Digital Health and Informatics Innovations for Sustainable Health Care Systems
J. Mantas et al. (Eds.)
© 2024 The Authors.
This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0).
doi:10.3233/SHTI240493

Quality Assessment of Brain MRI Defacing Using Machine Learning

Sina SADEGHI[a,b,1], Maryam KHODAEI[a,b], Lars HEMPEL[a,b,c] and Toralf KIRSTEN[a,b,c]
[a] *Department for Medical Data Science, Leipzig University Medical Center, Leipzig, Germany*
[b] *Institute for Medical Informatics, Statistics and Epidemiology, Leipzig University, Leipzig, Germany*
[c] *Faculty Applied Computer and Bio Sciences, Mittweida University of Applied Sciences, Mittweida, Germany*
ORCiD ID: Sina Sadeghi https://orcid.org/0000-0001-9914-1775, Toralf Kirsten https://orcid.org/0000-0001-7117-4268

Abstract. Defacing of brain magnetic resonance imaging (MRI) scans is a crucial process in medical imaging research aimed at preserving patient privacy while maintaining data integrity. However, existing defacing algorithms are prone to errors, potentially compromising patient anonymity. This paper investigates the feasibility and efficacy of automated quality assessment for defaced brain MRIs using machine learning (ML). Our findings demonstrate the promising capability of ML models in accurately distinguishing between properly and inadequately defaced MRI scans.

Keywords. MRI, Patient Privacy, Defacing, Machine Learning

1. Introduction

The sharing of medical imaging data, including structural magnetic resonance imaging (MRI) of the brain, is essential for advancing scientific knowledge and improving patient care. However, this practice presents inherent privacy risks, especially concerning facial recognition of patients, which violates data protection regulations such as the General Data Protection Regulation (GDPR). With advancements in scan quality and facial recognition software, the risk of re-identifying patients through 3D rendering of MRI scans has escalated significantly. This highlights the limitations of conventional anonymization methods that solely focus on removing sensitive metadata.

Defacing algorithms offer a potential solution to mitigate these risks by eliminating identifiable facial features while preserving diagnostic information. Nonetheless, these algorithms are not infallible and may occasionally leave detectable facial features [1]. Thus, this study aims to develop machine learning (ML) models capable of automatically assessing the efficacy of defacing algorithms integrated into the image processing pipeline.

[1] Corresponding Author: Sina Sadeghi; E-mail: sina.sadeghi@medizin.uni-leipzig.de.

2. Methods

Over 200 MRI scans from the Leukodystrophy registry of the clinic and polyclinic for neurology at the University of Leipzig Medical Center were utilized in this study, and three commonly used defacing programs, namely PyDeface, MRI-Deface, and QuickShare, were employed to remove facial features from the MRI scans. The original and defaced images were rendered in 3D and then converted to 2D for comparison. The defaced MRI scans were manually labeled as either properly or inadequately defaced using a graphical user interface (GUI).

ML models were developed to evaluate the output quality of defacing algorithms by generating several features, including the number and proportion of altered head/brain voxels, and brain-to-head ratio in 3D images, as well as peak signal to noise ratio, root mean squared error, and structural and feature similarity measures in 2D scans before and after defacing [2]. The dataset was divided into 80/20\% training/test sets, and experiments were conducted 100 times using Monte Carlo sampling with random splits.

3. Results

The results of the ML models, summarized in Table 1, demonstrate promising performance in distinguishing between properly and inadequately defaced MRI scans.

Table 1. ML models performance

Model	Accuracy	Specificity	Sensitivity	Precision	AUROC
Bagging	0.76	0.73	0.79	0.76	0.85
Random Forest	0.79	0.80	0.78	0.81	0.87
Gradient Boost	0.78	0.81	0.75	0.80	0.86
XGBoost	0.78	0.78	0.77	0.79	0.86

4. Discussion and Conclusions

This study has developed ML models for automatically evaluating the effectiveness of defacing algorithms, crucial for safeguarding patient privacy in medical imaging research. The models performance is comparable to that of deep learning models trained on much larger datasets [3]. Integrating such models into the image processing pipeline can enhance the security and compliance of data sharing practices, thereby promoting responsible and ethical research endeavors. Future work will focus on refining the models and expanding the dataset to improve performance and generalizability.

References

[1] Theyers et al, Multisite comparison of MRI defacing software across multiple cohorts. Frontiers in Psychiatry. 2021;12, doi: 10.3389/fpsyt.2021.617997.

[2] Sara et al, Image Quality Assessment through FSIM, SSIM, MSE and PSNR—A Comparative Study. Journal of Computer and Communications. 2019;7(3), doi: 10.4236/jcc.2019.73002.

[3] Delbarre et al, Application of a convolutional neural network to the quality control of MRI defacing. Computers in Biology and Medicine. 2022;151(A):106211, doi: 10.1016/j.compbiomed.2022.106211.