Synthetic Segmented Virtual Head Model Generation Using Generative Adversarial Network (GAN)

Nahian Ibn Hasan*

[a]*Elmore School of Electrical and Computer Engineering, Purdue University, West Lafayette, IN, USA*

Abstract

The structural information of the human brain assists in different research such as Transcranial Magnetic Stimulation (TMS), Transcranial Direct Current Stimulation (tDCS), brain tumor detection, uncertainty quantification analysis, etc. However, MRI images of less than 3T resolution are of little use since the tissue boundaries become uncertain in the brain. On the other hand, acquiring a substantial number of higher resolution images (3T, 5T, 7T) is significantly expensive. Hence, an expanded database is often necessary for population-based studies, which is often quite impossible to gather from a single machine, environment, or race of people. Therefore, we aim to develop a workflow to generate synthetic segmented MRI images, which will help in the population-based analysis and machine learning algorithms. Since generating "3D synthetic human head models" is computationally expensive, we aim to work with a single slice of MRI scans. The virtual head models (for ground-truth MRI scans), generated from SimNIBS/Fresurfer, are voxelated to a uniform grid space. Next, we utilize the Generative Adversarial Network (StyleGAN) to produce synthetic segmented MRI slices using over 800 MRI data from Human Connectome Project. The Frechet Inception Distance (FID) score between 5000 generated images and real dataset resembles realistic synthetic data.

Keywords: Synthetic, Generative Adversarial Network, MRI, Virtual Head Model, Segmentation

1. Introduction

The availability of structural and diffusion information of the human brain through MRI scans has led researchers to studies like Transcranial Magnetic Stimulation (TMS), Transcranial Direct Current Stimulation (tDCS), Uncertainty quantification for structural and conductivity analysis, neural connection analysis in different tissues, brain tumor segmentation, neuroscience studies, treatment of psychological disorders or depression. Most researches related to human subjects, especially brain structural/ diffusion information, require MRI data with a powerful enough MRI coil system such as 3T/ 5T/ 7T. Also, the diversity among the different populations from varying ethnic or environmental backgrounds makes the population-based or group-level analysis harder than it seems. Conclusion on any statistical inference for any population-based study requires the assurance of high enough diversity in the underlying database. That is why we propose an algorithm to generate synthetic data that will ensure that diversity in the population-based studies.

The first step of theoretical modeling and analysis of any study related to structural information of the human brain is to generate the virtual head model of the physical brain. Researchers and engineers have developed

*Corresponding author.
Email address: nahianhasan1994@gmail.com (Nahian Ibn Hasan)

automated tools for generating virtual head models over the years. SimNIBs[1–3] is one of the surface-based cortical segmentation tools that utilize Freesurfer[4–13] and FSL[14] to generate segmented tissue models of the human brain. The group-level analysis requires the MRI coregistration under the same platform, which the SimNIBS does automatically. There are other surface and volume-based cortical segmentation tools such as BrainSuite[15], CP-CHARM[16], SPM[17] etc. One of the main caveats of these ATLAS-based segmentation tools is the required time for segmenting a single head model. For example, SimNIBS takes about 6-24 hours on average for segmenting a single MRI scan. Hence, various deep learning models have been developed, to facilitate automatic and fast segmentation. One of the prime leaders is the FastSurfer[18] in this context. FastSurfer uses the segmented heads from FreeSurfer as ground truth data to train a deep neural network.

One requires a considerable amount of data before reaching any conclusion on group-level analysis or machine learning algorithms. There are two approaches for data enhancement - assembling more experimental data and synthetic data generation. The first one is a longer approach that requires a considerable amount of resources. The second one is a more practical approach with less accuracy. Synthetic data generation or, otherwise known as "data augmentation" is a common approach for increasing samples based on the feature distribution of actual limited data. Some of the basic techniques used for data augmentation for machine learning algorithms are morphological image transformations, image color space transformations and deep learning-based data augmentation. One of the prime candidates for deep learning-based synthetic data generation is the Generative Adversarial Networks (GANs)[19, 20]. Generative networks differ from image manipulation techniques by generating newer data rather than manipulating existing samples. After the invention of GANs in 2014, it has been developed further to design deep convolutional GAN (DCGAN)[19], conditional GAN (cGAN)[21], Information Maximizing GAN (InfoGAN)[22], Auxiliary Classifier GAN (AC-GAN)[23], StackGAN [24], Pix2Pix[25], Wasserstein GAN(WGAN)[26], Cycle-Consistent GAN (CycleGAN)[27], Progressive Growing GANProgressive GAN[28], BigGAN[29], Style-based GAN(StyleGAN)[20]. In this paper, we focus on StyleGAN to generate synthetic data.

In the paper, we use SimNIBS/ FreeSurfer for generating virtual head models from physically collected MRI scans. Next, the segmented models are voxelized in a uniform grid size across subjects. After that, the StyleGAN is trained with these segmented and voxelated data to generate synthetic and segmented head models. For analysis purposes, we focus on a single cortical slice. The workflow is easily extendable for producing 3D head models. Finally, a feasibility study confirms the resemblance between synthetic and ground-truth MRI.

2. Proposed Method

This section represents the overview of the database and the proposed methodology for synthetic data generation.

For our study, the structural MRI scans have been collected from [30]. The structural MRI scans are of either 3T or 7T MRI machines. There are 812 head models corresponding to 765 subjects. Some of the subjects have MRI MRI sessions. The data were collected by maintaining the Human Connectome Project (HCP) data usage terms and policies. All the subjects are young adults aged between 22 and 35 years.

Fig. 1 briefly illustrates the workflow of the proposed algorithm. First, all of the ground-truth MRI scans from the database are passed down to the SimNIBS/ FreeSurfer pipeline. The MRI scans are coregistered and a finite element mesh (FEM) is generated for each virtual head model. The FEM nodes contain the tissue

conductivity information of the brain. Next, all the FEM mesh are voxelized to a uniform $256 \times 256 \times 256$ grid.Each element of the grid is any of the five tissue classes, assigned by SimNIBS/Freesurfer. After that, the StyleGAN is trained with these ground-truth voxelized data to generate new synthetic data. The synthetic data are passed through a series of image noise reduction techniques to form a realistic head model. Finally, a feasibility study on the synthetic data confirms the robustness of the algorithm.

Table 1 briefly describes the signal processing technique used for the project. However, not all of the steps of the FreeSurfer pipeline have been listed in the table.

Table 1: Image Processing Steps Involved in Project

Description of the signal processing step	**Rationale**	**Inputs and Outputs of the step**
Image Co-registration - Geometrical transformation of MRI scans based on four fiducial points scattered over the cortical surface. These four fiducial points are Nz (Nasion), Iz (Inion), LPA (Left Preauricular point) and RPA (Right Preauricular point). Image coregistration facilitates group level analysis by transforming them them to a single geometrical orientation. The EEG co-ordinate system used for this co-registration is tderived from Cutini et al.[] and projected on the scalp of each subjects scalp surface. There are 76 electrodes are chosen from MNI template describeed by Jurak et al.[].	Facilitates Group-level analysis	Input = Raw MR data Output = Coregistered MRI
Motion Correction - Some of the subjects in the database have multiple MRI sessions and multiple runs. When there are multiple volumes, this technique averages over the volumes.	It removes the artifacts due to small movement of the head during scan. However, this is only applied for MRI data with multiple scan in the same session.	Input=Co-registered MRI, Output=Averaged MRI Volume
Intensity Correction - Non-parametric non-uniformity correction in MR data. It's originally implemented by MINC in the tool nu-correct'	Corrects the intensity distribution in the MRI data.	input=non-uniform intensity MR image, Output=Uniform intensity MR data
Affine transformation - Talairach Co-ordinate System	Transforms individula MR volume to the Talirach system to find the location of brain structure independent of subjects.	Input=Intensity uniformed MR data, Output=Transformed MR data to Talairach system
Intensity Normalization - Scales the intensity of each voxel of MRI scan such that the intensity of WM is 110. Performs intensity-based tissue segmentation. Hence, scaling the intensity for each tissue helps segmentation.	Normaliza intensity across the MR volume	Input=talairach registered head, Output=Normalized intensity MR image

Description of the signal processing step	Rationale	Inputs and Outputs of the step
Skull Strip - "Watershed" segmentation tool is applied to find the intensity of WM, GM and CSF. Next, a force field is applied to fit a spherical surface to the outer boundary of the brain tissues.	Segment out skull and other non-brain tissues	Input=Intensity normalized MR, Output=Segmented skull and brain tissues.
Tissue Segmentation - Intensity based tissue segmentation by utilizing multimodal gaussian distribution.	Segment and labeling 3 different tissue categories - WM, GM, CSF	Input = Brain tissue, Output=Segmented tissue voxels
Image voxelization - Voxelize image to a uniform grid size. Each voxel correspond to one of 5 different tissues.	uniform data structure construction for neural networks	Input=FEM mesh generated from SimNIBS, Output=Voxelized Image
Generative Adversarial Network (GAN) - synthetic image generation from StayleGAN architecture.	generate synthetic images	Input=Voxelized MR data, Output=Synthetic MR data
Image Quantization - labeling uniformly for a certain category of tissue. This method labels 5 unique conductivity values based on the histogram thresholding.	Tissue label uniforming	Input=Synthetic MR image generated by StyleGAN, Output=Quantized image.
Image Erosion & Dilation - Kernel based image convolution.	remove small noises and fills small holes in the images.	Input-Quantized image, Output=processed synthetic data.
FID Score - Calculates the Frechet Inception score. Takes the weights of the final fully-connected layer of the Inception-V3 neural network for both synthetic and ground-truth MR data, calculates the statistical distance between the two categories of wight.	evaluate the quality of synthetic images	Input=1000 Processed synthetic data, Output=Frechet score,
Tissue area - calculates the area of tissue categories with respect to the number of pixels for that category.	find out tissue area for evaluation	Input=Processed synthetic data, Output=Tissue area for each category .

3. Results and Discussions

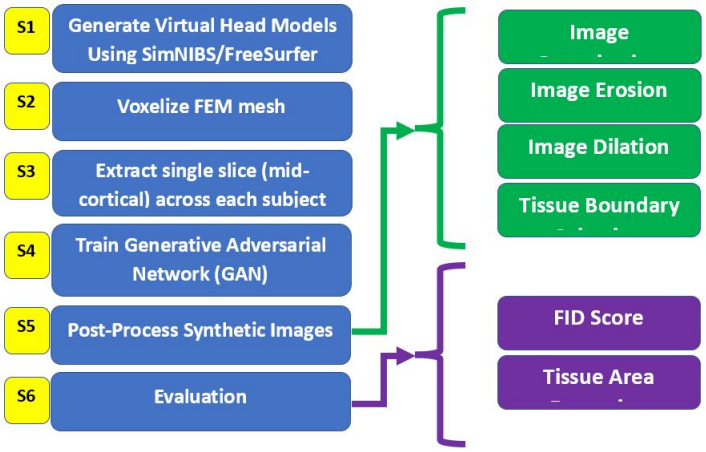


Figure 1: Synthetic data generation process. In step 1, segmented head models are generated with different tissue types [Atlas-based segmentation using tools like Freesurfer/ SimNIBS]. In step 2, voxelized head models of the segmented head tissues are produced. In step 3, a single slice (i.e. Middle coronal slice of the head models across different subjects) is extracted. Next, a Generative Adversarial Network (StyleGAN) is trained to generate synthetic slices. In the next step, the synthetic images are processed to reduce noise & artifacts and generate uniform tissue segments. This step involves different image processing techniques (i.e., image erosion, image dilation, and quantization). In the final step, a feasibility study shows the synthetic data quality. This step involves comparing pixel distribution across different tissues and similarity checks through Frechet Inception Distance (FID).

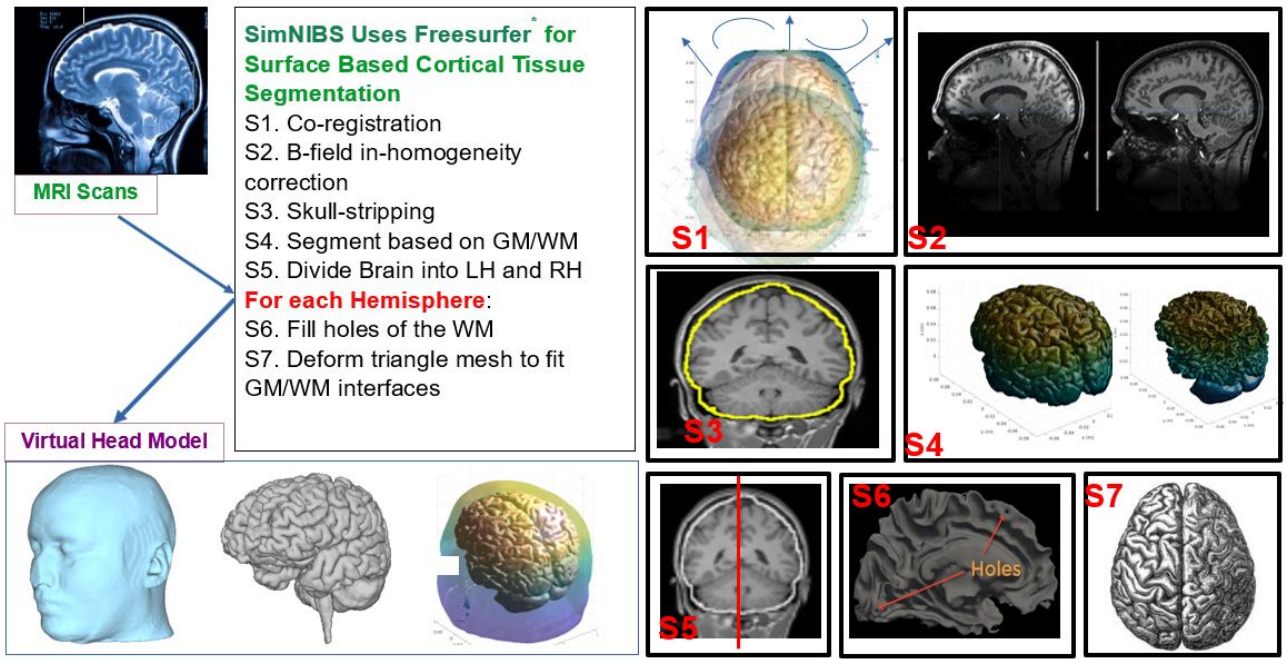


Figure 2: Virtual head model generation from ground-truth MRI data using SimNIBS/FreeSurfer pipeline. The MRI scans are passed down to the SimNIBS/ FreeSurfer pipeline, and the output is the co-registered FEM mesh. Several major steps are mentioned here, that is used to segment and mesh the whole brain. But there are other sub-processes. It's a huge workflow implementing different image processing steps. The toolbox co-registers the heads in a common frame (S1) facilitating group-level analysis, and the head models have high fidelity. Then it uses bias field correction to make the intensity uniform (S2). After that, it extracts the brain tissues from the skull (S3) and segments out different types of tissues such as GM, WM, CSF (S4). Next, it divides the brain into two hemispheres-LH and RH (S5). And at the end, it tries to fix some imperfections like holes and ambiguities in tissue boundaries (S6). Finally, it deforms the triangle mesh to fit the GM/WM interfaces and forms the pial surface (S7).

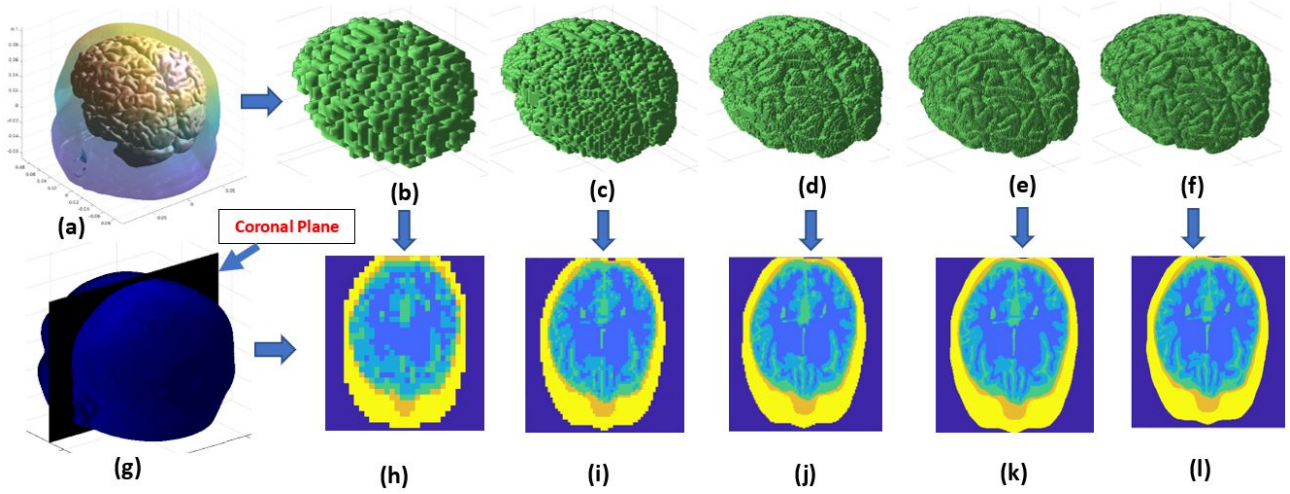


Figure 3: Image voxelization and 2D pixelated data generation. (a) Finite element mesh (FEM) generated from SimNIBS/FreeSurfer Pipeline. (b-f) represents uniform voxelized grid data with grid sizes of $32 \times 32 \times 32, 64 \times 64 \times 64, 128 \times 128 \times 128, 256 \times 256 \times 256$ and $512 \times 512 \times 512$, respectively. (g) The mid coronal plane is selected from the voxelized data. (h-l) represents the mid coronal slice for different voxelization grids. For lower grid sizes, the tissue boundaries are not well defined. The tissue boundaries become prominent with grid sizes beyond 128. So, the higher the grid size, the better the tissue boundary resolution. But for higher grid sizes, the computational effect comes into play. As you can see, the best quality of the voxelized head model is achieved by 512 grid sizes. But the data size becomes enormous. Training a 3D neural net with large data like this becomes impossible. For training purposes, we have selected voxelized heads of grid size 256.

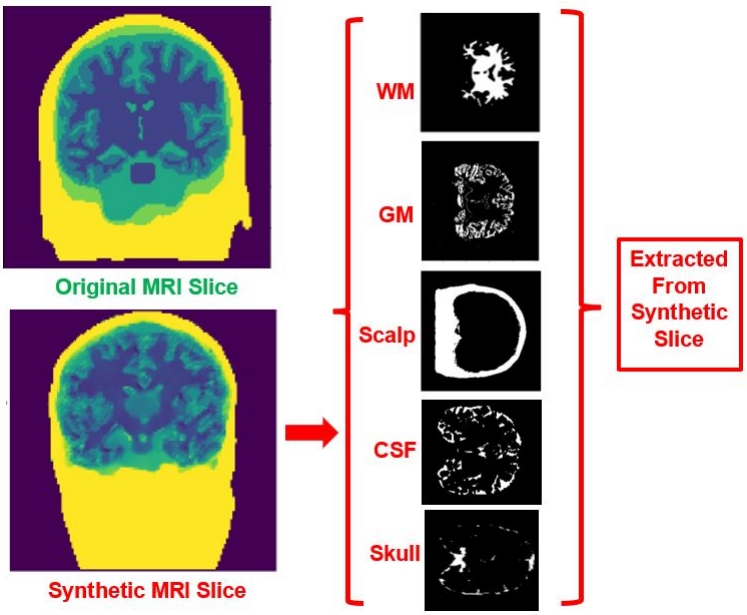


Figure 4: Synthetic image generation from StyleGAN. On the righ side, the extracted tissue regions are shown.

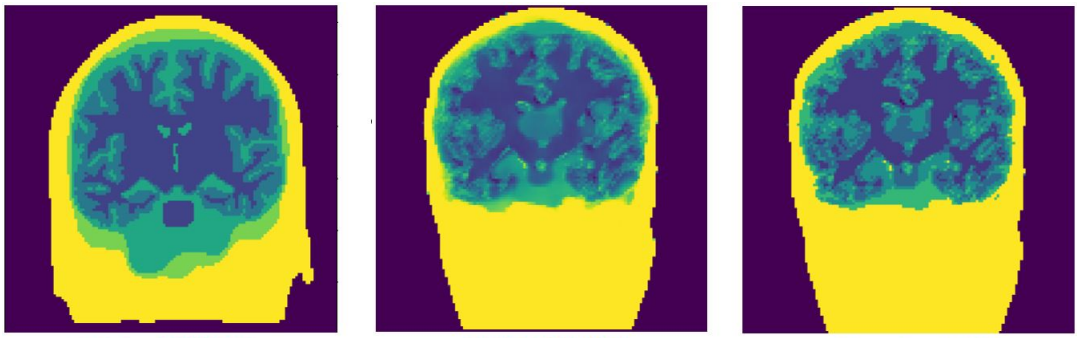
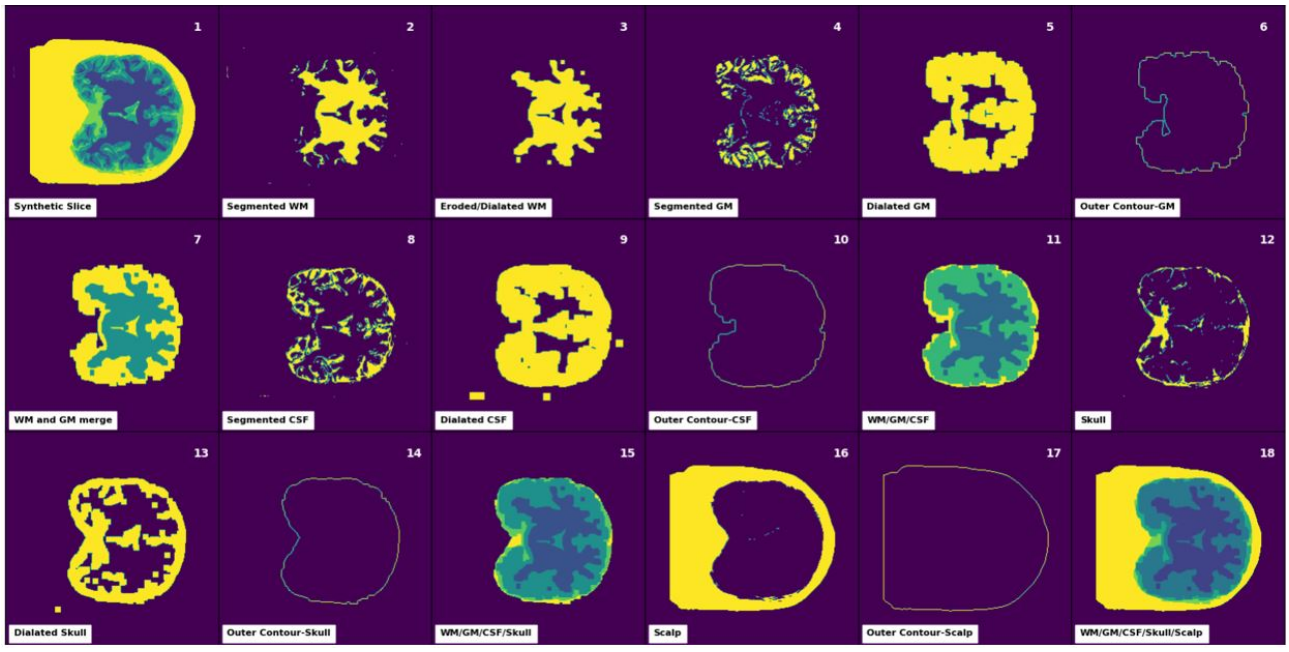
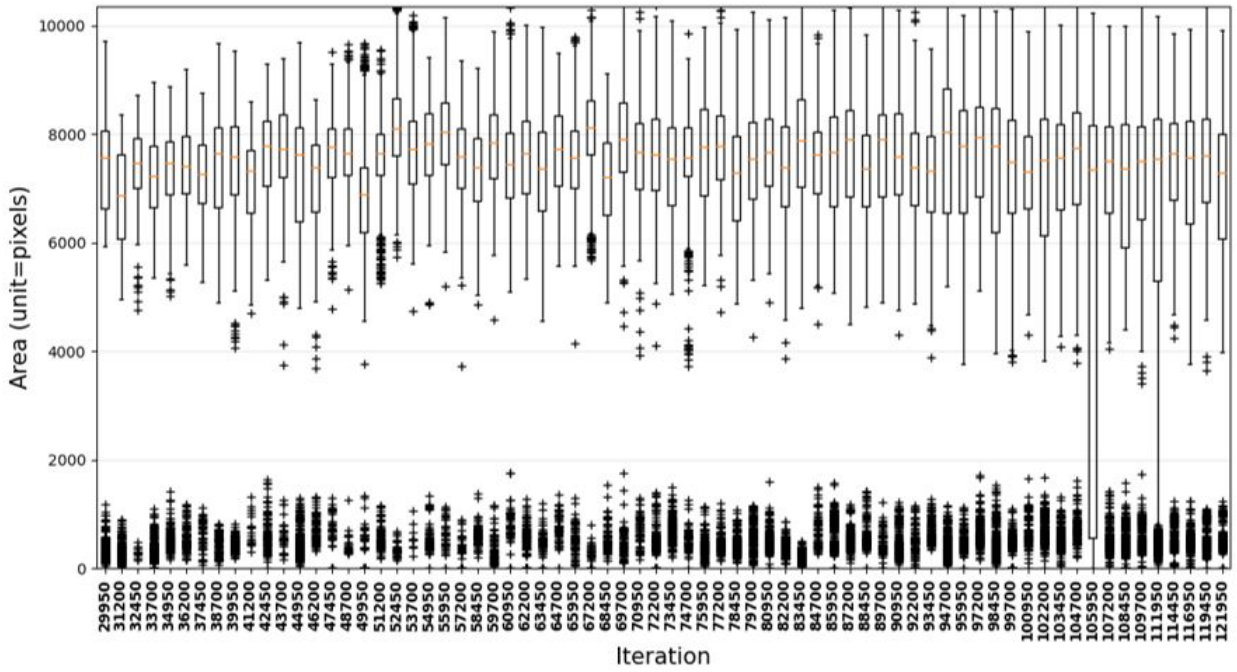


Figure 5: The first step of the post-process is image quantization. GAN generates images with distinct tissue numbers with multi-modal Gaussian distribution. In our case, we have 5 modalities corresponding to 5 different tissues. Hence, we need to label each pixel corresponding to each tissue with some unique labels. The quantized tissues are much more prominent than the immediate generated GAN image. However, the tissue boundaries are not well defined yet.

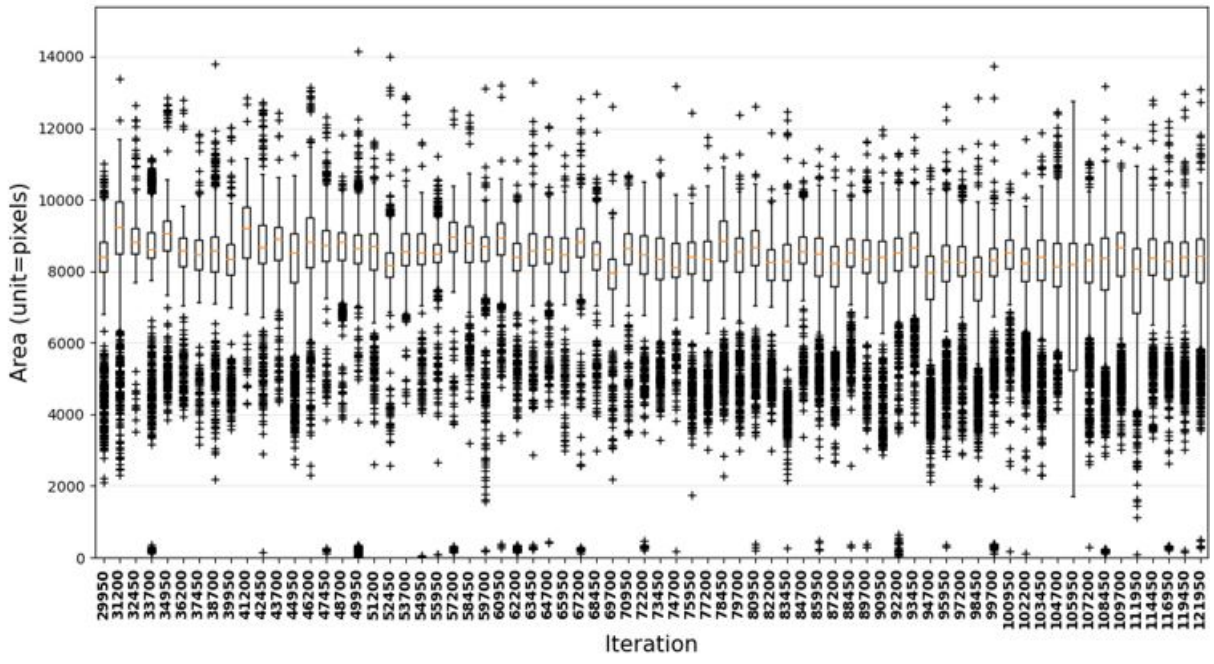


(a)

Figure 6: Summary of post-processing steps applied on synthetic images. Step 1 represents the quantized image. WM is extracted at step 2. Next, mage erosion and dilation remove most of the small island-like noises and reconstruct the tissue boundary at step 3. Next, GM is extracted at step 4 and dilated multiple times at step 5. Next GM outer boundary is extracted at step 6. At step 7, both the GM and WM are superimposed. However, the tissue boundary extracted at step 6 defines the outer boundary in the superimposition step. Next, a similar process is repeated for CSF from step 8 to 10 and superimposed with GM and WM in step 11. A similar process is applied on the skull from steps 12 to 15 and on the scalp from steps 16-18. Step 18 represents the processed synthetic image.



(a)



(b)

Figure 7: Sample tissue area distribution (WM in (a)) and GM in (b)) over 5000 synthetic images at any iteration. The generative network produces an almost uniform distribution of tissue area at different training steps. The StyleGAN starts to generate meaningful images after about 28000^{th} iteration. The network produces an almost equal amount of tissue pixels in different sample images for different tissue categories. That means, in the training process, the network first decides how many pixels should be assigned for each tissue, and after that, it starts to change the spatial distribution of the tissues. That's why the tissue area is not a good candidate for evaluating the performance of Generative networks.

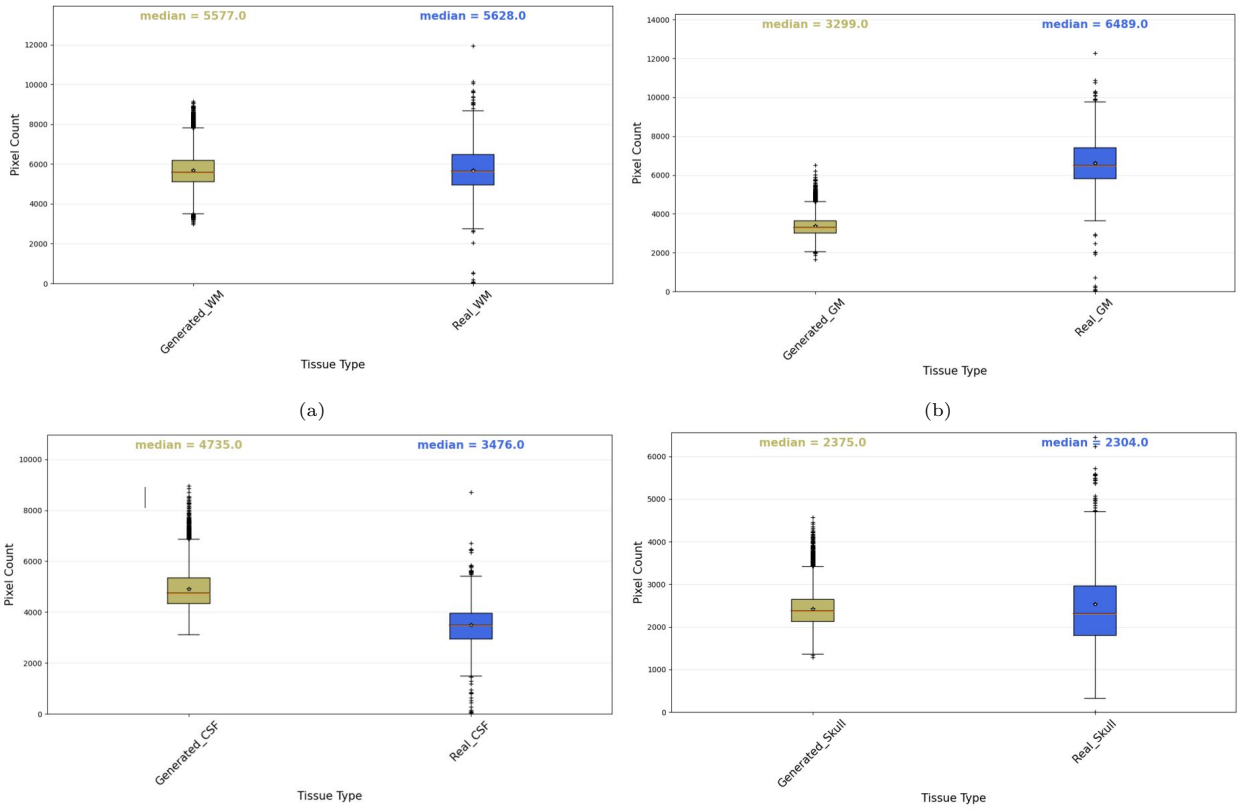


Figure 8: Tissue area (Number of tissue pixels × pixel area) distribution between 5000 synthetic images and real MR data for WM (a), GM (b), CSF (c), and Skull (d). The tissue area of synthetic images is almost equal and well within the 20-80 percentile of the actual distribution.

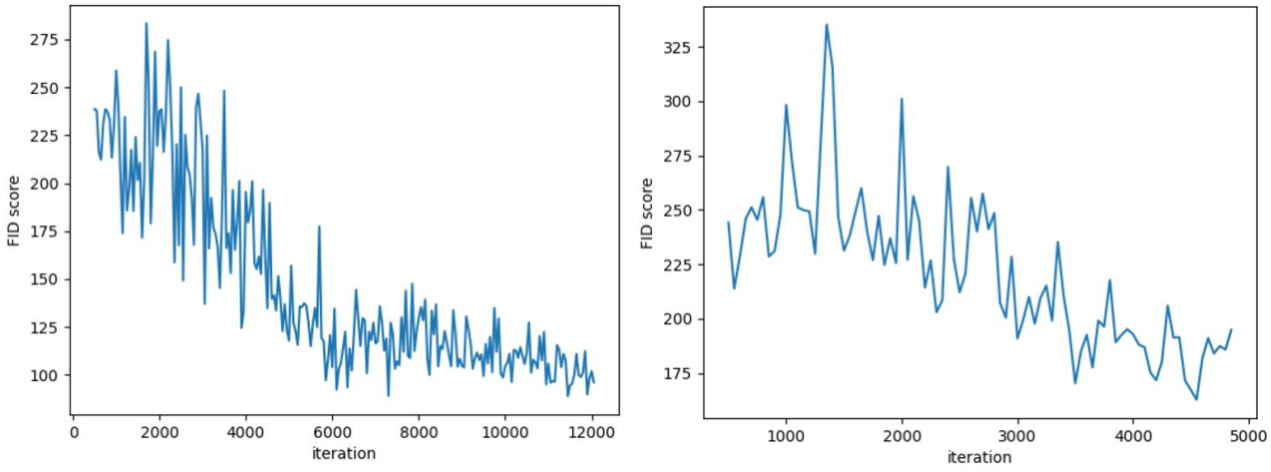


Figure 9: Frechet Inception Distance (FID) score for two different voxelization grid - $128 \times 128 \times 128$ (a) and $256 \times 256 \times 256$ (b). The lower the FID score, the better the image quality and the higher the similarity between synthetic data and ground-truth data. For large enough iterations, this score can be brought down to an expected range.

4. Conclusion

This paper focuses on generating synthetic data to be utilized for population-based studies. The MRI data are processed to generate virtual head models. Next, the FEM mesh from SimNIBS/ FreeSurfer pipeline are transformed into a uniform voxel grid size. The voxelization grid size affects the image quality, especially tissue boundary. Next, a single coronal plane is extracted from voxelized grid. StyleGAN utilizes these planar data to generate synthetic images and it require a considerable number of iterations before generating any meaningful image that resembles a ground truth MRI image. In the feasibility study, it appeared that the area of any tissue does not predict the quality of synthetic image at any training iteration. But the FID score improves with respect to training iterations. Therefore, the FID score can be used as a parameter for statistical inference. However, a manual inspection of the synthetic image always provides the actual synthetic image quality.

5. Acknowledgment

Declarations of interest: none. This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

References

[1] M. Windhoff, A. Opitz, A. Thielscher, Electric field calculations in brain stimulation based on finite elements: an optimized processing pipeline for the generation and usage of accurate individual head models, Tech. rep., Wiley Online Library (2013).

[2] J. D. Nielsen, K. H. Madsen, O. Puonti, H. R. Siebner, C. Bauer, C. G. Madsen, G. B. Saturnino, A. Thielscher, Automatic skull segmentation from mr images for realistic volume conductor models of the head: Assessment of the state-of-the-art, Neuroimage 174 (2018) 587–598.

[3] G. B. Saturnino, O. Puonti, J. D. Nielsen, D. Antonenko, K. H. Madsen, A. Thielscher, Simnibs 2.1: a comprehensive pipeline for individualized electric field modelling for transcranial brain stimulation, Brain and human body modeling (2019) 3–25.

[4] A. M. Dale, B. Fischl, M. I. Sereno, Cortical surface-based analysis: I. segmentation and surface reconstruction, Neuroimage 9 (2) (1999) 179–194.

[5] A. M. Dale, M. I. Sereno, Improved localizadon of cortical activity by combining eeg and meg with mri cortical surface reconstruction: a linear approach, Journal of cognitive neuroscience 5 (2) (1993) 162–176.

[6] R. S. Desikan, F. Ségonne, B. Fischl, B. T. Quinn, B. C. Dickerson, D. Blacker, R. L. Buckner, A. M. Dale, R. P. Maguire, B. T. Hyman, et al., An automated labeling system for subdividing the human cerebral cortex on mri scans into gyral based regions of interest, Neuroimage 31 (3) (2006) 968–980.

[7] B. Fischl, A. M. Dale, Measuring the thickness of the human cerebral cortex from magnetic resonance images, Proceedings of the National Academy of Sciences 97 (20) (2000) 11050–11055.

[8] B. Fischl, A. Liu, A. M. Dale, Automated manifold surgery: constructing geometrically accurate and topologically correct models of the human cerebral cortex, IEEE transactions on medical imaging 20 (1) (2001) 70–80.

[9] B. Fischl, D. H. Salat, E. Busa, M. Albert, M. Dieterich, C. Haselgrove, A. Van Der Kouwe, R. Killiany, D. Kennedy, S. Klaveness, et al., Whole brain segmentation: automated labeling of neuroanatomical structures in the human brain, Neuron 33 (3) (2002) 341–355.

[10] B. Fischl, D. H. Salat, A. J. Van Der Kouwe, N. Makris, F. Ségonne, B. T. Quinn, A. M. Dale, Sequence-independent segmentation of magnetic resonance images, Neuroimage 23 (2004) S69–S84.

[11] B. Fischl, M. I. Sereno, A. M. Dale, Cortical surface-based analysis: Ii: inflation, flattening, and a surface-based coordinate system, Neuroimage 9 (2) (1999) 195–207.

[12] M. Reuter, B. Fischl, Avoiding asymmetry-induced bias in longitudinal image processing, Neuroimage 57 (1) (2011) 19–21.

[13] H. Rosas, A. Liu, S. Hersch, M. Glessner, R. Ferrante, D. Salat, A. van Der Kouwe, B. Jenkins, A. Dale, B. Fischl, Regional and progressive thinning of the cortical ribbon in huntington's disease, Neurology 58 (5) (2002) 695–701.

[14] S. M. Smith, M. Jenkinson, M. W. Woolrich, C. F. Beckmann, T. E. Behrens, H. Johansen-Berg, P. R. Bannister, M. De Luca, I. Drobnjak, D. E. Flitney, et al., Advances in functional and structural mr image analysis and implementation as fsl, Neuroimage 23 (2004) S208–S219.

[15] D. W. Shattuck, R. M. Leahy, Brainsuite: an automated cortical surface identification tool, Medical image analysis 6 (2) (2002) 129–142.

[16] V. Uhlmann, S. Singh, A. E. Carpenter, Cp-charm: segmentation-free image classification made accessible, BMC bioinformatics 17 (1) (2016) 1–12.

[17] J. Ashburner, K. J. Friston, Voxel-based morphometry—the methods, Neuroimage 11 (6) (2000) 805–821.

[18] L. Henschel, S. Conjeti, S. Estrada, K. Diers, B. Fischl, M. Reuter, Fastsurfer-a fast and accurate deep learning based neuroimaging pipeline, NeuroImage 219 (2020) 117012.

[19] A. Radford, L. Metz, S. Chintala, Unsupervised representation learning with deep convolutional generative adversarial networks, arXiv preprint arXiv:1511.06434 (2015).

[20] T. Karras, S. Laine, T. Aila, A style-based generator architecture for generative adversarial networks, in: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, 2019, pp. 4401–4410.

[21] M. Mirza, S. Osindero, Conditional generative adversarial nets, arXiv preprint arXiv:1411.1784 (2014).

[22] X. Chen, Y. Duan, R. Houthooft, J. Schulman, I. Sutskever, P. Abbeel, Infogan: Interpretable representation learning by information maximizing generative adversarial nets, in: Proceedings of the 30th International Conference on Neural Information Processing Systems, 2016, pp. 2180–2188.

[23] A. Odena, C. Olah, J. Shlens, Conditional image synthesis with auxiliary classifier gans, in: International conference on machine learning, PMLR, 2017, pp. 2642–2651.

[24] H. Zhang, T. Xu, H. Li, S. Zhang, X. Wang, X. Huang, D. N. Metaxas, Stackgan: Text to photo-realistic image synthesis with stacked generative adversarial networks, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 5907–5915.

[25] P. Isola, J.-Y. Zhu, T. Zhou, A. A. Efros, Image-to-image translation with conditional adversarial networks, in: Proceedings of the IEEE conference on computer vision and pattern recognition, 2017, pp. 1125–1134.

[26] J. Adler, S. Lunz, Banach wasserstein gan, arXiv preprint arXiv:1806.06621 (2018).

[27] J.-Y. Zhu, T. Park, P. Isola, A. A. Efros, Unpaired image-to-image translation using cycle-consistent adversarial networks, in: Proceedings of the IEEE international conference on computer vision, 2017, pp. 2223–2232.

[28] T. Karras, T. Aila, S. Laine, J. Lehtinen, Progressive growing of gans for improved quality, stability, and variation, arXiv preprint arXiv:1710.10196 (2017).

[29] A. Brock, J. Donahue, K. Simonyan, Large scale gan training for high fidelity natural image synthesis, arXiv preprint arXiv:1809.11096 (2018).

[30] D. C. Van Essen, S. M. Smith, D. M. Barch, T. E. Behrens, E. Yacoub, K. Ugurbil, W.-M. H. Consortium, et al., The wu-minn human connectome project: an overview, Neuroimage 80 (2013) 62–79.