3D brain MRI synthesis utilizing 2D SPADE-GAN and 3D CNN architecture

A. Ayaz[1], R. Jong[1], S. Abbasi-Sureshjani[1], S. Amirrajab[1], C. Lorenz[2], J. Weese[2], M. Breeuwer[1,3]

1. Eindhoven University of Technology, Biomedical Engineering Department, The Netherlands

2. Philips Research Laboratories, Hamburg, Germany

3. Philips Healthcare, MR R&D – Clinical Science, Best, The Netherlands

Summary of key findings

A pipeline has been set-up for synthesizing realistic 3D brain MRI images in a cost-efficient manner. A large set of anatomically variable synthesized 3D T1w brain MRI is generated with enhanced realism and reduced artifacts in all image directions.

Synopsis

We propose a method to synthesize 3D consistent brain MRI by using multiple 2D conditional SPatially-Adaptive (DE)normalization (SPADE) GANs to preserve the spatial information of patient-specific brain anatomy and a 3D CNN based network to improve the image consistency in all directions in a cost-efficient manner. Three individual 2D SPADE-GAN networks are trained across the axial, sagittal and coronal slice directions and their outputs are thereafter used further to train a CNN based 3D image restoration network to combine them into one 3D volume, using real MRI as a reference. The resulting predicted-synthesized 3D brain MRI is evaluated quantitatively and qualitatively.

Introduction

The medical research community is in need of large sets of annotated MRI data for training and validating deep learning based medical image analysis algorithms. Many of the brain analysis algorithms rely on the usage of full brain 3D anatomical data and a comprehensive set of anatomical annotations. Such data availability is scarce due to lack of proper ground truth annotations, limited data sharing policies and associated patient confidentiality. However, such datasets can be artificially generated by means of data-driven image synthesis. 3D synthesis networks[1,2] enable the robust 3D image synthesis, but the implementation of such networks is a complex and computationally expensive task. 2D image synthesis networks do not take into account the spatial information of multi-dimensional data and hence the images synthesized using such may be inconsistent in the through-plane direction. We propose to use a state-of-the-art 2D conditional SPatially-Adaptive (DE)normalization (SPADE) GAN model[3] as a base to synthesize brain MRI across multiple slice directions. We further propose a CNN based model to combine the 2D-synthesized MRI volumes to a consistent 3D MRI, having real MRI as a reference. Our synthesized 3D brain MRI came out highly realistic.

Methods

The pipeline designed for 3D consistent synthetic brain MRI is shown in **Error! Reference source not found.**. 3T high-resolution T1w MR scans of young healthy subjects from the HCP[4] database are used. The data is fed to the Philips proprietary fully automated complete brain classification tool, based on the fully

convolutional network architecture that process images at multiple scales and different fields of view[5]. The structural data along with the computed labels are used for training image synthesis networks. A matrix size of 256x320 having 256 slices across sagittal and an isotropic resolution of 0.7mm for image and label set is used for training all image synthesis networks.

The 2D SPADE GAN is trained using the Adam optimizer with learning rate of 0.0002, batch size of 12 on four (04) NVIDIA TITAN Xp GPUs. The rest of the training parameters, the architectures of the generator and discriminator and the losses are the same as in Park et al[3]. Three separate 2D SPADE GANs are trained while slicing across axial, sagittal, and coronal direction respectively. At inference, multi slice 2D brain MRI volumes are synthesized for a set of labels that are not seen by the network during training. The training and inference data is kept same for all three image synthesis networks.

The output synthetic volumes are further used as 3 channel input to train a 3D CNN shown in Figure 2, while the associated real MRI volumes are used as target output volumes. A patch wise approach is used to train the network using patches of 32x32x32 voxels and an overlapping of 50% in all directions. The output 3D patches are merged to form a 3D MRI volume. Mean Square Error (MSE) is used as a loss function. The network is trained using the Adam optimizer with learning rate of 0.001, batch size of 50 on one NVIDIA GeForce 2080 GPU.

The final predicted volumes are evaluated qualitatively and quantitatively by comparison with the real 3D MRI volumes. Intensity profiles and image quality metrics of Structural Similarity Index (SSIM), MSE and Peak Signal to Noise Ratio (PSNR) are computed. A visual experiment is made showing real and our synthesized 3D MRI volumes as random slices in all directions. Readers are asked to rate the image quality on a scale of 1-5 varying from Very Good (as 1) to Very Poor (as 5). The responses are recorded for the visual quality assessment.

Results and Discussions

Single subject middle slices across all axis of generated and real samples are shown in Figure 3. Synthetic 2D MRI samples synthesized across axial, coronal, and sagittal plane are only close to the real MRI volume in the direction the network is trained on respectively but are blurry and missing detailed brain features along with some unwanted artifacts in the other, non-trained directions. However, our CNN-predicted 3D brain MRI and real brain MRI are extremely similar to each other across all directions. It is worth mentioning that these images are obtained for unseen data for both GAN and CNN networks at inference time.

Image quality metrics of our final synthesized-predicted MRI volumes when compared to the real for brain only structures are listed in Figure 4a. The metrics show low MSE and high SSIM and PSNR. The intensity profiles drawn in real and predicted MRI volume across all planes, as shown in Figure 4b, also show a close match. Visual image quality ratings from 2 readers are presented as a frequency chart in Figure 4c, showing overall high and almost same scorings for both real and predicted MRI volumes.

Conclusion

Using our proposed synthesis pipeline, highly realistic appearing 3D brain MRI volumes are generated without increasing the computational cost associated with 3D synthesis networks. CNN based combination of synthesized samples across different slice directions, while having real MRI as a reference, removes the slice incoherency and artifacts for images synthesized across individual 2D SPADE-GAN networks.

Acknowledgements

References

1. Kwon, Gihyun, et al. "Generation of 3D brain MRI using auto-encoding generative adversarial networks." *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Springer, Cham, 2019.
2. Lan, Haoyu, et al. "SC-GAN: 3D self-attention conditional GAN with spectral normalization for multi-modal neuroimaging synthesis." *bioRxiv*, 2020.
3. Park, Taesung, et al. "Semantic image synthesis with spatially-adaptive normalization." *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 2019.
4. Van Essen, David C., et al. "The WU-Minn human connectome project: an overview." *Neuroimage*, 2013.
5. Brosch, Tom, and Axel Saalbach. "Foveal fully convolutional nets for multi-organ segmentation." *Medical Imaging*, 2018.
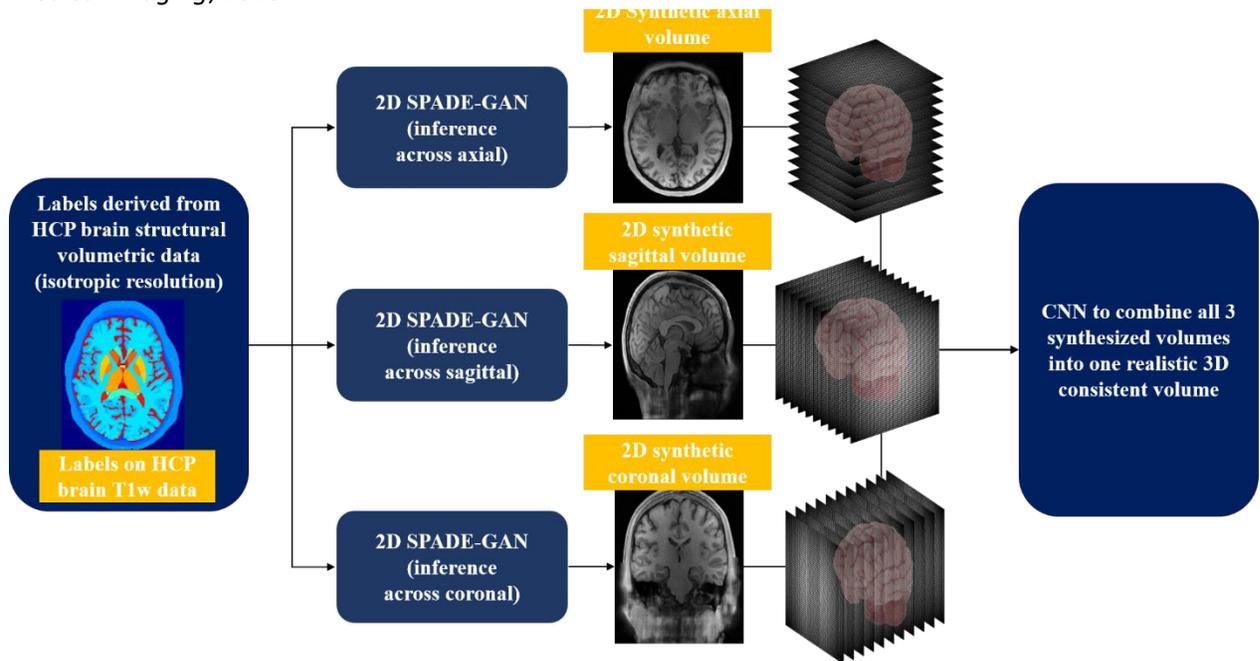
Figure 1: The pipeline designed for 3D consistent synthetic brain MR images, making use of isotropic patient-specific phantoms, multiple 2D SPADE-GANs and a 3D CNN.
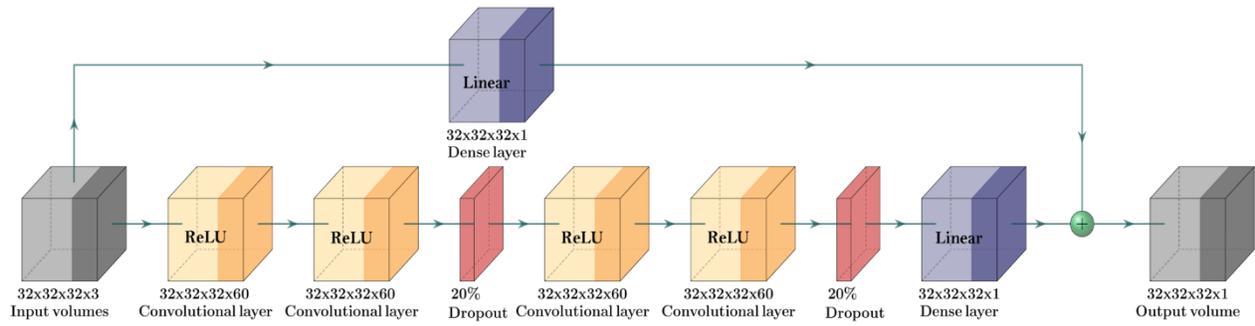
Figure 2: Our proposed 3D CNN architecture built on top of the multi input synthesized volumes with a paired real MRI volume as an output. The 3 channel inputs are 3D patches (of size 32x32x32) from axial, sagittal and coronal synthesized volumes from respective 2D SPADE-GANs.
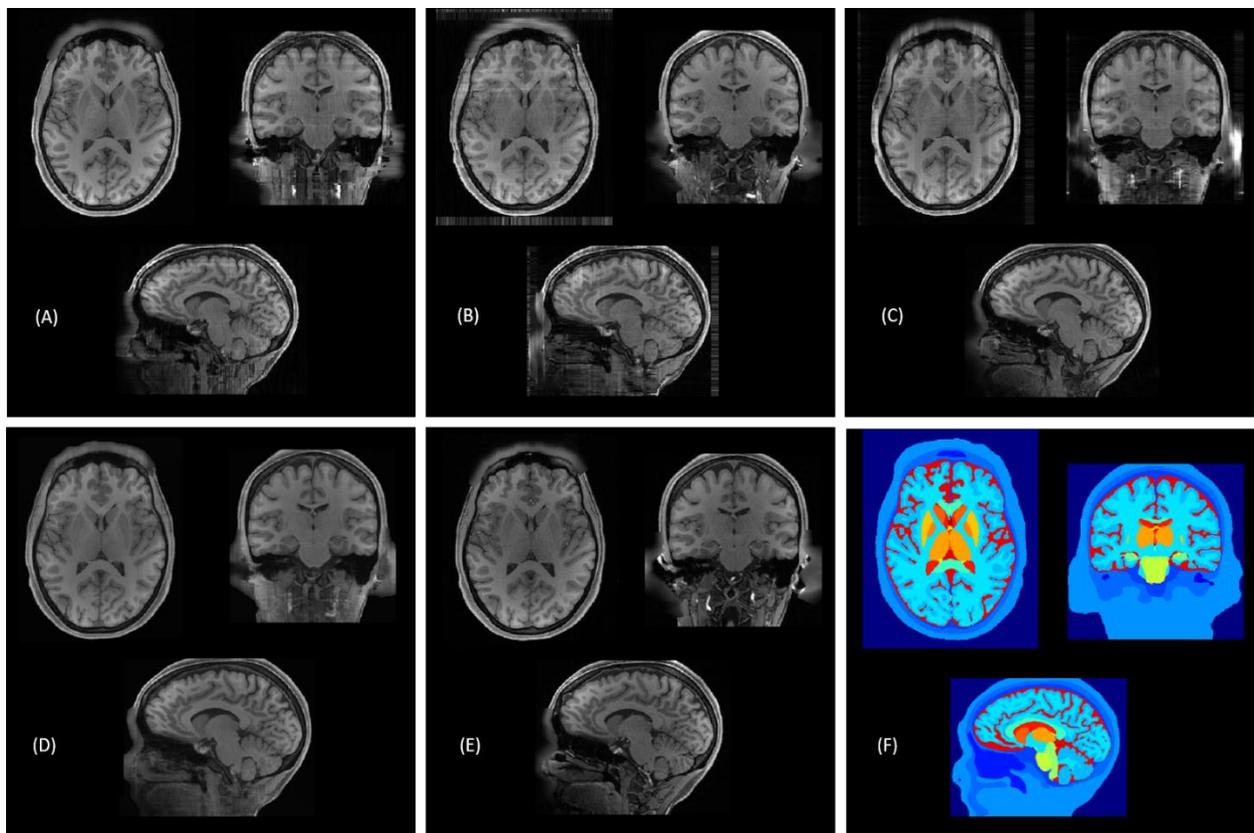

Figure 3: Real and synthesized samples of normal brain T1w MRI across middle axial, sagittal and coronal slices.(A) Synthetic 3D MRI samples synthesized across axial, (B) coronal, and (C) sagittal plane. (D) Our CNN predicted 3D consistent MRI volume. (E) Paired real MRI sample. (F) Ground truth labels used for synthesis.

(A) Image Quality Metrics

(B) Intensity Profiles

SSIM	MSE	PSNR
0.948 ± 0.006	0.002 ± 0.001	27.38 ± 0.001

Values as Mean ± Std
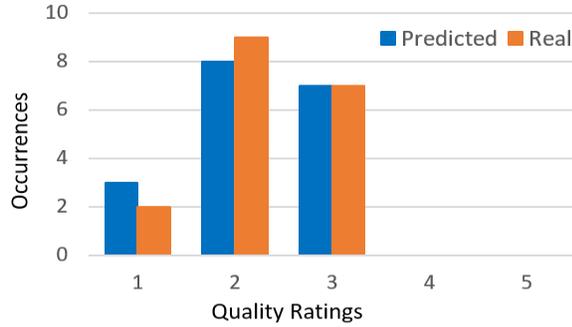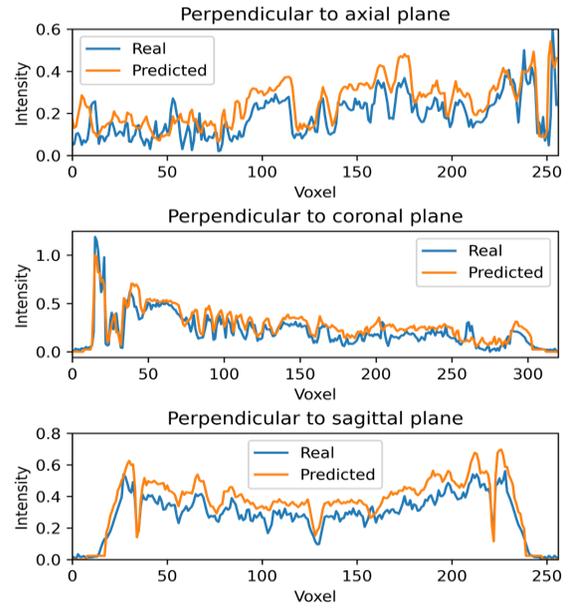
(C) Reader Scores

Figure 4: (A) Image quality metrics of SSIM, PSNR and MSE of predicted MRI volumes with real MRI as a ground truth as mean and standard deviation. (B) Intensity profiles drawn in real and predicted volume of one sample subject across all three brain MRI planes. (C) Reader scores for image quality ratings for some sample real and predicted axial, coronal and sagittal slices presented to user. The quality ratings are on a scale of 1-5 varying from Very Good (as 1) to Very Poor (as 5).