Anonymization and validation of three-dimensional volumetric renderings of computed tomography data using commercially available T1-weighted magnetic resonance imaging-based algorithms

Rahil Patel, Destie Provenzano, and Murray Loew*

George Washington University School of Engineering and Applied Science, Department of Biomedical Engineering, Washington, District of Columbia, United States

ABSTRACT. **Purpose:** Previous studies have demonstrated that three-dimensional (3D) volumetric renderings of magnetic resonance imaging (MRI) brain data can be used to identify patients using facial recognition. We have shown that facial features can be identified on simulation-computed tomography (CT) images for radiation oncology and mapped to face images from a database. We aim to determine whether CT images can be anonymized using anonymization software that was designed for T1-weighted MRI data.

Approach: Our study examines (1) the ability of off-the-shelf anonymization algorithms to anonymize CT data and (2) the ability of facial recognition algorithms to identify whether faces could be detected from a database of facial images. Our study generated 3D renderings from 57 head CT scans from The Cancer Imaging Archive database. Data were anonymized using AFNI (deface, reface, and 3Dskullstrip) and FSL's BET. Anonymized data were compared to the original renderings and passed through facial recognition algorithms (VGG-Face, FaceNet, DLib, and SFace) using a facial database (labeled faces in the wild) to determine what matches could be found.

Results: Our study found that all modules were able to process CT data and that AFNI's 3Dskullstrip and FSL's BET data consistently showed lower reidentification rates compared to the original.

Conclusions: The results from this study highlight the potential usage of anonymization algorithms as a clinical standard for deidentifying brain CT data. Our study demonstrates the importance of continued vigilance for patient privacy in publicly shared datasets and the importance of continued evaluation of anonymization methods for CT data.

© 2023 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JMI.10.6.066501]

Keywords: anonymization; deidentification; three-dimensional rendering; facial recognition; computed tomography

1 Introduction

Brain scan imaging data, such as computed tomography (CT) and magnetic resonance imaging (MRI), can create three-dimensional (3D) volumetric images of facial regions. These 3D renderings allow for improved visualization and analysis of medical images, which has led to

significant progress for clinical and research applications of medical data. As the implementation of 3D renderings of imaging data grew, however, increased attention was drawn to the unique challenges of neuroimaging data.[1] Specifically, with technological advancements, there is a possibility for faces to be recognized from the brain scans. Studies have now demonstrated this ability to identify patients from these renderings not only visually by human review but also using commonly available facial recognition algorithms with great accuracy.[2–7] This poses a major privacy concern for public data sharing and patient privacy as the identification of an individual can lead to access of private information such as medical diagnosis, genetic data, and other personal medical information.

This ability to identify facial features from 3D brain scan data renders the current standard of removing only metadata from medical images inadequate. As an attempt to further protect patient privacy, anonymization algorithms have been developed to deidentify the 3D rendering of a medical image, predominantly geared at T1-weighted (T1W) MRI data. These anonymization algorithms work by skull-stripping, defacing, or refacing the identifiable facial features, such as the eyes, nose, mouth, and ears. Skull-stripping algorithms isolate the brain from extracranial or nonbrain tissues from MRI head scans.[8] Defacing algorithms remove voxels containing data that could be used to render identifiable facial features and leave the remaining imaging data intact.[9] The complete stripping or large removal of facial surface features in skull-stripping and defacing algorithms, respectively, make it difficult for facial detection. Refacing algorithms do not remove any data and instead, replace the face on the image by applying scaled precreated masks of values from a population average.[10,11] Other current methods of anonymization of 3D brain scan data include local deformation[2] and blurring of the facial region[3]. Local deformation methods work by applying significant deformations to defining facial characteristics on the image while preserving the internal anatomy. Blurring techniques apply facial masks to the image that smooth the facial surface to blur the anatomical structures of the face and prevent identification. Previous studies have shown success in the ability of anonymization algorithms to reduce tracing back to the original medical data.[11–13]

We previously investigated if facial features could be identified on simulation-CT data of patients wearing an immobilization mask used for treatment for radiation oncology.[14] We found that patients were able to be reidentified with high accuracy. It is unknown whether CT images can be anonymized using publicly available anonymization software, which are typically designed for T1W MRI data. Our study sought to investigate the ability of off-the-shelf software anonymization software used for T1W MRI data to deidentify CT data and to assess the ability of facial recognition algorithms to then identify whether anonymized images could be identified in a database of facial images.

2 Methods

2.1 Overview
The purpose of this study is to evaluate if anonymization algorithms designed for T1W MRI data can work on CT data and, if so, if facial features can still be recognized and matched to an original patient image. This was done in three parts: (1) head CT and MRI data were collected from The Cancer Imaging Archive (TCIA) and OpenNeuro, respectively. (2) Data were anonymized using off-the-shelf T1W MRI-based anonymization algorithms (Afni_deface, Afni_reface, AFNI's 3Dskullstrip, and FSL's BET) and original and anonymized data were visualized in 3D Slicer. (3) Data were then run through various facial recognition softwares (VGG-Face, FaceNet, DLib, and SFace) and were then evaluated for true positive/false positive and true negative/false negative matches. The true positive (matched to original after anonymization) and false positive rates (matched to different face images) are reported.

2.2 Data Acquisition
This study is a follow-up to our previous study in which we found that facial features were found in CT images and correctly mapped to corresponding patient faces when included in a database of publicly available facial images (LFW: labeled faces in the wild).[15] For this study, we acquired new head CT data and tested whether the CT data could be anonymized. We acquired head CT scans of 57 patients from TCIA database and generated 3D renderings of the patients. The

Table 1 Description of CT and MRI datasets.

Dataset	Database	Description	Imaging modality used for our study	Number of patients acquired for our study
Low-dose CT image and projection data (LDCT-and-projection-data)	TCIA	The library contains a collection of medical scans, including noncontrast head CT scans used to assess cognitive or motor deficits, low-dose noncontrast chest scans for screening high-risk patients for lung nodules, and contrast-enhanced abdominal CT scans to detect metastatic liver lesions. The collection includes a total of 99 head scans, 100 chest scans, and 100 abdominal scans	CT	55
ACRIN-FMISO-Brain	TCIA	The ACRIN 6684 clinical trial aimed to link baseline FMISO PET uptake and MRI metrics with overall survival, disease progression time, and 6-month progression-free survival in individuals recently diagnosed with glioblastoma multiforme (GBM). The collection comprises MRI, PET, and low-dose CT images	CT	1
The CPTAC-LSCC	TCIA	This dataset consists of individuals from the CPTAC-LSCC cohort, established by the National Cancer Institute. TCIA hosts radiology (CT and PET) and pathology images of CPTAC patients	CT	1
Cortical myelin measured by the T1W/T2W ratio in individuals with depressive disorders and healthy controls	OpenNeuro	The dataset is a part of the ongoing longitudinal study that measures cortical myelin in individuals with depressive disorders and healthy controls. The current dataset includes T1W, T2W, and spin echo field maps for the participants	MRI	1
In-scanner head motion and structural covariance networks	OpenNeuro	This dataset consists of whole brain T1W MRI scans obtained in healthy adult controls. Each participant has two MRI scans: an MRI scan with motion-related artifacts and a usual still head MRI scan	MRI	1

datasets used were ACRIN-FMISO-Brain, the Clinical Proteomic Tumor Analysis Consortium Lung Squamous Cell Carcinoma Collection (CPTAC-LSCC), and low-dose CT image and projection data (LDCT-and-Projection-data).[16–22] Open-source T1W and T2-weighted (T2W) MRI data were also used from OpenNeuro.[23,24] MRI data were included as a control to demonstrate that the anonymization software worked prior to its application to CT data. Description of the used datasets is given in Table 1.

2.3 Data Processing and Anonymization

The data were first preprocessed to convert the original DICOM data to a usable form for the anonymization software. We generated 3D renderings of the head CT and MRI data using 3D Slicer prior to anonymization as a control.[25,26] DICOM files of the data were converted to the NIfTI format using icometrix: dicom2nifti.[27] Original CT and MRI DICOM files were converted to NIfTI as the anonymization algorithms used for our study only process NIfTI files. The converted NIfTI data were then anonymized using algorithms from AFNI (Analysis of Functional NeuroImages [NIH]: deface, reface, and 3Dskullstrip) and FSL (FMRIB Software Library [Oxford]: BET).[10,28–32] Anonymization algorithms were used to mask or remove identifiable facial features from the CT or MRI data. These algorithms work differently; descriptions of the T1W MRI-based algorithms used in our study are summarized in Table 2.

Final visualization used 3D slicer. MRI and CT images are two modalities of medical imaging that serve different functional purposes in terms of depth and clarity of imaging. A CT scan takes a fast series of X-rays and reconstructs them to create a full image of the area scanned. MRI uses strong magnetic fields and radiofrequency signals to image the inside of the body to visualize soft tissue. The standardized units and scales of the imaging modalities are different as well. CT uses Hounsfield units for each pixel grayscale representation that ranges from -1000 to 3000. MRI uses a different grayscale that typically ranges from 0 to 255. Additional processing was required to account for these structural differences between CT and MRI to obtain the appropriate results postanonymization. No additional processing of the data were needed for the defacing and refacing algorithms to function. The algorithms were able to process the unprocessed NIfTI file of the head CT data. Skull-stripping algorithms, however, required additional processing, including thresholding and smoothing the image and applying the filled skull-stripped mask to the original image.[33] All CT data required processing postanonymization using a custom algorithm. The MRI-based anonymization algorithms altered the CT data according to the value of air for MRI, which is 0. However, the CT value of air is -1000. This resulted in an image of a black box

Table 2 Summary of T1W MRI-based anonymization algorithms.

Anonymization algorithm	Method of anonymization	Description of algorithm	Citation
Afni_deface	Defacing	Uses a precreated mask to modify the image by replacing the voxel values of the face and ears with zero	Cox et al.[28]
Afni_reface	Refacing	Modifies the image by applying a scaled replacement "generic face" mask that was created by warping and averaging several public datasets; Modifies the same regions as Afni_deface	Cox et al.[28]
FSL's BET	Skull-stripping	A deformable model that makes an intensity-based estimation of the brain and nonbrain threshold, determines the center of gravity of the head and defines a sphere at that point, and expands until it reaches the brain edge	Smith et al.[32]
AFNI's 3Dskullstrip	Skull-stripping	Modified version of BET; uses additional processing such as nonuniformity correction and edge detection to minimize errors	"AFNI: 3Dskullstrip," National Institute of Mental Health[29]

that surrounded the head anatomy when rendered into a 3D image using 3D slicer. We developed a custom algorithm that modified the values of the anonymized CT image to match the appropriate Hounsfield units of air outside the patient volume, changing the pixel values of 0 to −1000. Our example code is found in Table S1 in the Supplementary Material. This provided a 3D rendering of the CT image that matched that of an anonymized MRI.

2.4 Facial Recognition

LFW was used as an expanded facial image dataset to determine the generalizability of our results. The LFW dataset consists of 13,233 images and 5749 people (with 1680 of them having two or more images). LFW was augmented with the 57 CT control images. Snapshots of the head in the form of PNG images were taken in the forward-facing position. Regions were limited to the head; any other region of the human body in the image was removed. The original and anonymized data were then passed through facial recognition algorithms (VGG-Face, FaceNet, DLib, and SFace) using LFW to determine whether matches between the control images and anonymized images could be found.[34–38] Description of the facial recognition algorithms used in this study are in Table 3. As in our previous study, any incorrect match of a putatively anonymized image to an LFW image was considered an incorrect match. During facial recognition, a parameter called "enforce detection" was turned off for VGG-Face, FaceNet, DLib, and SFace. Those algorithms would not run to completion if they were not able to detect a face in an image. With the "enforce detection" parameter off, if a face was not able to be detected, the algorithm would continue to analyze the image as a whole.

Facial recognition algorithms determine a match by comparing the distance between two images using measures of similarity, such as cosine distance or Euclidean distance with a determined tolerance value. Tolerance values are the set threshold distance for which two images are deemed a match. The lower the threshold is, the stricter the algorithm is in recognizing a match. For the purpose of our study, each facial recognition algorithm determined a match based on the cosine distance of the images. The set threshold values for our study for the cosine distance of VGG-Face, FaceNet, DLib, and SFace are 0.2, 0.2, 0.04, and 0.3, respectively.

2.5 Statistical Analysis

Statistical analysis was conducted to evaluate the performance of the facial recognition algorithms when applied to the images subjected to the different anonymization algorithms (Afni_deface, Afni_reface, FSL's BET, and AFNI's 3Dskullstrip). To gauge the efficacy of these anonymization methods, we computed specificity and sensitivity metrics which measure the

Table 3 Description of facial recognition algorithms.

Facial recognition algorithm	Description of algorithm	Citation
VGG-Face	A model that uses a deep CNN with 22 layers and 37 deep units that converts faces into feature vector representations and employs a triplet loss function to make them generalizable	Parkhi et al.[34]
FaceNet	A deep CNN model built by Google based on the inception model that uses a triple-based loss function to directly create 128D embeddings; similar to VGG-Face, the model represents the images as small dimension vectors and uses similarity to determine the identification	Schroff et al.[35]
DLib	A CNN model with a ResNet-34 network with built convolution layers that uses a facial map created by HOG to generate 128D vectors and checks similarity for facial recognition	King[36]
SFace	A CNN model that uses a loss function called sigmoid-constrained hypersphere loss that imposes intraclass and interclass constraints on a hypersphere manifold that allows for better balance between underfitting and overfitting, further improving generalizability of deep face models	Zhong et al.[37]

facial recognition algorithms' capacity to correctly identify true negative/false positive and true positive/false negative matches, respectively. The sensitivity represents the ability of the system to correctly match faces to their originals after anonymization. The specificity measures the ability of the system to correctly identify nonmatching images.

3 Results

3.1 Visual Anonymization Results

As shown in Figs. 1 and 2, 3D visualization of the control images and anonymized images reveal that all of the off-the-shelf anonymization algorithms were able to anonymize the CT data by altering the 3D rendering of the patient image. The skull-stripping algorithms, FSL's BET and AFNI's 3Dskullstrip, successfully extracted the brain from the surrounding tissue and completely removed all identifiable facial features from the rendering. Afni_deface worked to deface the image by replacing the values of the face and ear with zero. Afni_deface defaced the CT image and removed facial features comparably to the anonymized T1W and T2W images. Afni_deface has some minor gray artifacts present on the CT scan exterior of the face after processing. Afni_reface successfully refaced all medical images by replacing the face and ears with artificial values.
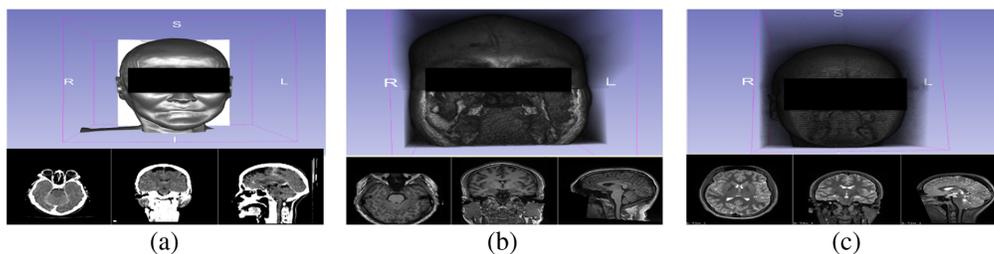

Fig. 1 3D volumetric rendering and imaging scan from (a) computed tomography, (b) T1W magnetic resonance imaging, and (c) T2W MRI.
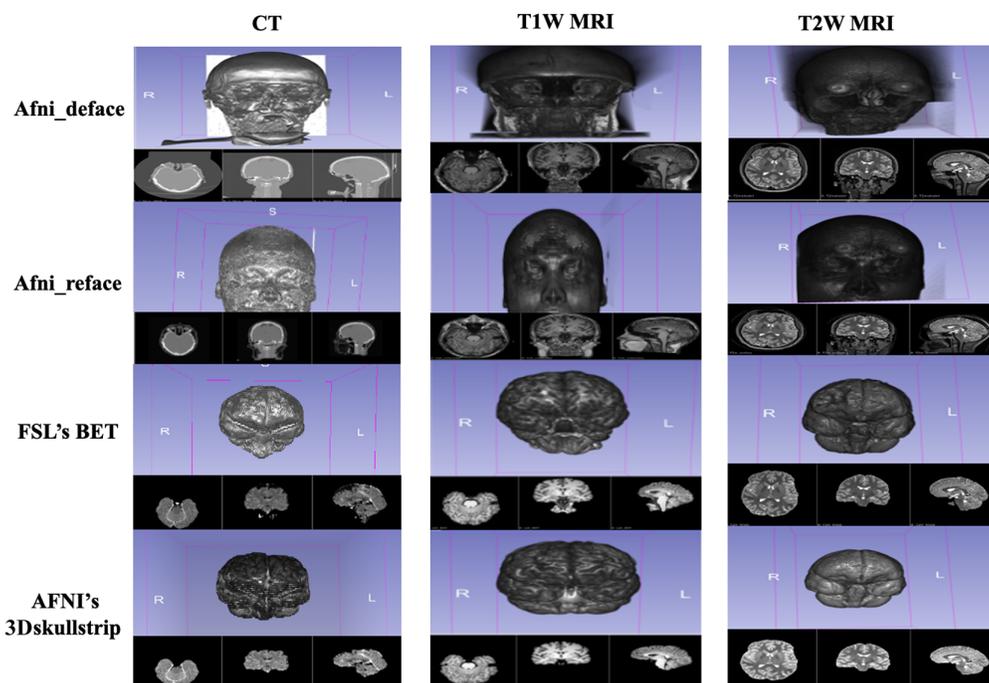

Fig. 2 Comparison of anonymized 3D volumetric rendering images from CT, T1W, and T2W modalities.

Table 4 Comparison of the total number of anonymized images matched to the original patient image by the facial recognition algorithm.

Facial recognition algorithm	Number of matches to original image				
	Original	Afni_deface	Afni_reface	FSL's BET	AFNI's 3Dskullstrip
VGG-Face	57	42	33	5	4
FaceNet	57	14	5	3	2
D-Lib	57	31	8	4	0
SFace	57	14	0	0	0

Table 5 Comparison of the total number of false positive matches of each image by the facial recognition algorithm.

Facial recognition algorithm	Number of false positive images matched				
	Original	Afni_deface	Afni_reface	FSL's BET	AFNI's 3Dskullstrip
VGG-Face	600	473	260	100	169
FaceNet	407	1674	2959	5893	4945
D-Lib	91	68	47	80	90
SFace	0	0	0	0	0

3.2 Facial Recognition Results

The total number of matched images for each anonymized image by the facial recognition algorithm is shown in Table 4. A baseline of all facial recognition algorithms was established by running the control images against themselves. For all facial recognition algorithms, all 57 head CT images were able to be matched back to each other. In comparison to the original Afni_deface, and Afni_reface images, the skull-stripping algorithms FSL's BET and AFNI's 3Dskullstrip showed significantly lower rates of reidentification for all facial recognition algorithms, with BET and 3Dskullstrip being zero for SFace and 3Dskullstrip being zero for D-Lib. Afni_reface also showed low rates of reidentification in all facial recognition algorithms except VGG-Face. Afni_deface comparatively showed the highest rates of reidentification across all facial recognition algorithms.

All algorithms except SFace had false positive matches to images in the LFW dataset as shown in Table 5. For both VGG-Face and D-Lib, the original image had the highest number of false positive matches, whereas for FaceNet, the original image had the lowest number of false positive matches. For FaceNet and D-Lib, FSL's BET and AFNI's 3Dskullstrip had higher false positives matches when compared with Afni_deface and Afni_reface.

3.3 Statistical Analysis

Specificity and sensitivity values for each anonymized image by facial recognition are given in Table 6. All anonymized images resulted in a decrease in sensitivity compared to the original. Sensitivity values of the skull-stripping algorithms, AFNI's 3Dskullstrip and FSL's BET, were the lowest across all images and all facial recognition algorithms. Afni_reface had comparatively lower sensitivity values for all facial recognition algorithms except VGG-Face. Afni_deface consistently had higher sensitivity values compared with the other anonymized images across all facial recognition algorithms. The specificity values varied across the anonymization methods and facial recognition algorithms, but the facial recognition algorithms, D-Lib and SFace, consistently had the highest specificity.

Table 6 Comparison of specificity and sensitivity metrics across anonymized images and facial recognition algorithms.

Facial recognition algorithm	Original		Afni_deface		Afni_reface		FSL's BET		AFNI's 3Dskullstrip	
	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity	Specificity	Sensitivity
VGG-Face	0.95	1	0.96	0.88	0.98	0.87	0.99	0.08	0.99	0.07
FaceNet	0.97	1	0.87	0.20	0.78	0.08	0.56	0.05	0.63	0.03
D-Lib	0.99	1	0.99	0.54	0.99	0.12	0.99	0.07	0.99	0
SFace	1	1	1	0.20	1	0	1	0	1	0

4 Discussion

The clinical and research community commonly practices sharing of medical imaging data. The current standard of deidentifying medical imaging data is to remove metadata. With the ability to render medical images as 3D volumetric images, however, it makes the current standard inadequate because the underlying image remains identifiable. Schwartz et al.[7] published a study in the New England Journal of Medicine that demonstrated a key finding that patients could be identified based on 3D renderings of MRI images of the brain. The study found that 83% of the patients were able to be identified by facial recognition algorithms. The findings from that study raised serious concerns about current privacy practices. The identification of an individual can lead to illegal exploitation of health information.[39] Acknowledging this concern, we investigated whether a similar privacy concern existed in the sharing of CT simulation radiation therapy data of brain tumor or head and neck cancer patients. Similarly, our work demonstrated that a thermoplastic mask used for radiation oncology treatment planning was not sufficient to prevent identification and that 83% of patients were able to be recognized with a publicly available facial recognition algorithm.[14] Our findings further emphasized the need to change the current standard of only removing metadata for head and neck imaging data not only for MRI data but also for CT data.

In this study, we investigated the ability of off-the-shelf anonymization software designed for T1W MRI data to anonymize CT data and prevent facial recognition identification of the original data. CT and MRI imaging are different imaging modalities not only in their usage but also in how they function and how the data are structured. The grayscale pixels in each modality are represented by different values. The significance of our work is that none of the anonymization algorithms used in this study have been designed for CT data. All of the anonymization software packages were able to anonymize the CT data to a certain extent. When the anonymized images were run against the LFW dataset modified with the control image using various facial recognition algorithms, we found that the skull-stripped images, FSL's BET and AFNI's 3Dskullstrip, showed significantly lower rates of reidentification when compared across all facial recognition algorithms. Data anonymized by Afni_reface and Afni_deface varied in terms of detection performance among the facial recognition algorithms. As sensitivity was defined as the ability of the system to correctly match faces to their originals after anonymization, the skull-stripped images resulted in the greatest decrease in performance of all facial recognition algorithms. The results from this study highlight the effectiveness of anonymization algorithms in preventing reidentification of CT data and the potential ability to use T1W MRI-based anonymization software as a clinical standard for deidentifying CT data. Further, they emphasize the need for greater attention to be drawn to anonymization methods for CT data and anonymization software in general for all medical imaging modalities.

This study was limited in its sample size. It sought to use only publicly available medical imaging data in its evaluation of anonymization. There is a limited amount of publicly available CT data, and the publicly available MRI data released are already anonymized. Visual evaluation of the control anonymized images for T1W and T2W MRI was already partially anonymized when run through the anonymization software, which explains the differences in results

compared with CT. Usage of additional medical CT and MRI data may provide greater insight into the effectiveness of publicly available anonymization algorithms. Additionally, this study did not evaluate the loss of anatomical structures within the head and brain but focused on the surface alterations to prevent identification. Anonymization algorithms could potentially induce alterations that reduce the effective utilization of data derived from medical data images by changing the appearance of the brain.

Additional considerations include the differences between CT and MRI volume renderings versus standard facial images. CT and MRI are primarily designed for clinical diagnostics and research, resulting in different imaging protocols and resolutions compared to the standard photographs for which facial recognition algorithms are predominantly used. This leads to variations in image quality and appearances for medical imaging data. Facial recognition with traditional facial images considers cues such as lighting and expressions that aid recognition. CT and MRI volumetric renderings lack such cues, however, and may include medical devices, thermoplastic masks, or artifacts that further complicate recognition. Facial recognition of a 3D medical imaging rendering does not consider the full volumetric ability of the medical data as distinguishing features could be present on the lateral or dorsal region of the head. For the purposes of our study, we used face-forward images of the patient and removed any external artifacts to the best of our ability to account for such considerations and differences between the CT volumetric rendering and the traditional facial images in the LFW dataset.

Techniques such as supervised and unsupervised machine learning on extensive datasets as well as the concept of data fusion have added increased complexity to the challenge of anonymization. These new machine learning techniques leverage the power of algorithms to discern patterns that may not be immediately apparent through manual analysis. This could potentially allow for methods to reidentify patient data despite current efforts to conceal the patient's identity while still allowing data exchange for research. In addition, the concept of data fusion is becoming increasingly prevalent as medical imaging data is being fused with other relevant sources, such as genetic profiles and demographic information. The amalgamation of the different types of data enhances the ability to reidentify the patient. Such techniques cause significant concern about the navigating the balance between leveraging the available data for clinical and research purposes and ensuring the protection of patient identities. Our results along with the rising concern of new techniques underscores the pressing need to continue evaluating and improving methods for the anonymization of medical imaging data, including CT, to uphold individual privacy as medical research advances.

The exploration of anonymization in medical imaging requires further inquiry, driven by the need to safeguard patient privacy while preserving diagnostic value. Future endeavors to enhance anonymization techniques might involve the development of purpose-built AI algorithms explicitly designed to uncover vulnerabilities in reidentifying anonymized images. Other possibilities include delving into the realm of image quality and the effects of machine variations as they hold the potential to unveil patient identities, imaging sites, scanner configurations, and geographic origins. The spectrum of possible investigations is extensive, reflecting the ever-evolving landscape of technology. As technology progresses, the importance of protecting patient privacy grows, necessitating an increased focus on avenues for robust patient data protection and anonymization evaluation.

5 Conclusion

Our study found that all modules were able to process CT data in addition to T1W and T2W data and that all modules anonymized patient data such that there were reductions in the reidentification rate of the patients. AFNI's 3Dskullstrip and FSL's BET anonymization modules resulted in the most significant reduction in reidentification. Afni_reface, AFNI's 3Dskullstrip, and FSL's BET had zero matches to the original patient image for SFace. Our study demonstrates the importance of continued vigilance for patient privacy in publicly shared datasets and the importance of evaluation of anonymization methods for CT data. The next step for further validation of the results is radiologist evaluation of the clinical utility of the anonymized CT data.

Disclosures

The authors of this study have no conflicts of interests to declare.

Code and Data Availability

All software code used in this study can be publicly accessed online on the associated anonymization algorithms and facial recognition software's website. The custom algorithm used to change the pixel values of the CT images for postanonymization processing is given in Table S1 in the Supplementary Material. The CT data used in this study were obtained from TCIA and are not publicly available. Data are available upon request and permission from TCIA. The CT datasets used in this study are available with permission in the Low Dose CT Image and Projection Dataset at DOI:10.7937/9NPB-2637, ACRIN-FMISO-Brain Dataset at DOI:10.7937/K9/TCIA.2018.vohlekok, and CPTAC-LSCC Dataset at DOI:10.7937/K9/TCIA.2018.6EMUB5L2. The MRI data used in this study are publicly available in the OpenNeuro database. The T1W-MRI is available in the in-scanner head motion and structural covariance networks dataset at DOI:10.18112/openneuro.ds003639.v1.0.0. The T2W-MRI is available in cortical myelin measured by the T1w/T2w ratio in individuals with depressive disorders and healthy controls dataset at DOI:10.18112/openneuro.ds003653.v1.0.0.

References

1. A. W. Toga, "Neuroimage databases: the good, the bad and the ugly," *Nat. Rev. Neurosci.* **3**(4), 302–309 (2002).
2. F. Budin et al., "Preventing facial recognition when rendering MR images of the head in three dimensions," *Med. Image Anal.* **12**(3), 229–239 (2008).
3. M. Milchenko and D. Marcus, "Obscuring surface anatomy in volumetric imaging data," *Neuroinformatics* **11**(1), 65–75 (2012).
4. N. Schimke, M. Kuehler, and J. Hale, "Preserving privacy in structural neuroimages," *Lect. Notes Comput. Sci.* **6818**, 301–308 (2011).
5. J. J. Chen et al., "Observer success rates for identification of 3D surface reconstructed facial images and implications for patient privacy and security," *Proc. SPIE* **6516**, 65161B (2007).
6. F. W. Prior et al., "Facial recognition from volume-rendered magnetic resonance imaging data," *IEEE Trans. Inf. Technol. Biomed.* **13**(1), 5–9 (2009).
7. C. G. Schwarz et al., "Identification of anonymous MRI research participants with face-recognition software," *New Engl. J. Med.* **381**(17), 1684–1686 (2019).
8. P. Kalavathi and V. B. Prasath, "Methods on skull stripping of MRI head scan images: a review," *J. Digit. Imaging* **29**(3), 365–379 (2015).
9. A. Bischoff-Grethe et al., "A technique for the deidentification of structural brain MR images," *Hum. Brain Mapp.* **28**(9), 892–903 (2007).
10. R. W. Cox, "AFNI: software for analysis and visualization of functional magnetic resonance neuroimages," *Comput. Biomed. Res.* **29**(3), 162–173 (1996).
11. C. G. Schwarz et al., "Changing the face of neuroimaging research: comparing a new MRI de-facing technique with popular alternatives," *NeuroImage* **231**, 117845 (2021).
12. Y. U. Jeong et al., "Deidentification of facial features in magnetic resonance images: software development using deep learning technology," *J. Med. Internet Res.* **22**(12), e22739 (2020).
13. A. E. Theyers et al., "Multisite comparison of MRI defacing software across multiple cohorts," *Front. Psychiatry* **12**, 617997 (2021).
14. D. Provenzano et al., "Unmasking a privacy concern: potential identification of patients in an immobilization mask from 3-dimensional reconstructions of simulation computed tomography," *Pract. Radiat. Oncol.* **12**(2), 120–124 (2022).
15. G. B. Huang et al., "*Labeled faces in the wild: a database for studying face recognition in unconstrained environments*," Tech. Rep. 07-49, University of Massachusetts, Amherst (2007).
16. P. Kinahan et al., "Data from ACRIN-FMISO-Brain," https://vis-www.cs.umass.edu/lfw/lfw.pdf (2018).
17. E. R. Gerstner et al., "ACRIN 6684: assessment of tumor hypoxia in newly diagnosed glioblastoma using 18F-FMISO PET and MRI," *Clin. Cancer Res.* **22**(20), 5079–5086 (2016).

18. E.-M. Ratai et al., "ACRIN 6684: multicenter, phase II assessment of tumor hypoxia in newly diagnosed glioblastoma using magnetic resonance spectroscopy," *PLoS One* **13**(6), e0198548 (2018).
19. National Cancer Institute Clinical Proteomic Tumor Analysis Consortium (CPTAC), "The clinical proteomic tumor analysis consortium lung squamous cell carcinoma collection (CPTAC-LSCC)," The Cancer Imaging Archive (2018).
20. C. McCollough et al., "Low dose CT image and projection data (LDCT-and-Projection-data)," The Cancer Imaging Archive (2020).
21. T. R. Moen et al., "Low-dose CT image and projection dataset," *Med. Phys.* **48**(2), 902–911 (2021).
22. K. Clark et al., "The Cancer Imaging Archive (TCIA): maintaining and operating a public information repository," *J. Digit. Imaging* **26**(6), 1045–1057 (2013).
23. A. Manelis et al., "Cortical myelin measured by the T1w/T2w ratio in individuals with depressive disorders and healthy controls," OpenNeuro [Dataset] (2021).
24. H. R. Pardoe and S. P. Martin, "In-scanner head motion and structural covariance networks," OpenNeuro [Dataset] (2021).
25. A. Fedorov et al., "3D Slicer as an image computing platform for the Quantitative Imaging Network," *Magn. Reson. Imaging* **30**(9), 1323–1341 (2012).
26. 3D Slicer, "3D slicer image computing platform," https://www.slicer.org/ (accessed 9 January 2023).
27. A. Brys, "DICOM2NIfTI GitHub repository," (2016). https://icometrix.github.io/dicom2nifti/ (accessed 18 July 2023).
28. R. W. Cox and P. A. Taylor, "Why de-face when you can re-face?" in *26th Annu. Meet. of the Organization for Hum. Brain Mapp.* (2020).
29. National Institute of Mental Health, "AFNI program: 3DSkullStrip," https://afni.nimh.nih.gov/pub/dist/doc/program_help/3dSkullStrip.html (accessed 25 January 2023).
30. S. Smith et al., "FSL: new tools for functional and structural brain image analysis," *NeuroImage* **13**(6), 249 (2001).
31. F. Alfaro-Almagro et al., "Image processing and quality control for the first 10,000 brain imaging datasets from UK Biobank," *NeuroImage* **166**, 400–424 (2018).
32. S. M. Smith, "Fast robust automated brain extraction," *Hum. Brain Mapp.* **17**(3), 143–155 (2002).
33. J. Muschelli, "Recommendations for processing head CT data," *Front. Neuroinf.* **13**, 61 (2019).
34. O. M. Parkhi, A. Vedaldi, and A. Zisserman, "Deep face recognition," in *Proc. Br. Mach. Vis. Conf.* (2015).
35. F. Schroff, D. Kalenichenko, and J. Philbin, "FaceNet: a unified embedding for face recognition and clustering," in *IEEE Conf. Comput. Vis. and Pattern Recognit. (CVPR)* (2015).
36. D. E. King, "DLIB-ML: a machine learning toolkit," *J. Mach. Learn. Res.* **10**, 1755–1758 (2009).
37. Y. Zhong et al., "SFace: sigmoid-constrained hypersphere loss for robust face recognition," *IEEE Trans. Image Process.* **30**, 2587–2598 (2021).
38. S. I. Serengil and A. Ozpinar, "LightFace: a hybrid deep face recognition framework," in *Innov. Intell. Syst. and Appl. Conf. (ASYU)* (2020).
39. B. L. Filkins et al., "Privacy and security in the era of digital health: what should translational researchers know and do about it?" *Am. J. Transl. Res.* **8**(3), 1560–1580 (2016).

Rahil Patel is a medical student at George Washington University School of Medicine and Health Sciences. His research is focused on the applicability of machine learning to medicine.

Destie Provenzano is a PhD student who specializes in machine learning for medical imaging at Loew Laboratory of George Washington University School of Engineering and Applied Science.

Murray Loew is the principal investigator of the Medical Imaging and Image Analysis Laboratory at George Washington University School of Engineering and Applied Science.