Demystifying the Risk of Reidentification in Neuroimaging Data

A Technical and Regulatory Analysis

Anita S. Jwa,[1] Oluwasanmi Koyejo[2] & Russell A. Poldrack[1]

[1]Department of Psychology, Stanford University

[2] Department of Computer Science, Stanford University

I have no disclosures

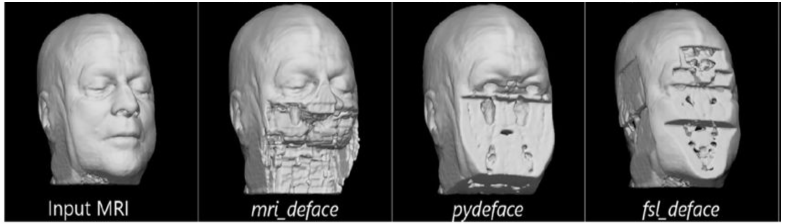
Introduction

- Sharing research data has been widely promoted in the field of neuroimaging and has enhanced the rigor and reproducibility of neuroimaging studies.[1]
- Emergence of novel software tools and algorithms, such as face recognition, has raised concerns due to their potential to reidentify neuroimaging data that are thought to have been deidentified.[2]
- However, the risk of reidentification via these tools and algorithms has not yet been examined outside the limited settings for demonstration purposes.
- In this study, we will examine 1) the likelihood of reidentification via face recognition in real-world settings and 2) the regulatory implications of this risk by taking US jurisdiction as a case study.

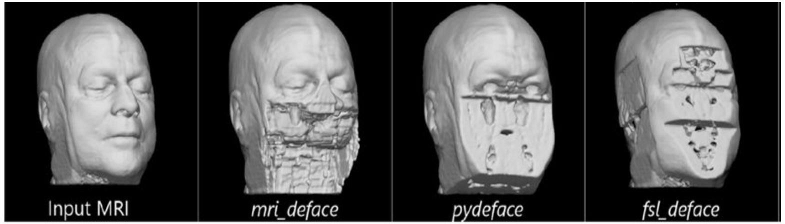
[1] Poldrack & Gorgolewski, 2014. *Nat Neurosci*.

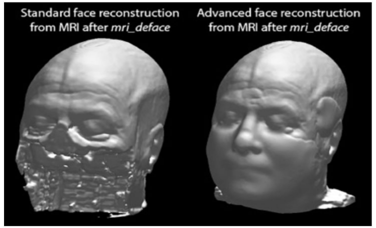
[2] Schwarz et al., 2021. *Neuroimage*.

Previous Study: Schwarz et al. (2021)



Previous Study: Schwarz et al. (2021)

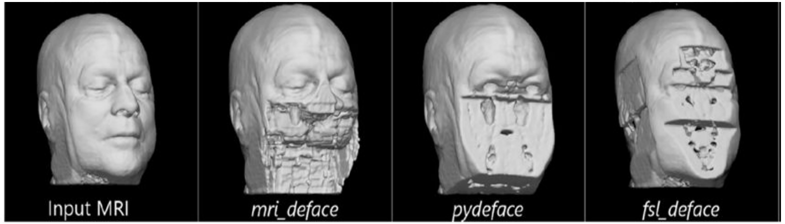


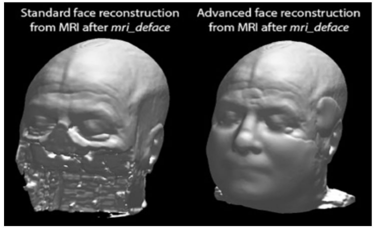

Standard face reconstruction from MRI after mri_deface
Advanced face reconstruction from MRI after mri_deface

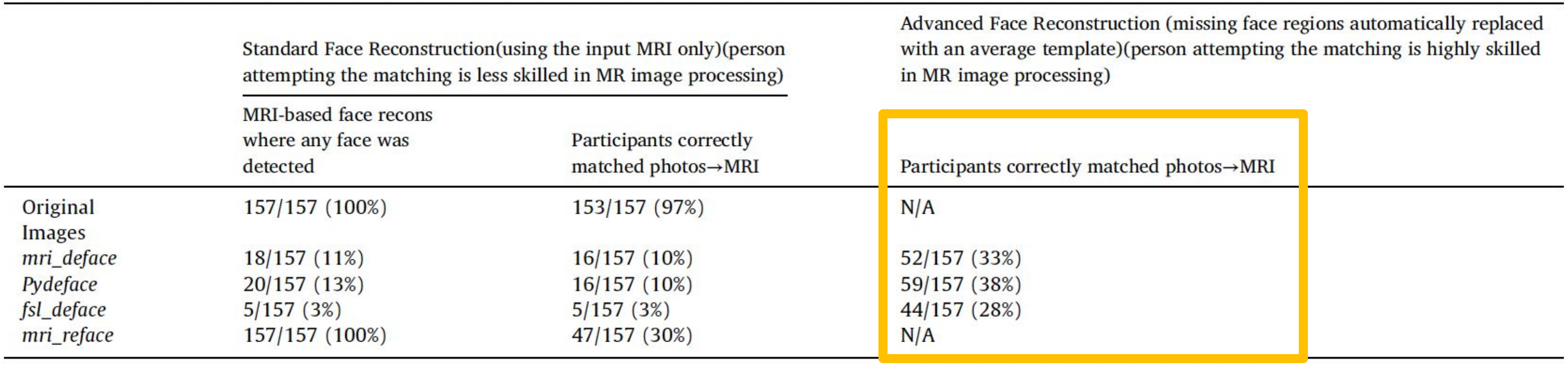
	Standard Face Reconstruction(using the input MRI only)(person attempting the matching is less skilled in MR image processing)		Advanced Face Reconstruction (missing face regions automatically replaced with an average template)(person attempting the matching is highly skilled in MR image processing)
	MRI-based face recons where any face was detected	Participants correctly matched photos→MRI	Participants correctly matched photos→MRI
Original Images	157/157 (100%)	153/157 (97%)	N/A
mri_deface	18/157 (11%)	16/157 (10%)	52/157 (33%)
Pydeface	20/157 (13%)	16/157 (10%)	59/157 (38%)
fsl_deface	5/157 (3%)	5/157 (3%)	44/157 (28%)
mri_reface	157/157 (100%)	47/157 (30%)	N/A

(Schwarz et al., 2021. *Neuroimage*)

Previous Study: Schwarz et al. (2021)




Standard face reconstruction from MRI after *mri_deface*
Advanced face reconstruction from MRI after *mri_deface*



	Standard Face Reconstruction(using the input MRI only)(person attempting the matching is less skilled in MR image processing)		Advanced Face Reconstruction (missing face regions automatically replaced with an average template)(person attempting the matching is highly skilled in MR image processing)
	MRI-based face recons where any face was detected	Participants correctly matched photos→MRI	Participants correctly matched photos→MRI
Original Images	157/157 (100%)	153/157 (97%)	N/A
mri_deface	18/157 (11%)	16/157 (10%)	52/157 (33%)
Pydeface	20/157 (13%)	16/157 (10%)	59/157 (38%)
fsl_deface	5/157 (3%)	5/157 (3%)	44/157 (28%)
mri_reface	157/157 (100%)	47/157 (30%)	N/A

(Schwarz et al., 2021. *Neuroimage*)

Methods (1)

- The reidentification problem is a multi-class classification problem, wherein each individual represents a separate class
- Classification accuracy is necessarily tied to the number of classes being distinguished.
- The pool of potential matches is much larger than the size of samples in schwarz et al.'s study (N = 157).

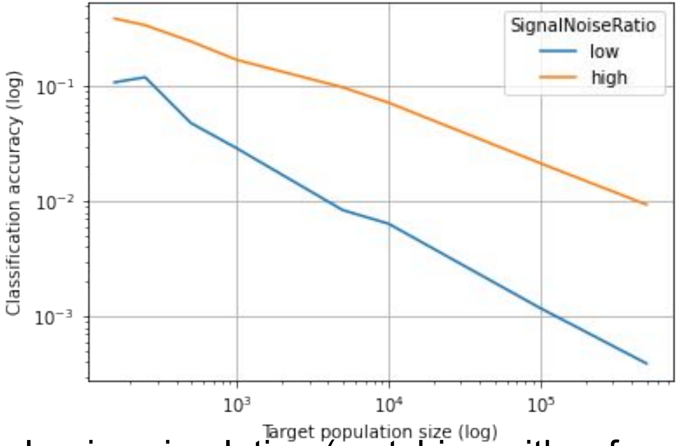
Methods (1)

- To test the generalizability of these reported accuracies in real-world situations, we designed a classification problem using simplified data.
    - Test data are generated from a normal distribution by adding random noise to each individual calibrated to provide two target levels of reidentification performance (10% for defaced images (using pydeface) and 38% for refaced images).
    - Matching accuracies are assessed as the population size varied from 157 to a size large enough to be realistic for a population of potential research subjects (500,000).

Methods (2)

- We further examined whether this novel risk of reidentification would place defaced neuroimaging data out of compliance with existing standards for data deidentification provided in relevant US regulations.

Results: Simulation Analysis



- For the higher signal-noise simulation (matching with refaced images), identification accuracy dropped from **38.2%** for the initial population size of 157 to **0.9%** at a population size of 500,000.
- For the lower signal-noise simulation (matching with defaced images), identification accuracy dropped from **10.8%** for the initial population size of 157 to **0.04%** at a population size of 500,000.
- The relationship between accuracy and population size is roughly linear in log-log space, consistent with theoretical results.

Results: Regulatory Analysis (1)

Results: Regulatory Analysis (1)

- **Common Rule**
  - US federal regulations that provide the standards for ethical conduct of biomedical and behavioral research involving human subjects

Results: Regulatory Analysis (1)

- **Common Rule**
  - Secondary research on identifiable private information is subject to Common Rule requirements (informed consent & IRB review)
    - ✓ Private information for which the identity of the subject is or may <u>readily be ascertained</u> by the investigator or associated with the information (45 CFR §46.102 (e)(5)).
    - ✓ Common Rule does not define the term “readily ascertainable,” and it is left to individual IRB’s discretion to interpret and apply this standard.

Results: Regulatory Analysis (1)

- **Common Rule**
  - Research involving shared data that are not individually identifiable falls outside the scope of human subject research
  - The Rule does *not* provide specific methods to make private information not individually identifiable
    - ✔ Office for Human Subject Research(OHRP) guidance(2008): Data are not individually identifiable
      - If identifying information has been coded and
      - The investigators and the holder of the key to the code enter into an agreement prohibiting the release of the key under any circumstances.

Results: Regulatory Analysis (1)

- **Common Rule**
  - The OHRP guidance would not be applicable
    - ✔ Reidentification via face recognition would directly link neuroimaging data back to individual subjects without needing to decipher the code.
  - The core question here is whether the identity of the subject is readily ascertainable.
    - ✔ "[T]he Rule's bar for rendering data nonidentifiable is fairly low."[3]
  - It would be difficult to argue that the minimal real-world likelihood of reidentification via face recognition would make the identities of subjects readily ascertainable.

[3] Meyer, 2020. *J. L. Med. & Ethics.*

Results: Regulatory Analysis (2)

- **Health Insurance Portability and Accountability Act of 1996 (HIPAA)**
  - The HIPAA Privacy Rule addresses the required and permitted use/disclosure of protected health information (PHI) held or transmitted by a covered entity

Results: Regulatory Analysis (2)

- **Health Insurance Portability and Accountability Act of 1996 (HIPAA)**
  - Data deidentification standards under HIPAA
    - ✔ Expert determination (45 CFR §164.514(b)(1))
      - Formal determination by a qualified expert that there exists a very small risk that the information could be used to identify the individual to whom the the information pertains
      - No explicit numerical level of risk that is deemed to meet the “very small” level

  - The minimal likelihood of reidentification posed by face recognition would not render defaced neuroimaging data noncompliant under the expert determination

Results: Regulatory Analysis (2)

- **Health Insurance Portability and Accountability Act of 1996 (HIPAA)**
  - Data deidentification standards under HIPAA
    - ✔ Safe Harbor (45 CFR §164.514(b)(2))
      - Removal of 18 unique identifiers from PHI ***and***
      - A covered entity *does not have actual knowledge* that the information could be used alone or in combination with other information to identify the subject

  - Whether being aware of recent studies on face recognition could constitute having actual knowledge that defaced neuroimaging data could still be used to reidentify data subjects
  - Mere knowledge of studies about methods to reidentify health information does not necessarily count as actual knowledge under the safe harbor standard (OCR Guidance, 2012)

Discussion

- The results of our study suggest a more *balanced view* of the real-world likelihood of reidentification in neuroimaging data when weighed against the benefits of data sharing and open science practice.
- Any limitations on sharing and secondary use of data should be carefully calibrated, corresponding to the specific risk associated with individual data sets to avoid chilling effects on open sharing of neuroimaging data.
- Best practices for sharing human neuroimaging data would be needed to better inform researchers of the standards and due diligence for sharing their data.
- Data repositories should also equip themselves with adequate privacy measures for sharing of human neuroimaging data.
- Future development in technical countermeasures to the novel privacy attack would further aid open sharing of neuroimaging data.

Thank you!

RESEARCH FUNDED BY

NIH 5R24MH117179

STANFORD HUMAN-CENTERED ARTIFICIAL INTELLIGENCE (HAI)