High-Resolution MRI Brain Inpainting

Mohamed Almansour
Alexandria University
Alexandria, Egypt
mismail@alexu.edu.eg

Nagia M. Ghanem
Alexandria University
Alexandria, Egypt
nagia.ghanem@alexu.edu.eg

Soheir Bassiouny
Alexandria University
Alexandria, Egypt
saf@alexu.edu.eg

2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI) | 978-1-6654-0358-0/21/$31.00 ©2021 IEEE | DOI: 10.1109/BHI50953.2021.9508561

***Abstract*—Analyzing diseased brain images is often challenging as the presence of abnormal brain tissues hinders several image processing tasks, like tissue segmentation and non-rigid registration, leading to distorted or biased results. This paper presents a solution to this problem by replacing abnormal brain tissues with healthy ones using free-form image inpainting to restore spatial consistency. Inspired by the recent success of Gated Convolution, we propose a user-guided deep adversarial inpainting model to fill irregularly shaped holes in high-resolution T1-weighted MR brain images. Besides, we propose a novel linear-time algorithm to generate irregular masks on the fly for training images. Comparative qualitative and quantitative analyses demonstrated that the proposed model outperforms recently proposed deep learning methods for brain lesion filling and significantly reduces blurriness and boundary artifacts. Moreover, the proposed network shows promising performance in inpainting brain tumors and improving tissue segmentation outputs.**

***Index Terms*—Brain Image Inpainting, Gated Convolution, Human Connectome Project, Magnetic Resonance Imaging.**

I. INTRODUCTION

Brain image analysis theory and application development have recently featured enormous growth, which certainly supported dependent medical applications and neuroscience research. Reliable software packages, such as FreeSurfer [1], Advanced Normalization Tools (ANTs) [2], and the Oxford Center for Functional Magnetic Resonance Imaging of the Brain (FMRIB) Software Library (FSL) [3], provided accurate and robust solutions for important preprocessing and post-processing tasks in multiple modalities (e.g., Bias field correction, linear and nonlinear registration, tissue segmentation, and network modeling). However, most of these tools were designed for spatially consistent healthy brains or treat pathological brains as exceptional cases of their healthy counterparts. As pathological brain analysis has critical clinical and research values, brain image inpainting has been previously suggested as a preprocessing step to achieve spatial consistency by replacing damaged tissues with healthy ones. This approach showed promising performance in different tasks, such as tissue segmentation [4], non-rigid registration [5], and tumor segmentation [6].

Image inpainting is an important computer vision task that aims to fill missing pixels of an incomplete image with the most visually consistent and indistinguishable values. Besides brain imaging, image inpainting has a wide range of applications, including Object Removal [7] and Image Restoration [8]. As pointed out in [9], image inpainting methods could be classified into two main categories: structural inpainting (which relies on geometric techniques) and texture inpainting (which uses texture synthesis). Classical methods provided solutions by structural inpainting, texture inpainting, or a combination of both to estimate the missing pixels depending on low-level features and patch similarity [10]–[14]. These classical methods are plausible for simple textures but less accurate for complex, finely detailed, and inhomogeneous textures, especially when filling large irregular holes. Several tools, including ANTs and FSL, provide classical brain lesion filling modules, which produce plausible results for small lesions but critically fail in handling large lesions.

On the other hand, deep learning approaches moved towards synthesizing high-level features over the entire background (available pixels) and propagating them to the foreground (missing pixels) to improve the inpainting accuracy and quality. Deep image inpainting substantially progressed in the last decade after Convolutional Neural Networks (CNNs) achieved state-of-the-art performance in several image understanding tasks, including Image Classification [15], Object Detection [16], and Semantic Segmentation [17]. Numerous networks were proposed inspired by Context Encoders [18], which was the first model based on Generative Adversarial Networks (GANs) designed for the image inpainting task. For example, [19] made a remarkable improvement to the network design and performance by using a dilated convolutional bottleneck and building local and global discriminators for patch and global consistencies. More recently, the Context Attention Layer [20] was proposed to coarse-to-fine models, which reported better refinement results and successfully replaced the Poisson blending post-processing in [19]. Other layers were also proposed to substitute for vanilla convolutional layers to fit irregular masks better. These layers upgraded vanilla convolution by designing methods of discrimination between the foreground and the background pixels for the convolution operation. The most prominent examples are the Partial Convolutional Layer [21] and the Gated Convolutional Layer [22].

In this work, we present a deep adversarial network for user-guided free-form inpainting of high-resolution magnetic resonance (MR) brain images. The approach we present is simple yet efficient in dealing with several challenges, including data cleansing, training optimization, and model generalization. The contributions of our work can be summarized as follows. First, we present the first gated convolutional network for brain

978-1-6654-0358-0/21/$31.00 ©2021 IEEE
Authorized licensed use limited to: University Bern. Downloaded on May 16,2024 at 07:50:27 UTC from IEEE Xplore. Restrictions apply.

image inpainting. Second, the proposed network is trained using 1000 T1-weighted (T1-w) MR brain volumes, which is the largest training set used in deep brain image inpainting to the best of our knowledge. Third, we propose a novel iterative algorithm to generate irregular masks on the fly in linear time. Finally, we achieve the best qualitative and quantitative results compared to previous similar methods.

The rest of the paper is structured as follows. Section II describes the dataset and the methods used in building the model. Section III includes the network training configuration and the evaluation results. Finally, Section IV concludes the paper and discusses the potential future work based on the model limitations.

II. Dataset and Methods

A. Data Description

The training dataset was downloaded from the Human Connectome Project 1200 Subjects Release (S1200) [23], which contains 1113 subjects scanned on Siemens® 3 Tesla Connectome Skyra in Washington University. Each subject has a T1-w structural MR volume of dimensions 256 × 320 × 320 and 0.7 millimeters (mm) isotropic voxel size. After applying the same preprocessing steps to all subjects, they were randomly split into 1000 for training and 113 for testing.

B. Preprocessing

The initial preprocessing steps are performed using ANTs, including resampling to 1 mm isotropic voxel size (output dimensions: 179 × 224 × 224), denoising using the spatially adaptive filter proposed in [24], bias field correction using the algorithm proposed in [25], intensity normalization to range 0–1, and brain extraction. Afterward, cropping and minor resizing are performed to transform the brains to 180 × 180 × 180 volumes, preserving the same aspect ratios. Finally, rotation, cropping, flipping, noising, and brightness scaling are randomly used as augmentation techniques to create the training set. Fig. 1 demonstrates the preprocessing and augmentation steps.

C. Irregular Mask Generator

When it comes to model generalization, generating a new mask for each training image is much better than using a predefined mask set, provided that it does not hinder the training efficiency. Accordingly, we propose an iterative geometric method (Algorithm 1) to generate a randomly shaped mask in linear time and memory complexities.

The algorithm draws a mask for a corresponding image of shape $h \times w$ by deforming a circle of initial center (x, y) and initial radius r. The deformation rate of the circle is controlled by the drift parameters $C_{x,y,r}$, whereas the interface $random\_sign$ returns 0, 1, or -1 with equal probabilities. Fig. 3 shows the effect of changing the drift parameters on a centrally located shape of $r = 30$.

There are multiple advantages of this method, including linear time complexity $O(hw)$, tunability, and transferability to 3D space to create irregular volumes. Reproducibility could be also achieved by replicating $h, w, x, y, r, C_{x,y,z}$, and the random sign generator seed. Finally, the exact mask area could be assigned or calculated by counting the number of pixels set to 1 (line 7).

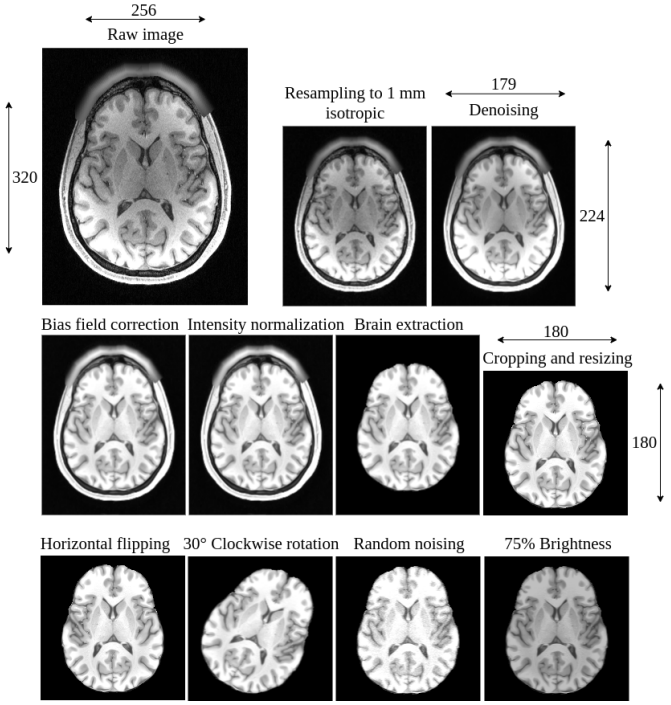

Figure 1. Visual illustration of preprocessing and augmentation steps with scale ratios preserved.

Algorithm 1 Generate Irregular Mask

1: **procedure** GENERATE_MASK($h, w, x, y, r, C_x, C_y, C_r$)
2: $mask = array\_of\_zeros(shape = (h, w))$
3: **for** every row i **do**
4: **for** every column j **do**
5: $d = \sqrt{(i-x)^2 + (j-y)^2}$
6: **if** $d \leq r$ **then**
7: $mask[i][j] = 1$
8: x += $random\_sign() * C_x$
9: y += $random\_sign() * C_y$
10: r += $random\_sign() * C_r$
11: **return** $mask$

Training masks were created by combining a potentially overlapping random number of blobs (m) of random initial centers. The initial radius of each blob was estimated by

$$r = \sqrt{\frac{hw}{\lambda \pi m}} \tag{1}$$

where λ is the ratio of the image area to the mask area.

D. Network and Loss Functions

The generator network is similar to the coarse-to-fine model proposed in [22]. However, the refinement attention branch is omitted for optimization as its effect was found negligible

Authorized licensed use limited to: University Bern. Downloaded on May 16,2024 at 07:50:27 UTC from IEEE Xplore. Restrictions apply.

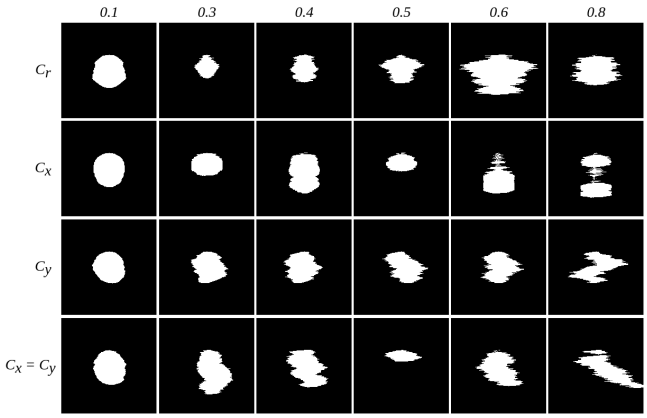

Figure 3. The results of changing the values (top) of every drift parameter (left) while all the other drift parameters equal zero. In all shapes $r = 30$. In the last row $C_x = C_y$ (= top value) and $C_r = 0$.

in our experiments. The coarse and refine units are both encoder-decoder fully gated convolutional networks of dilated bottlenecks, where gated convolution is given by

$$y(x) = \sigma\left(\sum\sum W_{feature} \cdot x\right) \odot \phi\left(\sum\sum W_{gating} \cdot x\right) \quad (2)$$

Equation (2) defines the gated convolutional layer output (y) as a function of the input image/feature (x), where $W_{feature}$ is the feature convolutional filter, W_{gating} is the gating convolutional filter, σ is the sigmoid function, ϕ is an arbitrary activation function, and $\odot$ is element-wise product.

The generator network (G) has three input channels (input image, foreground mask, and user guidance sketch) and one output channel (refine output). Fig. 2 shows the generator architecture details. Furthermore, the generator pixel-wise reconstruction loss $(\mathcal{L}_{Rec})$ is given by

$$\mathcal{L}_{Rec} = \mathbb{E}_{x \sim P_{data}(x)}\left[\left\|G(x \odot (1-m)) - x\right\|_1 \odot m\right] \quad (3)$$

where x is an arbitrary data point, m is the foreground pixels mask, and $\|\cdot\|_1$ is ℓ_1-norm (Manhattan norm).

For adversarial training, we use the Spectral-Normalized Markovian Discriminator proposed in [22], which computes adversarial and discriminator hinge losses for each point on its output map. The adversarial hinge loss $(\mathcal{L}_{Adv})$ is given by

$$\mathcal{L}_{Adv} = -\mathbb{E}_{z \sim P(z)}\left[D(G(z))\right] \quad (4)$$

where D is the discriminator network and z is an arbitrary incomplete input image. And the discriminator hinge loss $(\mathcal{L}_D)$ is given by

$$\begin{aligned}\mathcal{L}_D = \mathbb{E}_{x \sim P_{data}(x)}[ReLU(1 - D(x))] + \\ \mathbb{E}_{z \sim P_z(z)}[ReLU(1 + D(G(z)))]\end{aligned} \quad (5)$$

where $ReLU$ is the rectified linear activation function. Accordingly, the generator total loss $(\mathcal{L}_D)$ is given by

$$\mathcal{L}_G = \lambda_{Rec}\mathcal{L}_{Rec} + \lambda_{Adv}\mathcal{L}_{Adv} \quad (6)$$

where λ_{Rec} and λ_{Adv} are the balancing hyperparameters, which were both set to 1 during training.

E. User Guidance Sketch Generation

User guidance sketches are expected to be less accurate and dense than the corresponding ground truth edges because of the irregularity of the cortical gyrification and the inpainting masks. To address this issue, we use Canny Edge Detector [26] to produce sketches for training images on the fly, with the hysteresis minimum and maximum thresholds selected at random. Edge maps are then smoothed by a 5×5 Gaussian kernel to decrease the sharpness of the results. Fig. 4 shows the effect of changing the hysteresis thresholds on the edges densities and the subsequently developed sketches by Gaussian smoothing.

III. Training and Evaluation

A. Training Hyperparameters

The model was trained for 1000 iterations using Adam optimizer [27], where $\beta_1 = 0.9$, $\beta_2 = 0.999$, the generator learning rate $\eta_G = 0.0001$, the discriminator learning rate $\eta_D = 0.001$, and the batch size $= 5$. These hyperparameters were estimated by performing a grid search procedure. Furthermore, the batch size was increased to 10 after the first 500 iterations as proposed in [28].

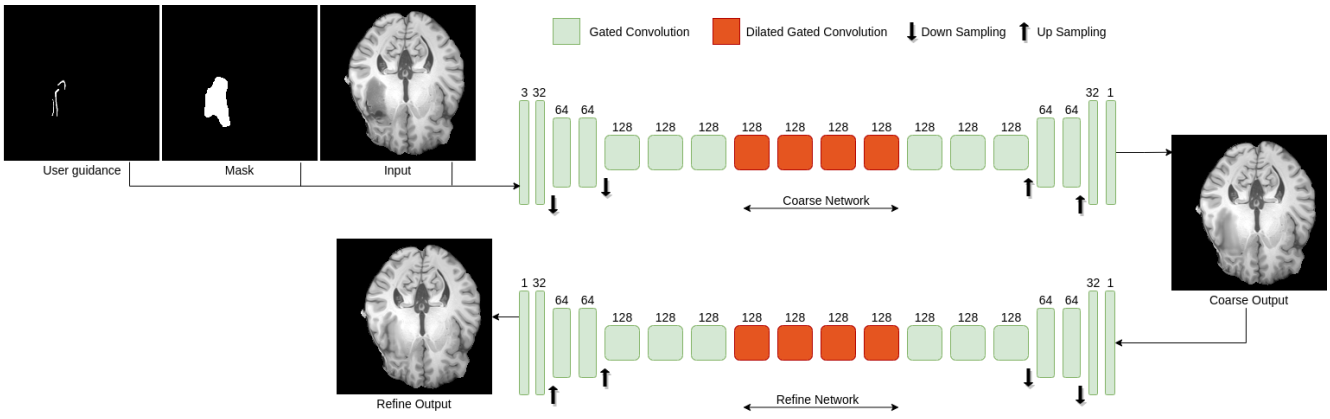

Figure 2. The generator network (G) architecture and the number of channels per layer.

Authorized licensed use limited to: University Bern. Downloaded on May 16,2024 at 07:50:27 UTC from IEEE Xplore. Restrictions apply.

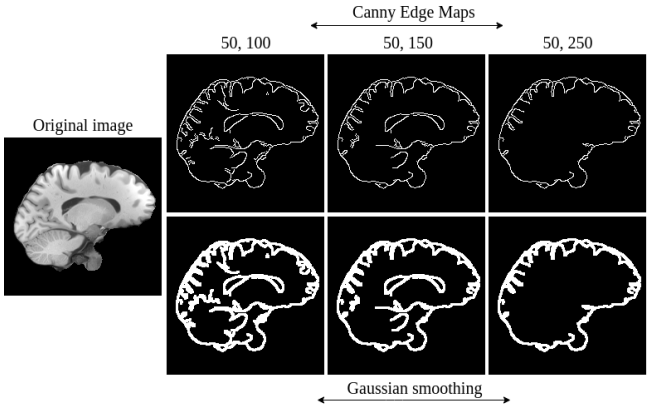

Figure 4. Canny edge maps of different hysteresis thresholds (top values) of a sagittal slice. 5×5 Gaussian smoothing is displayed below each edge map.

B. Masks and User Guidance Sketches

For each training image we generated a new mask and a new user guidance sketch on the fly. Masks were generated using the method in Section II-C, where $m \sim U(1, 20)$, $\lambda \sim U(3, 10)$, $C_x = 0$, $C_y = 0$, and $C_r \sim U(0.1, 0.5)$. User guidance sketches were generated using the method in Section II-E, where the hysteresis minimum and maximum thresholds were randomly sampled from $U(50, 80)$ and $U(100, 400)$, respectively. Fig. 5 shows the inpainting results of different user sketches.

C. Comparative Evaluation

We compared our model with three previous methods: the blind inpainting network proposed in [4], the basic inpainting algorithm proposed in [5], and the context-aware network proposed in [29]; which, for simplicity, were abbreviated to BIN, BIA and CAN, respectively. The quantitative comparison was made using mean absolute error (MAE), mean square error (MSE), euclidean-normalized root mean square error (NRMSE), peak signal-to-noise ratio (PSNR), and the structural similarity index measure (SSIM) proposed in [30]. Testing images were obtained by slicing the 113 testing brain volumes in the directions of the sagittal, coronal, and axial planes, ignoring vacant images, while testing masks were generated using the method in Section II-C. Evaluation metrics were then measured between the model output and the corresponding ground truth. Table I shows the average values of the evaluation metrics and Fig. 8 displays examples of the qualitative comparative evaluation using the testing image set.

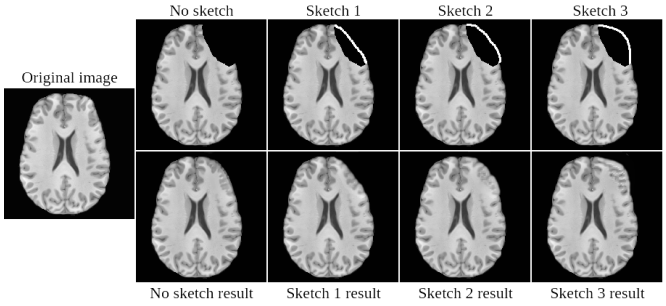

Figure 5. The inpainting results of different user sketches and the corresponding ground truth.

Table I
QUANTITATIVE COMPARATIVE EVALUATION.

Method	MAE	MSE	NRMSE (%)	PSNR (decibels)	SSIM
BIN	0.014	0.0006	17.2	31.6	0.961
BIA	0.02	0.0008	21.1	28.56	0.942
CAN	0.005	0.0004	13.4	34.6	0.976
Ours	**0.0024**	**0.00017**	**8.7**	**44.1**	**0.987**

D. Tumor Inpainting and Tissue Segmentation Refinement

The model was further tested using glioma and meningioma samples obtained from [31], which underwent the ANTs preprocessing steps mentioned in Section II-B. Inpainting masks were manually drawn along the lesion boundaries and the surrounding perilesional edema using FSL. As shown in Fig. 9, the model sustainably filled tumors of different sizes and positions. However, restoring the subject's pretumoral brain anatomy is not always possible as the bulk of the tumor and the peritumoral edema exert mass effect on the adjacent healthy brain tissue with subsequent distortion of its normal anatomy, which depends on the site of tumor origin. This condition is illustrated in Fig. 6. Moreover, tumor inpainting substantially alleviated the adverse effect caused by the tumor and edema on the tissue segmentation quality of FMRIB's Automated Segmentation Tool (FAST) [32]; Fig. 7 displays the results.

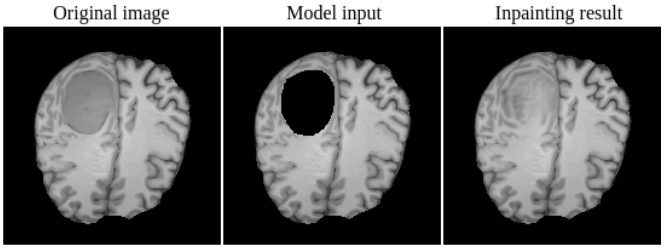

Figure 6. Meningioma mass effect and the corresponding inpainting attempt.

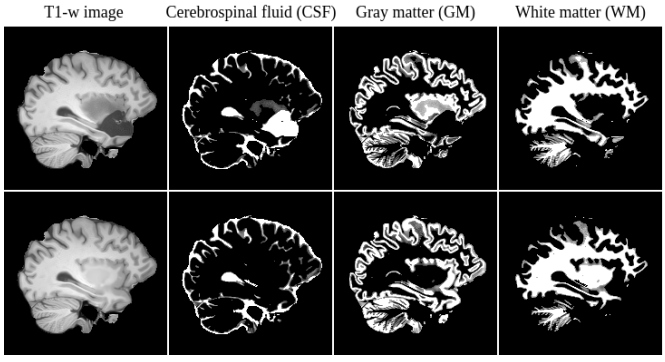

Figure 7. FAST partial tissue segmentation of a glioma patient before (top) and after (bottom) lesion inpainting.

Authorized licensed use limited to: University Bern. Downloaded on May 16,2024 at 07:50:27 UTC from IEEE Xplore. Restrictions apply.

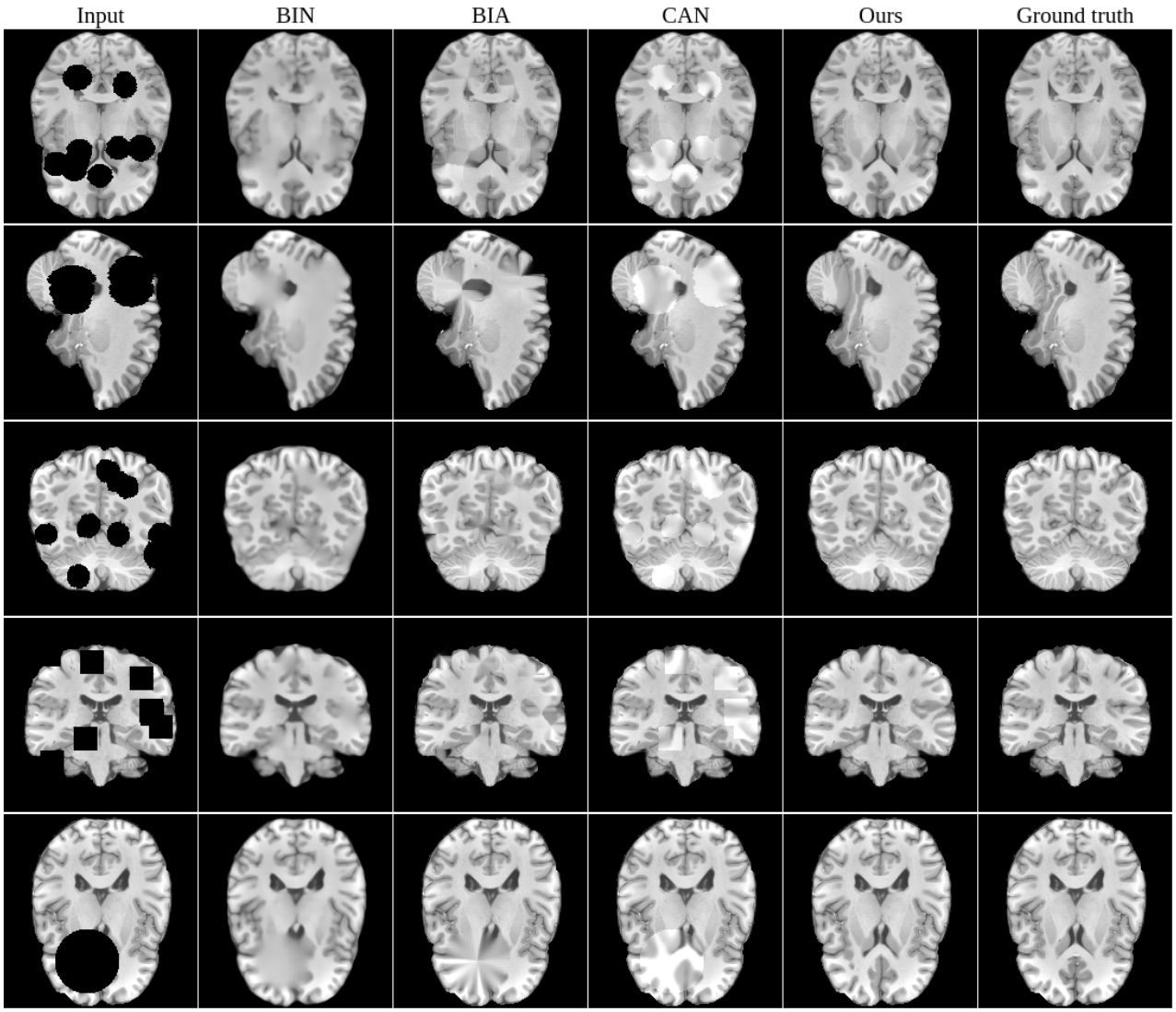

Figure 8. Qualitative comparative evaluation examples.

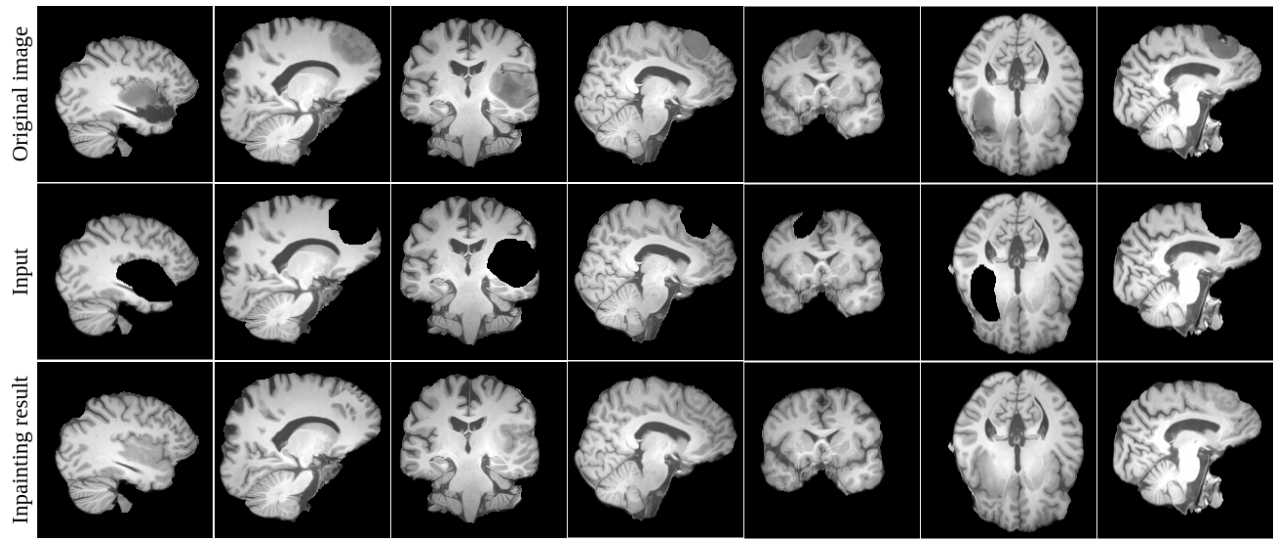

Figure 9. Glioma and meningioma inpainting examples.

Authorized licensed use limited to: University Bern. Downloaded on May 16,2024 at 07:50:27 UTC from IEEE Xplore. Restrictions apply.

IV. Conclusions and Future Work

In this work, we presented a user-guided deep gated convolutional network for high-resolution brain image inpainting. To improve the inpainting quality, we trained the network using a large dataset of 1000 brain volumes acquired from the Human Connectome Project. Besides, we proposed a novel linear-time algorithm to generate a new mask for every training image on the fly. Experiments revealed that gated convolution outperforms vanilla convolution in terms of convergence time, reconstruction error, and adversarial loss, especially for irregular masks. And compared to previous methods, our network provided better quantitative and qualitative inpainting results. Furthermore, the proposed network showed promising performance in filling brain tumors and accordingly improved the results of tissue segmentation.

However, there are several limitations of this work, which represent challenges for future work. First, the proposed network is limited to inpainting 2D images, and transferring to 3D might need a larger dataset. Second, increasing the input resolution will result in a longer training time. Hence, for higher resolutions, smaller networks are preferred for faster convergence. Finally, sometimes tumors and their surrounding edema cause distortions to the surrounding healthy tissues, which have an adverse effect on the inpainting quality. Hence, more complex strategies should be integrated into brain inpainting methods to neutralize this effect.

References

[1] B. Fischl, "Freesurfer," *Neuroimage*, vol. 62, no. 2, pp. 774–781, 2012.

[2] B. B. Avants, N. Tustison, and G. Song, "Advanced normalization tools (ants)," *Insight j*, vol. 2, no. 365, pp. 1–35, 2009.

[3] M. Jenkinson, C. F. Beckmann, T. E. Behrens, M. W. Woolrich, and S. M. Smith, "Fsl," *Neuroimage*, vol. 62, no. 2, pp. 782–790, 2012.

[4] J. V. Manjón, J. E. Romero, R. Vivo-Hernando, G. Rubio, F. Aparici, M. de la Iglesia-Vaya, T. Tourdias, and P. Coupé, "Blind mri brain lesion inpainting using deep learning," in *International Workshop on Simulation and Synthesis in Medical Imaging*. Springer, 2020, pp. 41–49.

[5] M. Sdika and D. Pelletier, "Nonrigid registration of multiple sclerosis brain images using lesion inpainting for morphometry or lesion mapping," Wiley Online Library, Tech. Rep., 2009.

[6] B. Nguyen, A. Feldman, S. Bethapudi, A. Jennings, and C. G. Willcocks, "Unsupervised region-based anomaly detection in brain mri with adversarial image inpainting," *arXiv preprint arXiv:2010.01942*, 2020.

[7] A. Criminisi, P. Perez, and K. Toyama, "Object removal by exemplar-based inpainting," in *2003 IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2003. Proceedings.*, vol. 2. IEEE, 2003, pp. II–II.

[8] D. Ulyanov, A. Vedaldi, and V. Lempitsky, "Deep image prior," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 9446–9454.

[9] V. Caselles, "Exemplar-based image inpainting and applications," *SIAM News*, vol. 44, no. 10, pp. 1–3, 2011.

[10] A. A. Efros and T. K. Leung, "Texture synthesis by non-parametric sampling," in *Proceedings of the seventh IEEE international conference on computer vision*, vol. 2. IEEE, 1999, pp. 1033–1038.

[11] M. Bertalmio, G. Sapiro, V. Caselles, and C. Ballester, "Image inpainting," in *Proceedings of the 27th annual conference on Computer graphics and interactive techniques*, 2000, pp. 417–424.

[12] A. A. Efros and W. T. Freeman, "Image quilting for texture synthesis and transfer," in *Proceedings of the 28th annual conference on Computer graphics and interactive techniques*, 2001, pp. 341–346.

[13] C. Ballester, M. Bertalmio, V. Caselles, G. Sapiro, and J. Verdera, "Filling-in by joint interpolation of vector fields and gray levels," *IEEE transactions on image processing*, vol. 10, no. 8, pp. 1200–1211, 2001.

[14] C. Barnes, E. Shechtman, A. Finkelstein, and D. B. Goldman, "Patchmatch: A randomized correspondence algorithm for structural image editing," *ACM Trans. Graph.*, vol. 28, no. 3, p. 24, 2009.

[15] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "Imagenet classification with deep convolutional neural networks," *Advances in neural information processing systems*, vol. 25, pp. 1097–1105, 2012.

[16] S. Ren, K. He, R. Girshick, and J. Sun, "Faster r-cnn: Towards real-time object detection with region proposal networks," *arXiv preprint arXiv:1506.01497*, 2015.

[17] O. Ronneberger, P. Fischer, and T. Brox, "U-net: Convolutional networks for biomedical image segmentation," in *International Conference on Medical image computing and computer-assisted intervention*. Springer, 2015, pp. 234–241.

[18] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, pp. 2536–2544.

[19] S. Iizuka, E. Simo-Serra, and H. Ishikawa, "Globally and locally consistent image completion," *ACM Transactions on Graphics (ToG)*, vol. 36, no. 4, pp. 1–14, 2017.

[20] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Generative image inpainting with contextual attention," in *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2018, pp. 5505–5514.

[21] G. Liu, F. A. Reda, K. J. Shih, T.-C. Wang, A. Tao, and B. Catanzaro, "Image inpainting for irregular holes using partial convolutions," in *Proceedings of the European Conference on Computer Vision (ECCV)*, 2018, pp. 85–100.

[22] J. Yu, Z. Lin, J. Yang, X. Shen, X. Lu, and T. S. Huang, "Free-form image inpainting with gated convolution," in *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 2019, pp. 4471–4480.

[23] H. WU-Minn, "1200 subjects data release reference manual," 2017.

[24] J. V. Manjón, P. Coupé, L. Martí-Bonmatí, D. L. Collins, and M. Robles, "Adaptive non-local means denoising of mr images with spatially varying noise levels," *Journal of Magnetic Resonance Imaging*, vol. 31, no. 1, pp. 192–203, 2010.

[25] N. J. Tustison, B. B. Avants, P. A. Cook, Y. Zheng, A. Egan, P. A. Yushkevich, and J. C. Gee, "N4itk: improved n3 bias correction," *IEEE transactions on medical imaging*, vol. 29, no. 6, pp. 1310–1320, 2010.

[26] J. Canny, "A computational approach to edge detection," *IEEE Transactions on pattern analysis and machine intelligence*, no. 6, pp. 679–698, 1986.

[27] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *arXiv preprint arXiv:1412.6980*, 2014.

[28] S. L. Smith, P.-J. Kindermans, C. Ying, and Q. V. Le, "Don't decay the learning rate, increase the batch size," *arXiv preprint arXiv:1711.00489*, 2017.

[29] X. Liu, F. Xing, C. Yang, C.-C. J. Kuo, G. ElFakhri, and J. Woo, "Symmetric-constrained irregular structure inpainting for brain mri registration with tumor pathology," *arXiv preprint arXiv:2101.06775*, 2021.

[30] Zhou Wang, A. C. Bovik, H. R. Sheikh, and E. P. Simoncelli, "Image quality assessment: from error visibility to structural similarity," *IEEE Transactions on Image Processing*, vol. 13, no. 4, pp. 600–612, 2004.

[31] H. Aerts and D. Marinazzo, ""btc_preop"," 2018.

[32] Y. Zhang, M. Brady, and S. Smith, "Segmentation of brain mr images through a hidden markov random field model and the expectation-maximization algorithm," *IEEE transactions on medical imaging*, vol. 20, no. 1, pp. 45–57, 2001.

Authorized licensed use limited to: University Bern. Downloaded on May 16,2024 at 07:50:27 UTC from IEEE Xplore. Restrictions apply.