
A APPENDIX

A.1 STATEMENT ON ETHICS AND PRIVACY

The value of medical data has increased dramatically in recent years. With it, the need for privacy has intensified. Concerns surround companies and government agencies that collect information from users, amid worries that medical records could be matched with data from smartphone use, internet search, or social media to identify patients. This sensitive information can be monetized in the form of ads served to those identified, or sold to insurance companies and hospitals. Bad actors could potentially extort patients with threats to reveal their medical history.

In response, legislation to protect the right to privacy has been enacted in many countries, notably the General Data Protection Regulation (GDPR) in the EU. While such measures give patients control over their identifiable data, regulations are at odds with society's interest to advance healthcare technology. In particular, deep neural networks depend on big data. Roadblocks imposed on medical data collection delay progress, potentially at the expense of our health. *Anonymization* or strong de-identification is an attractive workaround, as unidentifiable data can be used without patient consent. However, fool-proof removal of identifiable information from medical data, especially images, remains a challenging and unsolved problem.

In this work, we offer a new approach to better ensure privacy through remodeling of sensitive information. This allows data to be shared without compromising the value of the medically relevant information that is crucially important for improving technology in healthcare.

A.2 MODEL-AGNOSTIC PREPROCESSING

To ensure quality, comparable signal intensity distributions, and a consistent orientation of the acquired MR images, we apply various preprocessing steps. All steps can be seen as embodying standard procedure and have been proven various times (*e.g.* Jude Hemanth & Anitha (2012); Roy et al. (2013)) to positively affect algorithms. These preprocessing steps are applied prior to all de-identification methods in our study.

Orientation Correction Images within MR datasets are often not consistently aligned which ultimately hampers the learning process. Thus, we leverage FSL (FMRIB Software Library) Smith et al. (2004) to re-orient all images to the *radiological orientation convention*, the so-called *LAS* orientation.

Bias Field Correction The *bias field* is a low-frequency degradation of an MR scan due to magnetic field inhomogeneities which is typically imperceptible to humans. If unaccounted for, this degradation can induce different grayscale values on the same tissue type which, in turn, might impair downstream algorithms (Juntu et al., 2008). We use the nonparametric N4BiasFieldCorrection algorithm that comes with the ANTs library (Tustison et al., 2013).

Registration MR scans expose a high degree of variability primarily because of anatomical reasons but also because patients are typically not consistently positioned within the MRI tube. We therefore apply a non-rigid, double-affine *TRSAA* registration offered by the ANTs library to mitigate these effects. The necessary registration template is chosen uniformly at random *once* for *each* dataset.

Region-of-Interest Segmentation The toolkit Robex (Robust Brain Extraction, Iglesias et al. (2011)) was used to separate the brain from surrounding tissues. We have chosen this method as its segmentations turned out to be more robust than competing methods (*e.g.* Ségonne et al. (2004)). On the flip side, this surplus in accuracy comes at the expense of longer execution times.

White Stripe Normalization¹ As no significant performance differences between intensity normalization schemes has been reported (Reinhold et al., 2019), we opt for the comparatively simple *white strip normalization* that estimates both the mean μ_b and the (biased) standard deviation σ_b over the image voxels $x_{u,v,w}$ belonging to the brain tissue (as indicated by $b_{u,v,w}(x) = b_{u,v,w}$):

$$\left\{ \mu_b = \frac{\sum_{u,v,w \in \{1,\dots,S\}} b_{u,v,w} \cdot x_{u,v,w}}{\sum_{u,v,w \in \{1,\dots,S\}} b_{u,v,w}}, \sigma_b = \frac{\sum_{u,v,w \in \{1,\dots,S\}} b_{u,v,w} \cdot (x_{u,v,w} - \mu_b)^2}{\sum_{u,v,w \in \{1,\dots,S\}} b_{u,v,w}} \right\}$$

¹Only done for our model as other algorithms expect no change in scaling

Finally *all* voxel values are shifted and rescaled using the *z-score* transformation $\hat{x} = \frac{x - \mu_b}{\sigma_b}$.

A.3 LOSS FUNCTION CONSIDERATIONS

We have initially experimented with the relativistic *average* loss *Ra-LSGAN* variant suggested by (Jolicoeur-Martineau, 2018):

$$\begin{aligned}\mathcal{L}_G^{\text{RaLSGAN}} &= \mathbb{E}_{x_f \sim \mathcal{P}_Y} [(D_\Theta(x_f) - \mathbb{E}_{x_r \sim \mathcal{P}_X} D_\Theta(x_r) - 1)^2] + \mathbb{E}_{x_r \sim \mathcal{P}_X} [(D_\Theta(x_r) - \mathbb{E}_{x_f \sim \mathcal{P}_Y} D_\Theta(x_f) + 1)^2] \\ \mathcal{L}_D^{\text{RaLSGAN}} &= \mathbb{E}_{x_r \sim \mathcal{P}_X} [(D_\Theta(x_r) - \mathbb{E}_{x_f \sim \mathcal{P}_Y} D_\Theta(x_f) - 1)^2] + \mathbb{E}_{x_f \sim \mathcal{P}_Y} [(D_\Theta(x_f) - \mathbb{E}_{x_r \sim \mathcal{P}_X} D_\Theta(x_r) + 1)^2]\end{aligned}$$

where $\mathcal{P}_X, \mathcal{P}_Y$ denote the original resp. the fake distribution induced by G_Φ and we drop the conditioning variable $\gamma(x)$ from the notation. Observe, however, that this loss function is incompatible with our *conditional* scenario as $\mathbb{E}_{x_r \sim \mathcal{P}_X} D_\Theta(x_r)$ and $\mathbb{E}_{x_f \sim \mathcal{P}_Y} D_\Theta(x_f)$ are computed by averaging *across* scans associated to *different* conditional information. To solve this problem, we make sure that every patient can occur at most once in a batch and get rid of the aforementioned expectations:

$$\begin{aligned}\mathcal{L}_G &= \mathbb{E}_{(x_f, x_r) \sim (\mathcal{P}_Y, \mathcal{P}_X)} [(D_\Theta(x_f) - D_\Theta(x_r) - 1)^2] + \mathbb{E}_{(x_r, x_f) \sim (\mathcal{P}_X, \mathcal{P}_Y)} [(D_\Theta(x_r) - D_\Theta(x_f) + 1)^2] \\ &= 2\mathbb{E}_{(x_r, x_f) \sim (\mathcal{P}_X, \mathcal{P}_Y)} [(D_\Theta(x_r) - D_\Theta(x_f) + 1)^2] \\ \mathcal{L}_D &= \mathbb{E}_{(x_r, x_f) \sim (\mathcal{P}_X, \mathcal{P}_Y)} [(D_\Theta(x_r) - D_\Theta(x_f) - 1)^2] + \mathbb{E}_{(x_f, x_r) \sim (\mathcal{P}_Y, \mathcal{P}_X)} [(D_\Theta(x_f) - D_\Theta(x_r) + 1)^2] \\ &= 2\mathbb{E}_{(x_r, x_f) \sim (\mathcal{P}_X, \mathcal{P}_Y)} [(D_\Theta(x_f) - D_\Theta(x_r) + 1)^2]\end{aligned}$$

Observe that this is identical to the construction of *R-LSGAN* if we follow the principles of Jolicoeur-Martineau (2018).

A.4 BINARY DOWNSAMPLING

Suppose that we have some given input mask $m_0 \in \{0, 1\}^{2^n \times 2^n \times 2^n}$ for some fixed $n \in \mathbb{N}$. Let $(m')_{i=1, \dots, n-p} \in [0, 1]^{2^{n-i} \times 2^{n-i} \times 2^{n-i}}$ further denote the result of applying *average pooling* i times on m_0 and stopping at some minimal resolution $2^p \times 2^p \times 2^p$, $p \in \mathbb{N}$. A new sequence $(m')_{i=1, \dots, n-p}$ of binary representations can then be constructed by interpreting each voxel value of m'_i as a *Bernoulli* parameter determined by a *maximum likelihood* estimation over a (flattened) patch of $2^{i+1} \cdot 2^{i+1} \cdot 2^{i+1} = 2^{3(i+1)}$ (binary) realizations stemming from m_0 . Accordingly, this interpretation permits us to view m'_i as a volume of *Bernoulli parameters* from which we can derive m''_i by sampling in a voxel-wise fashion. Most importantly, this construction *preserves*² the non-zero $\zeta(m_0)$ proportion of m_0 in expectation (i.e. $(\frac{1}{2^n})^3 \sum_{u,v,w} m_0^{(u,v,w)} = \mathbb{E} \left[(\frac{1}{2^{n-i}})^3 \sum_{u,v,w} m_i^{(u,v,w)} \right]$ for $i = 1, \dots, n-p$). Although this downsampling scheme is stochastic in essence, we observe that the mapping becomes deterministic if a specific *Bernoulli* parameter is found by averaging over a region (patch) of constant realizations. Consequently, voxels of m'_i corresponding to the cerebral cortex (“brain boundary”) expose higher entropy than voxels from the interior of the brain. We conjecture that this makes the proposed generator more robust as it cannot rely on not having to model certain voxels.

A.5 EXECUTION TIME MEASUREMENT

As anonymization tools are meant to be applied within a clinical environment, it is important for them not to be overly time-consuming. Therefore, we measure the elapsed time³ that each method takes. To put all methods on an equal footing, we decide to start measuring the time just before the *NifTi-I*⁴ image is loaded and stop measuring once the de-identified *NifTi-I* scan was generated. Moreover, we decide to omit the time that was spent in *Model-Agnostic Preprocessing* since it is the same for all methods. All methods are single-threaded and we only execute one process at a time. The benchmark system is given by a Intel(R) Xeon(R) Silver 4216 CPU @ 2.10GHz, 252 GiB

²A proof of this can be found Section A.6

³Wall clock time

⁴NifTi-I is a widely-established format for MR imagery, used to ship OASIS-3 & ADNI

Method	Elapsed Time / sample [sec]
QUICKSHEAR	28 ± 3.7
MRI WATERSHED	101 ± 8.9
FACE MASK	91 ± 5.0
DEFACE	71 ± 3.0
C-DeID-GAN	71 ± 2.9

Table 1: *Execution Time Measurement*: Mean and standard deviations are aggregated over 200 runs on different images.

of RAM and 6 Quadro RTX 6000 GPUs (24 GiB main memory) of which we only use one. All methods are executed exactly 200 times.

In Table 1 we observe that C-DeID-GAN’s execution time is in line with the other de-identification methods. It ranks second, however QUICKSHEAR is more than two times faster than C-DeID-GAN and DEFACE, with the other two methods, FACE MASK and MRI WATERSHED, taking the most time per sample. It is worthwhile to mention that both QUICKSHEAR and C-DeID-GAN leverage the Robex (Iglesias et al., 2011) algorithm to compute a brain mask, which, when run in isolation, already takes 27 seconds per sample on average. This insight provides motivation to speed up either method by supplanting Robex with a faster algorithm if time constraints are a priority.

A.6 PROOFS

Sparsity Preservation of Binary Downsampling. Assume that $m \in \{0, 1\}^{2S \times 2S \times 2S}$ denotes some arbitrary binary image. Let $m' \in \{0, 1\}^{S \times S \times S}$ the result of performing $2 \times 2 \times 2$ average pooling on m . We interpret each voxel value $m'_{i,j,k}$ as a parameter to a *Bernoulli* distribution estimated by averaging over 8 voxel values from m allowing us to draw a sample from each voxel. The sampled result, denoted by m'' , is binary again and preserves the degree of sparsity $\zeta(m)$ of m in expectation, *i.e.*:

$$\zeta(m'') = \mathbb{E} \left[\frac{1}{S^3} \sum_{i,j,k} m''_{i,j,k} \right] = \frac{1}{(2S)^3} \sum_{i,j,k} m_{i,j,k} = \zeta(m)$$

Proof:

$$\begin{aligned}
\zeta(m'') &= \mathbb{E} \left[\frac{1}{S^3} \sum_{i,j,k=0}^{S-1} m''_{i,j,k} \right] \\
&= \mathbb{E} \left[\frac{1}{S^3} \sum_{i,j,k=0}^{S-1} z \sim \mathcal{B} \left(p = 1/8 \sum_{i_0=0}^1 \sum_{j_0=0}^1 \sum_{k_0=0}^1 m_{2i+i_0, 2j+j_0, 2k+k_0} \right) \right] \\
&= \frac{1}{S^3} \sum_{i,j,k=0}^{S-1} \mathbb{E} \left[z \sim \mathcal{B} \left(p = 1/8 \sum_{i_0=0}^1 \sum_{j_0=0}^1 \sum_{k_0=0}^1 m_{2i+i_0, 2j+j_0, 2k+k_0} \right) \right] \\
&= \frac{1}{S^3} \sum_{i,j,k=0}^{S-1} 1/8 \sum_{i_0=0}^1 \sum_{j_0=0}^1 \sum_{k_0=0}^1 m_{2i+i_0, 2j+j_0, 2k+k_0} \\
&= \frac{1}{(2S)^3} \sum_{i,j,k=0}^{2S-1} m_{i,j,k} = \zeta(m)
\end{aligned}$$

where the last step follows from the observation that every element in m occurs exactly once in the summation.

A.7 HYPERPARAMETERS

The table below provides a list of the hyperparameter values used in the experiments appearing in the main text.

Hyperparameter	Value
Batch size	2
Number of steps (OASIS-3, ADNI)	30,000
Min. generated resolution s	4
Max. generated resolution S	128
Number of blocks N_B	6
Kernel size k	5
Leaky ReLU steepness	0.2

A.8 BENCHMARK DE-IDENTIFICATION METHODS

We compare our result with four publicly available and widely-established methods for de-identification of MRI head scans, depicted in Figure 1 in the main text. All methods have in common that they (1) are not deep-learning-driven, (2) require no additional training and (3), are used on a day-to-day basis by clinical- and neuroscientists. All procedures were applied with default settings on images of resolution $128 \times 128 \times 128$.

QUICKSHEAR (Schimke et al., 2011) computes a plane to divide a given MRI into two parts: one containing facial structures, and the other containing the remainder of the scan. Voxels in the first part are set to zero.

FACE MASK (Milchenko & Marcus, 2013) uses a filtering method to blur the facial features. Based on registration to an atlas, the face region is identified, normalized and filtered. The result is transformed back to the original image space.

DEFACE (Bischoff-Grethe et al., 2007) estimates the probabilities of voxels belonging to the face based on an atlas of healthy control subjects. Intensities of voxels whose probabilities are small enough are set to zero.

MRI WATERSHED (Ségonne et al., 2004) de-identifies images by extracting only the brain using a watershed-based approach.

A.9 STUDY ON DE-IDENTIFICATION QUALITY

In the main text, we report the identification rate with worker uncertainty combined. Cases where the worker selected “uncertain” were randomly assigned one of the five de-identification choices. The reason for combining the uncertain responses is so that the results for each method reflect the same number of responses. In the table and figures below, we explicitly separate the correct identification rate from the uncertainty rate. The overall results do not change appreciably, but we notice that workers were most frequently uncertain for MRI WATERSHED. Apart from the evaluation in Figure 3, we also provide a box-whisker plot in Figure 8 which depicts the bootstrapped *mean identification rate* distribution of each method. The *percentage of uncertain cases* is visualized in Figure 7. We also report disentangled results

Method	Identification Rate [%]		Uncertainty Rate [%]	
	ADNI	OASIS-3	ADNI	OASIS-3
BLURRED	49	43	2.4	4.1
FACE MASK	45	35	7.6	8.8
DEFACE	45	40	6.9	8.6
QUICKSHEAR	37	37	11.2	6.9
MRI WATERSHED	24	18	43.8	41.1
C-DeID-GAN	21	25	6.8	9.5

Table 2: *Disentangled evaluation.* Rather than assigning uncertain cases to a random label, we can also only consider certain cases to calculate the identification rate. The results are slightly different but generally deliver the same message.

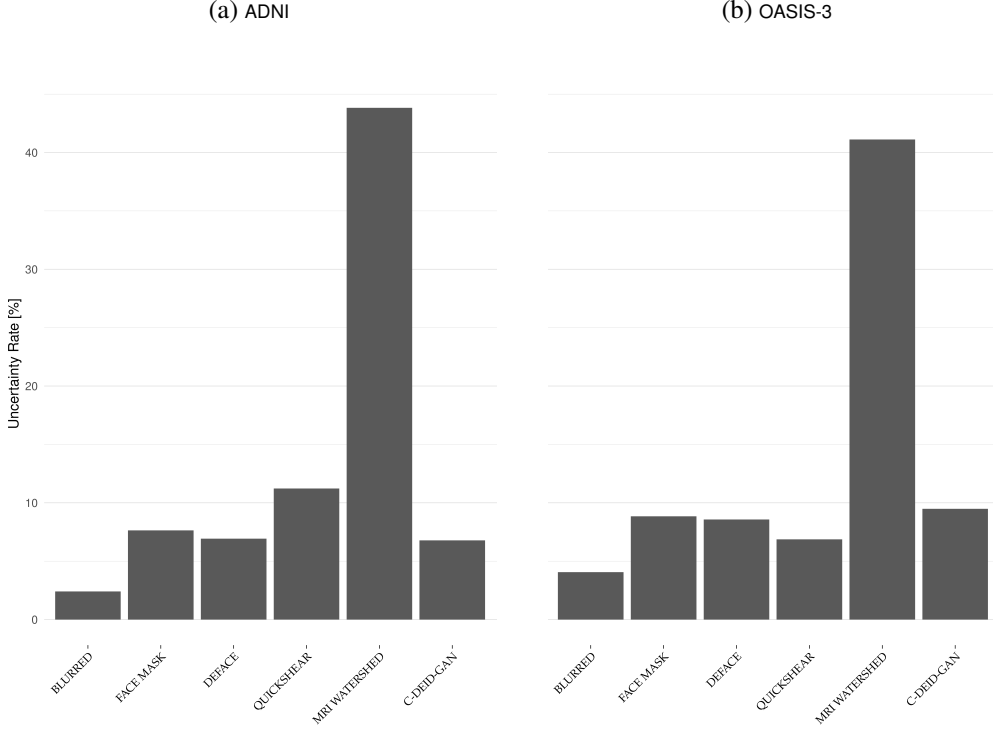


Figure 7: *Uncertainty in User Study*. Workers can explicitly express uncertainty by clicking on the question mark option (see Figure 3 in the main text). The *uncertainty rate* measures how frequently workers are uncertain on average. The high uncertainty rate registered for MRI WATERSHED can be explained by the nature of said method which solely retains the brain. As it is not clear how to map the original 2D rendering to one of the brains, workers thus resort to expressing uncertainty more often than when confronted to other methods.

A.10 DEEP LEARNING-BASED AGE PREDICTION

We use an adapted ResNet-18 He et al. (2015) architecture with 3D convolutions and Instance Normalization (Ulyanov et al., 2016) for the age prediction. A supervised setting with L_2 loss function and chronological age as ground truth labels is used to train the model. Random cropping (cropped input size $64 \times 96 \times 96$) and *gamma contrast* data augmentation is applied. Furthermore, we use the Adam optimizer (Kingma & Ba, 2015) with a learning rate of 10^{-4} . The batch size was chosen to be 16.

A.11 CONTROLLING SYNTHESIS VIA THE PRIVACY TRANSFORM

The scaling limits $(\alpha_{\min}, \alpha_{\max})$ shown in Figure 5 in the main text were chosen to reflect the distribution of brain sizes present in the data. We are interested in showing that the GAN is able to realistically synthesize volumes in approximately the same conditions that it has seen during training. We identify the 5% and 95% quantiles of the brain mask volumes, η_5 resp. η_{95} . From this, we set limits $\alpha_{\min} = \eta_5/\eta_{50} \approx 0.88$ and $\alpha_{\max} = \eta_{95}/\eta_{50} \approx 1.1$. Thus, $(\alpha_{\min}, \alpha_{\max})$ reflect the extremes of head size appearing in the training data.

A.12 LIMITATIONS AND FUTURE WORK

We note that certain existing privacy methods affect downstream tasks such as brain volume estimation and age prediction to a lesser extent than C-DeID-GAN, and a future line of research will be to improve the fidelity of GAN-generated volumes to mitigate this effect. This will likely entail addressing data limitations typical for medical imaging, especially by experimenting with data augmentation as recently suggested by (Karras et al., 2020). Another limitation of our work is the

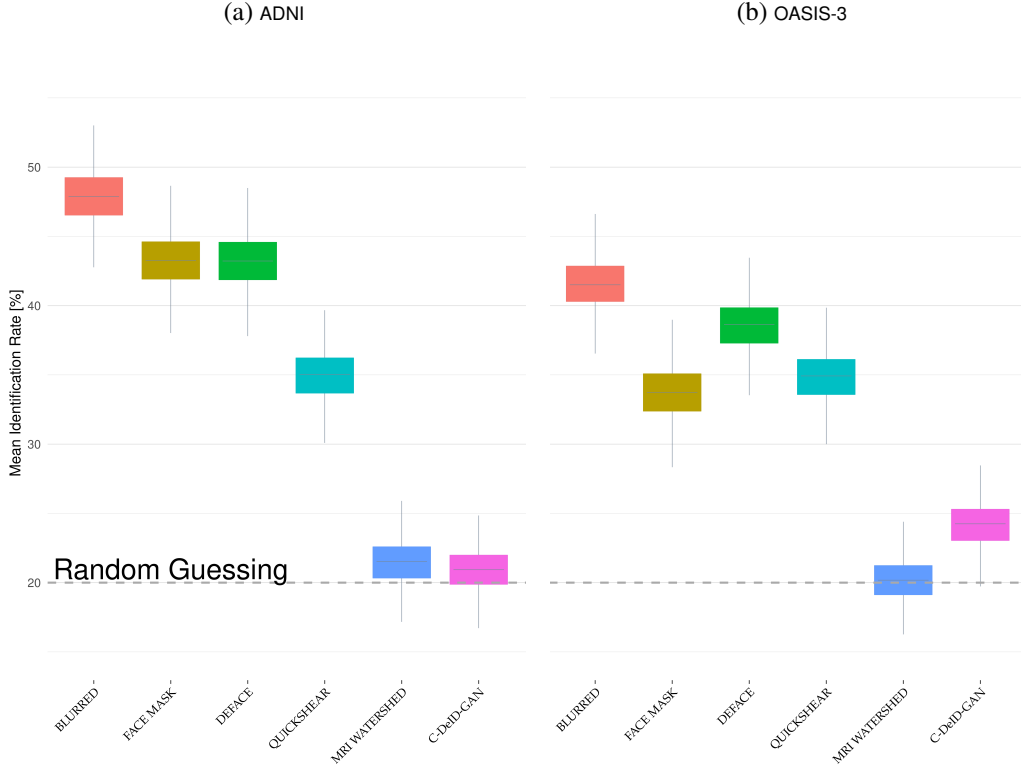


Figure 8: *Outcome of the User Study*: The above distributions are obtained by calculating the *mean identification rate* on 1,000 bootstrap samples for each method and dataset. Whenever a worker expressed uncertainty, a decision was sampled uniformly at random.

resolution. Other methods operate at the native $256 \times 256 \times 256$ resolution with little issue, but GANs have not yet been proven to generate content of such dimension due to several reasons, chief among them: memory limitations. Possible solutions may include network structures with more economic designs such as *Octree Generating Networks* (Tatarchenko et al., 2017). Finally, we are interested in seeing our model being deployed to preserve privacy in clinical and research settings. This will necessitate the creation of a model that works regardless of the source domain (*e.g.* scanner, population), motivating future work in domain adaptation.

A.13 SYNTHESIZED IMAGES

Below, we provide examples of synthesized images and renderings from our model and other de-identification methods.

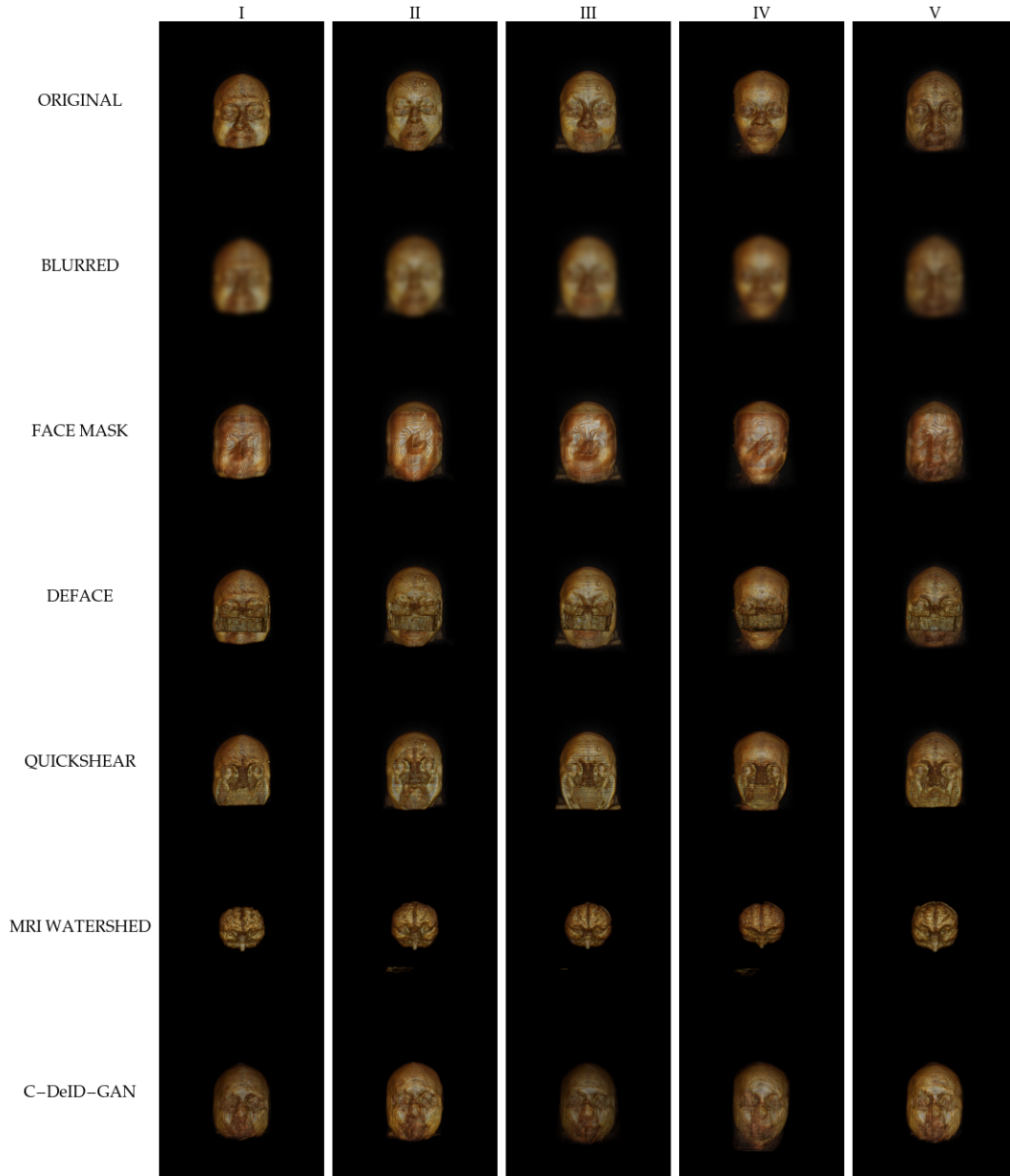


Figure 9: *Amazon Mechanical Turk User Study*: Renderings of five OASIS-3 subjects



Figure 10: *Amazon Mechanical Turk User Study*: Renderings of five ADNI subjects

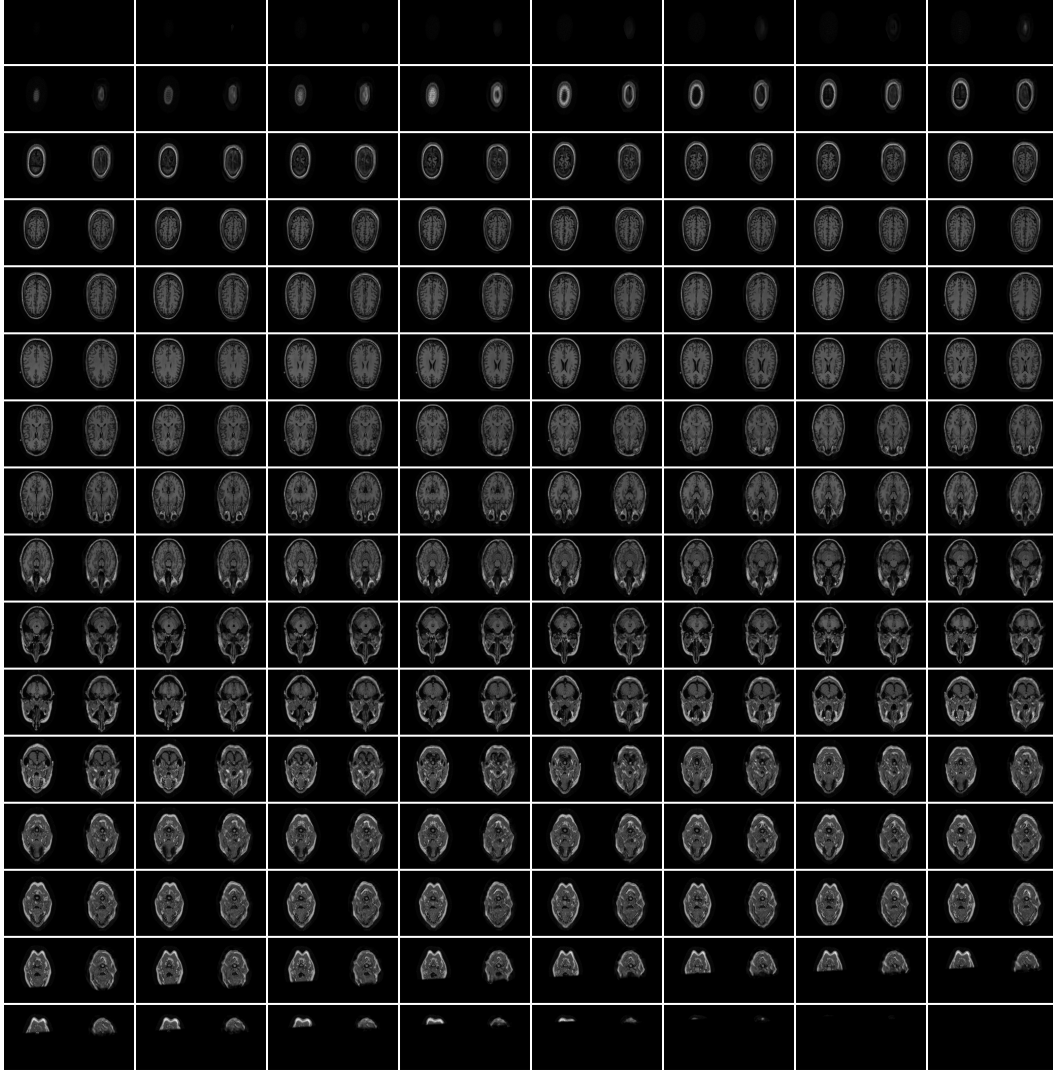


Figure 11: *Original MRI slices and those generated from C-DeID-GAN (OASIS-3): Slices (of a single patient) run from left to right and from top to bottom. Each box corresponds to one slice index and contains the original on the left and the synthesized counterpart on the right.*

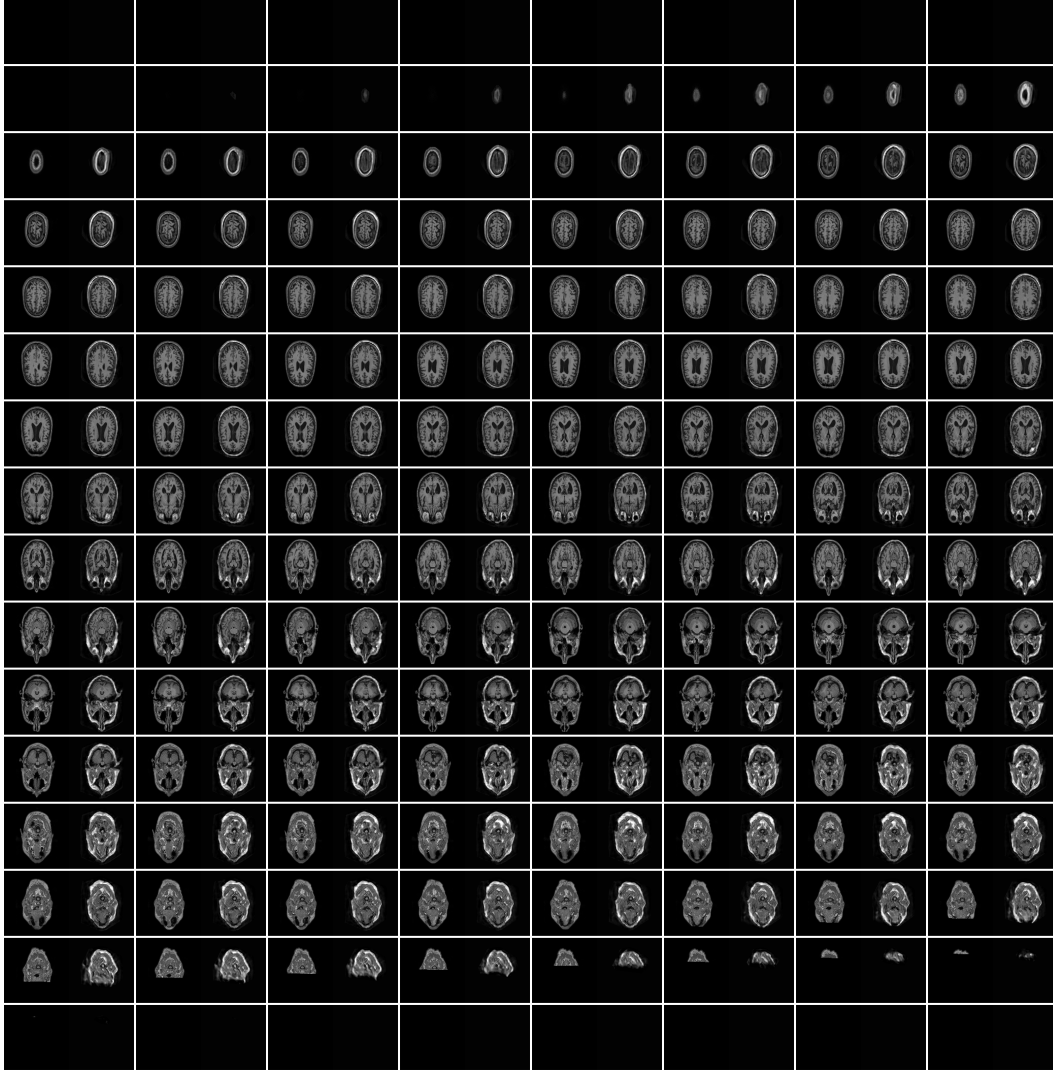


Figure 12: *Original MRI slices and those generated from C-DeID-GAN (ADNI): Slices (of a single patient) run from left to right and from top to bottom. Each box corresponds to one slice index and contains the original on the left and the synthesized counterpart on the right.*

REFERENCES

- Amanda Bischoff-Grethe, I. Burak Ozyurt, Evelina Busa, Brian T. Quinn, Christine Fennema-Notestine, Camellia P. Clark, Shaunna Morris, Mark W. Bondi, Terry L. Jernigan, Anders M. Dale, Gregory G. Brown, and Bruce Fischl. A technique for the deidentification of structural brain MR images. *Human Brain Mapping*, 28(9):892–903, sep 2007. ISSN 10659471. doi: 10.1002/hbm.20312.
- Kaiming He, Xiangyu Zhang, Shaoqing Ren, and Jian Sun. Deep residual learning for image recognition. *CoRR*, abs/1512.03385, 2015. URL <http://arxiv.org/abs/1512.03385>.
- Juan Eugenio Iglesias, Cheng Yi Liu, Paul M. Thompson, and Zhuowen Tu. Robust brain extraction across datasets and comparison with publicly available methods. *IEEE Transactions on Medical Imaging*, 30(9):1617–1634, sep 2011. ISSN 02780062. doi: 10.1109/TMI.2011.2138152.
- Alexia Jolicoeur-Martineau. The relativistic discriminator: a key element missing from standard GAN. *7th International Conference on Learning Representations, ICLR 2019*, jul 2018. URL <http://arxiv.org/abs/1807.00734>.
- D. Jude Hemanth and J. Anitha. Image pre-processing and feature extraction techniques for magnetic resonance brain image analysis. In Tai-hoon Kim, Dae-sik Ko, Thanos Vasilakos, Adrian Stoica, and Jemal Abawajy (eds.), *Computer Applications for Communication, Networking, and Digital Contents*, pp. 349–356, Berlin, Heidelberg, 2012. Springer Berlin Heidelberg. ISBN 978-3-642-35594-3.
- Jaber Juntu, Jan Sijbers, Dirk Dyck, and Jan Gielen. Bias Field Correction for MRI Images. pp. 543–551. Springer, Berlin, Heidelberg, oct 2008. doi: 10.1007/3-540-32390-2.64.
- Tero Karras, Miika Aittala, Janne Hellsten, Samuli Laine, Jaakko Lehtinen, and Timo Aila. Training generative adversarial networks with limited data. *arXiv preprint arXiv:2006.06676*, 2020.
- Diederik P. Kingma and Jimmy Lei Ba. Adam: A method for stochastic optimization. In *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*. International Conference on Learning Representations, ICLR, dec 2015.
- Mikhail Milchenko and Daniel Marcus. Obscuring surface anatomy in volumetric imaging data. *Neuroinformatics*, 11(1):65–75, jan 2013. ISSN 15392791. doi: 10.1007/s12021-012-9160-3. URL <http://www.ncbi.nlm.nih.gov/pubmed/22968671><http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=PMC3538950>.
- Jacob C. Reinhold, Blake E. Dewey, Aaron Carass, and Jerry L. Prince. Evaluating the impact of intensity normalization on MR image synthesis. In *Proceedings of SPIE—the International Society for Optical Engineering*, volume 10949, pp. 126. SPIE-Intl Soc Optical Eng, mar 2019. ISBN 9781510625457. doi: 10.1117/12.2513089.
- Sudipta Roy, Sanjay Nag, Indra Kanta Maitra, and Samir Kumar Bandyopadhyay. A review on automated brain tumor detection and segmentation from MRI of brain. *CoRR*, abs/1312.6150, 2013. URL <http://arxiv.org/abs/1312.6150>.
- Nakeisha Schimke, Mary Kuehler, and John Hale. Preserving privacy in structural neuroimages. In *Lecture Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, volume 6818 LNCS, pp. 301–308. Springer, Berlin, Heidelberg, 2011. ISBN 9783642223471. doi: 10.1007/978-3-642-22348-8_26.
- F. Ségonne, A. M. Dale, E. Busa, M. Glessner, D. Salat, H. K. Hahn, and B. Fischl. A hybrid approach to the skull stripping problem in MRI. *NeuroImage*, 22(3):1060–1075, jul 2004. ISSN 10538119. doi: 10.1016/j.neuroimage.2004.03.032.
- Stephen M. Smith, Mark Jenkinson, Mark W. Woolrich, Christian F. Beckmann, Timothy E.J. Behrens, Heidi Johansen-Berg, Peter R. Bannister, Marilena De Luca, Ivana Drobnjak, David E. Flitney, Rami K. Niazy, James Saunders, John Vickers, Yongyue Zhang, Nicola De Stefano, J. Michael Brady, and Paul M. Matthews. Advances in functional and structural MR image analysis and implementation as FSL. In *NeuroImage*, volume 23, pp. S208–S219. Academic Press, jan 2004. doi: 10.1016/j.neuroimage.2004.07.051.

Maxim Tatarchenko, Alexey Dosovitskiy, and Thomas Brox. Octree generating networks: Efficient convolutional architectures for high-resolution 3d outputs. *CoRR*, abs/1703.09438, 2017. URL <http://arxiv.org/abs/1703.09438>.

Nicholas J. Tustison, Brian B. Avants, Philip A. Cook, Gang Song, Sandhitsu Das, Niels van Strien, James R. Stone, and James C. Gee. The ANTs cortical thickness processing pipeline. In John B. Weaver and Robert C. Molthen (eds.), *Medical Imaging 2013: Biomedical Applications in Molecular, Structural, and Functional Imaging*, volume 8672, pp. 86720K. SPIE, mar 2013. doi: 10.1117/12.2007128. URL <http://proceedings.spiedigitallibrary.org/proceeding.aspx?doi=10.1117/12.2007128>.

Dmitry Ulyanov, Andrea Vedaldi, and Victor Lempitsky. Instance Normalization: The Missing Ingredient for Fast Stylization. jul 2016. URL <http://arxiv.org/abs/1607.08022>.