✓	67.5	87.8	75.3
T2 Synthesis	✓	⊙	✓	✓	66.3	87.3	75.6
	✓	•	✓	✓	66.1	87.2	75.4
T1ce Synthesis	✓	✓	⊙	✓	24.8	87.3	54.0
	✓	✓	•	✓	24.1	85.9	53.9
FLAIR Synthesis	✓	✓	✓	⊙	66.8	83.6	73.1
	✓	✓	✓	•	62.9	81.3	71.5

Table 1: Multi-class brain tumor segmentation results on the BraTS17 Validation Dataset. Notation: Real MRI (✓), synthesized MRI RS-Net (⊙), and synthesized MRI R-Net (•). Quantitative segmentation results based on Dice coefficients for: enhancing tumor, whole tumor, and tumor core.

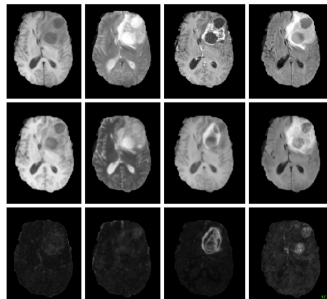


Figure 2: Example slice from synthetic MR volumes generated using the proposed RS-Net along with its associated uncertainties. Real MRI (Row 1); synthesized volumes (Row 2) and its associated uncertainties (Row 3) produced as mean and variance across 20 MC dropout samples. Columns from left to right: T1, T2, T1ce, and FLAIR. Notice that uncertainties are highest where predicted tumour enhancements in T1ce are incorrect.

synthesis convolution block is removed. To quantitatively evaluate the performance of the synthesized MRI, we replace the real MRI with synthesized one in S-Net. Note that we do not retrain the S-Net with the synthesized volume. This allows us to measure quality of the synthesized volumes as compared to the real volumes. To evaluate the effectiveness of the combined Regression-Segmentation task, we train a separate network (**R-Net**) for the task of synthesis alone. This network is similar to RS-Net but modified such that the segmentation block is removed as well as the additional input to the regression block. We compare the results of the competing methods.

In Table 1, we compare the tumour segmentation performance for real MR images vs synthesized MR images for all three types of Dice Coefficients (Enhancement, Tumour, and Core) on the BraTS 2017 Validation Dataset. The results indicate that by swapping out real MR volumes with the synthesized T1 or T2 MR volumes generated by the RS-Net leads to comparable segmentation performance on all three tumour sub-types. While for FLAIR synthesis, we observe small degradation in tumour segmentation performance for all three tumour sub-types. T1ce synthesis results in no loss of whole tumour segmentation performance but significant reduction in performance in terms of enhancement and necrotic core. This result was not surprising as T1ce images show enhancements that appear post-injection of a contrast agent, and therefore contain information not present in the other MR sequences. We can also observe that RS-Net performs better than R-Net, especially for FLAIR and T1ce, which indicates that performing synthesis and segmentation together allows the network to focus its attention on the tumour, and in turn gives better quality of the synthesized volume.

4 Conclusion

In this paper, a full resolution 3D end-to-end CNN was developed for the task of MR volume synthesis in the presence of brain tumours. The network was trained for the concurrent tasks of synthesizing a missing MRI sequence and tumour segmentation. Experiments on BraTS 2017 datasets indicated that multi-task learning helps in synthesizing high quality volumes over synthesis alone. Results also indicated that real MRIs can be replaced with synthesized volumes with minimum degradation in segmentation accuracy.

Acknowledgment

This work was supported by a Canadian Natural Science and Engineering Research Council (NSERC) Collaborative Research and Development Grant (CRDPJ 505357 - 16) and Synaptive Medical. We gratefully acknowledge the support of NVIDIA Corporation for the donation of the Titan X Pascal GPU used for this research.

References

- [1] Menze BH, Jakab A, Bauer S, Kalpathy-Cramer J, Farahani K, Kirby J, Burren Y, Porz N, Slotboom J, Wiest R, Lanczi L. "The multimodal brain tumor image segmentation benchmark (BRATS).", *IEEE Transactions on Medical Imaging (TMI)*. 2015 Oct;34(10):1993.
- [2] Havaei M, Guizard N, Chapados N, Bengio Y. "HeMIS: Hetero-modal image segmentation.", In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2016 Oct 17* (pp. 469-477). Springer, Cham.
- [3] van Tulder G, de Bruijne M. "Why does synthesized data improve multi-sequence classification?.", In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015 Oct 5* (pp. 531-538). Springer, Cham.
- [4] Jog A, Carass A, Roy S, Pham DL, Prince JL. "Random forest regression for magnetic resonance image synthesis.", *Medical image analysis*. 2017 Jan 31;35:475-88.
- [5] Roy S, Carass A, Prince J. "A compressed sensing approach for MR tissue contrast synthesis.", In *Biennial International Conference on Information Processing in Medical Imaging (IPMI) 2011 Jul 3* (pp. 371-383). Springer, Berlin, Heidelberg.
- [6] Van Nguyen H, Zhou K, Vemulapalli R. "Cross-domain synthesis of medical images using efficient location-sensitive deep network.", In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2015 Oct 5* (pp. 677-684). Springer, Cham.
- [7] Chatsias A, Joyce T, Giuffrida MV, Tsiftaris SA. "Multimodal mr synthesis via modality-invariant latent representation.", *IEEE Transactions on Medical Imaging (TMI)*. 2018 Mar;37(3):803-14.
- [8] Çiçek Ö, Abdulkadir A, Lienkamp SS, Brox T, Ronneberger O. "3D U-Net: learning dense volumetric segmentation from sparse annotation.", In *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI) 2016 Oct 17* (pp. 424-432). Springer, Cham.
- [9] Gal Y, and Ghahramani Z. "Dropout as a Bayesian approximation: Representing model uncertainty in deep learning." In *International Conference on Machine Learning (ICML)*, pp. 1050-1059. 2016.