



DEEPHEALTH

Hackathon - Course 4

De-identification of Biomedical Images



The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



DEEPHEALTH

Index

1. Introduction
2. Facial information contained in medical images
3. Applications
4. Image annotation with text
5. Text removal methods





DEEPHEALTH

Introduction



The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



Introduction

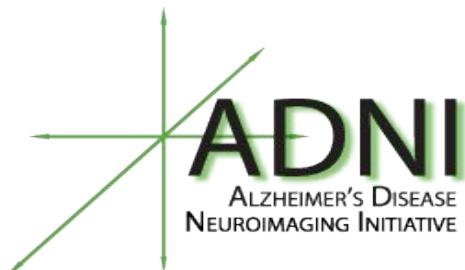
Datasets



Can find many image dataset on internet

A lot of them:

- Images obtained within a study
- With high control of acquisition protocols
- Cleaned dataset
- Labeled
- High quality resolution





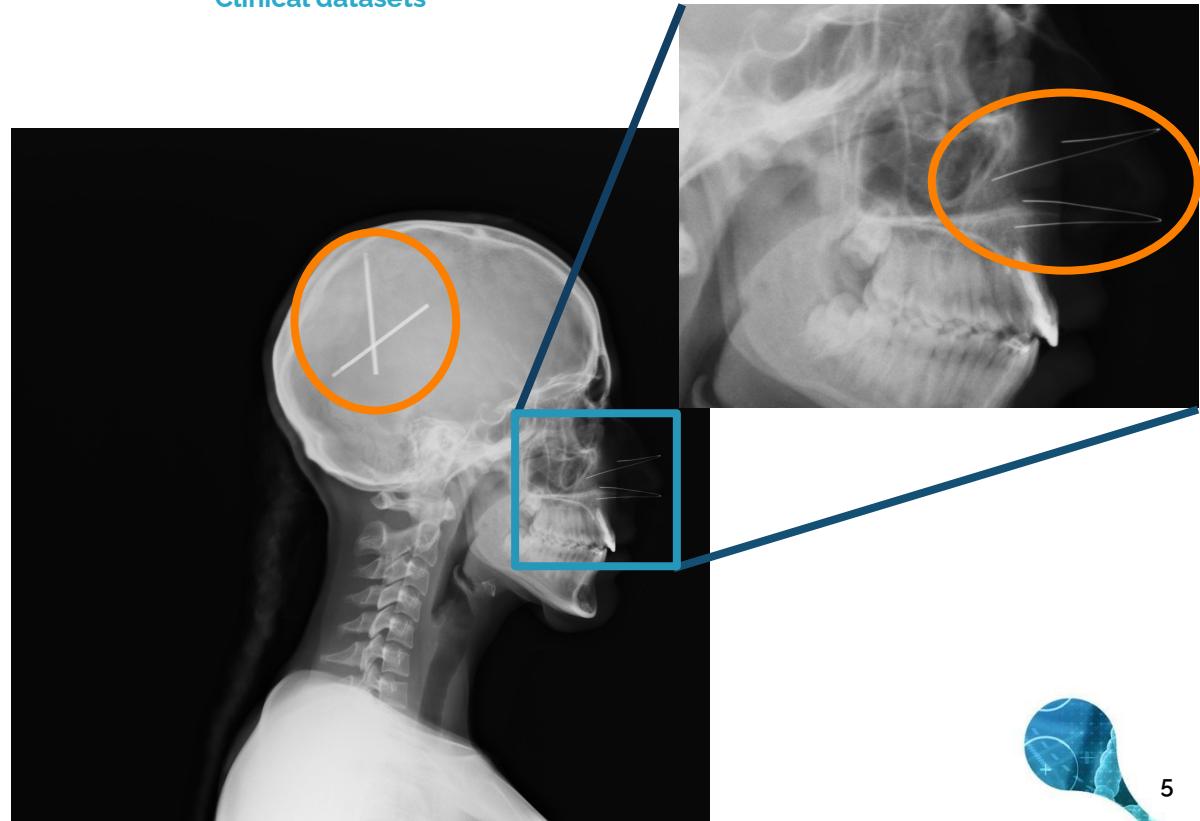
Introduction

Clinical datasets



We obtain a clinical dataset:

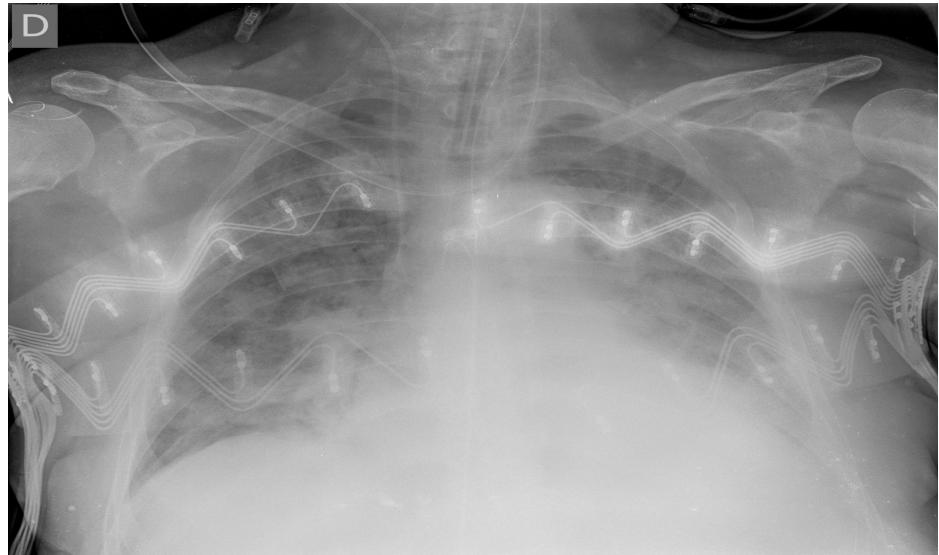
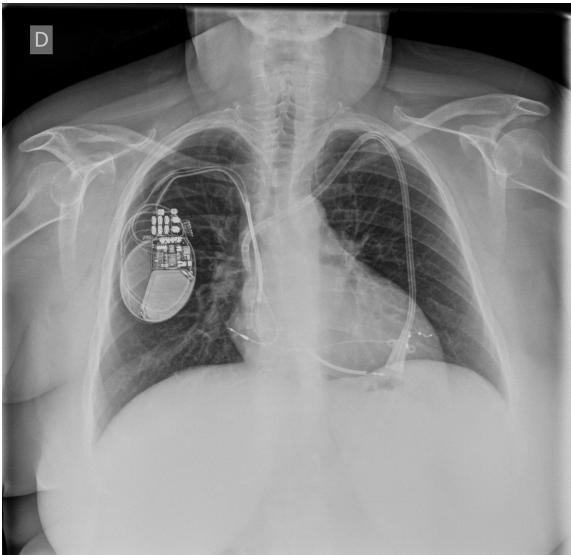
- Raw images
- With noises
- With sensitive information
- Many distinct protocols





Introduction

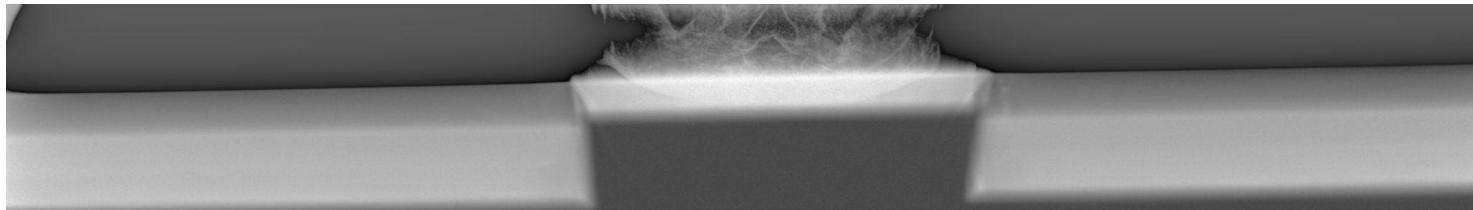
Artifacts





Introduction

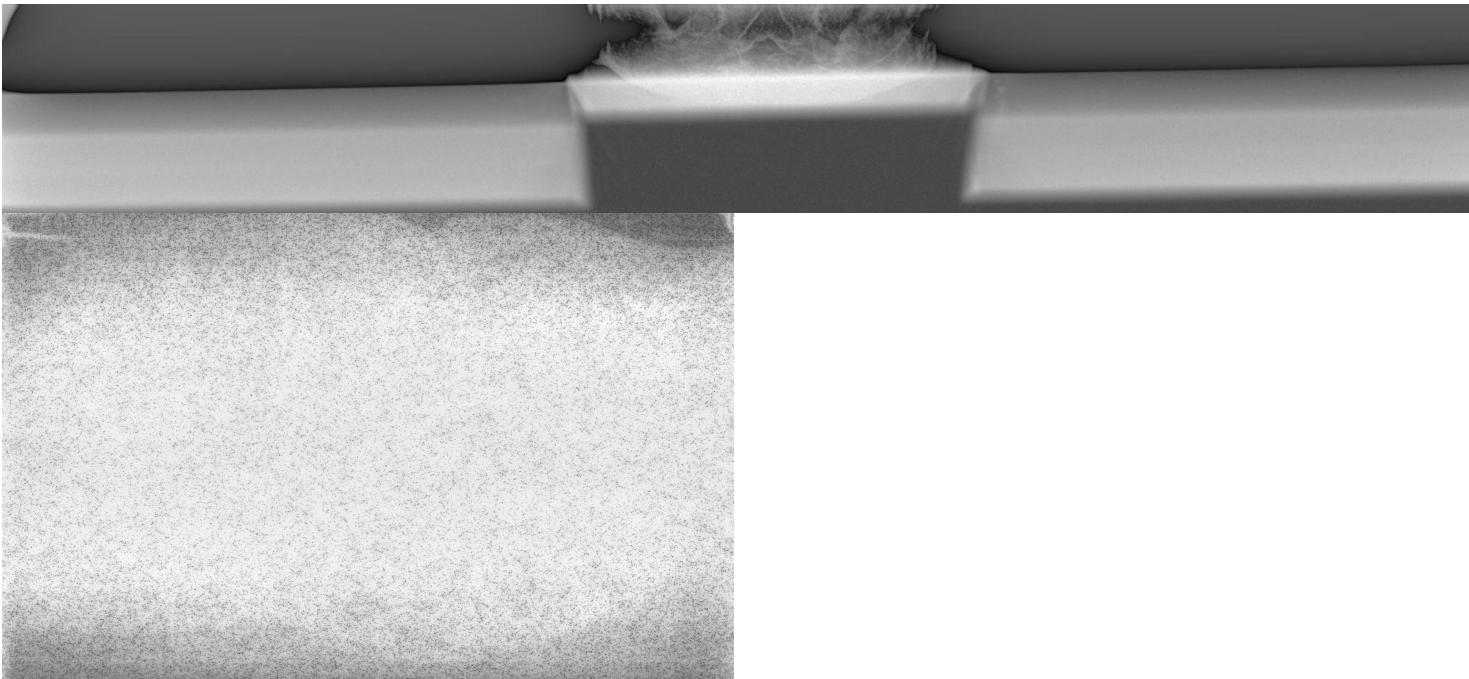
unable images





Introduction

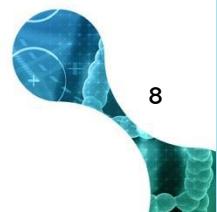
unable images



8



The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.

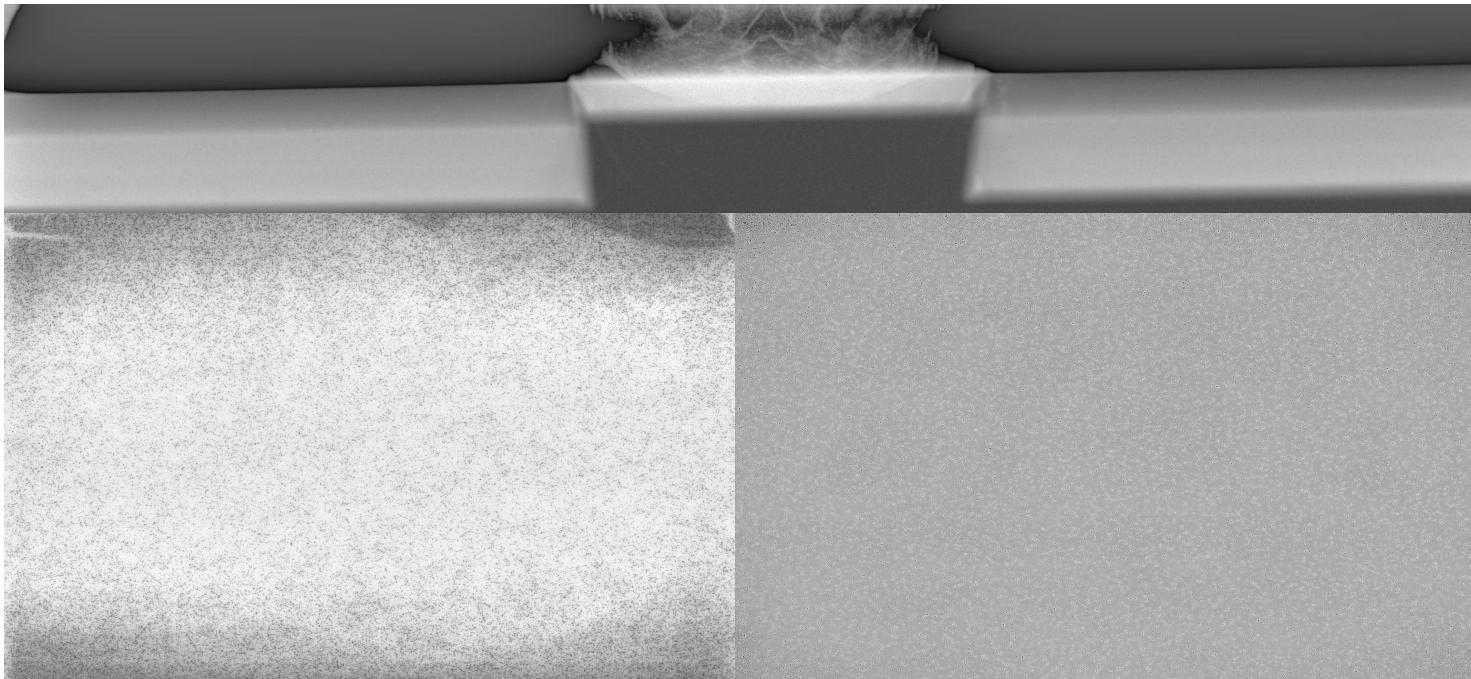




DEEPHEALTH

Introduction

unable images



9

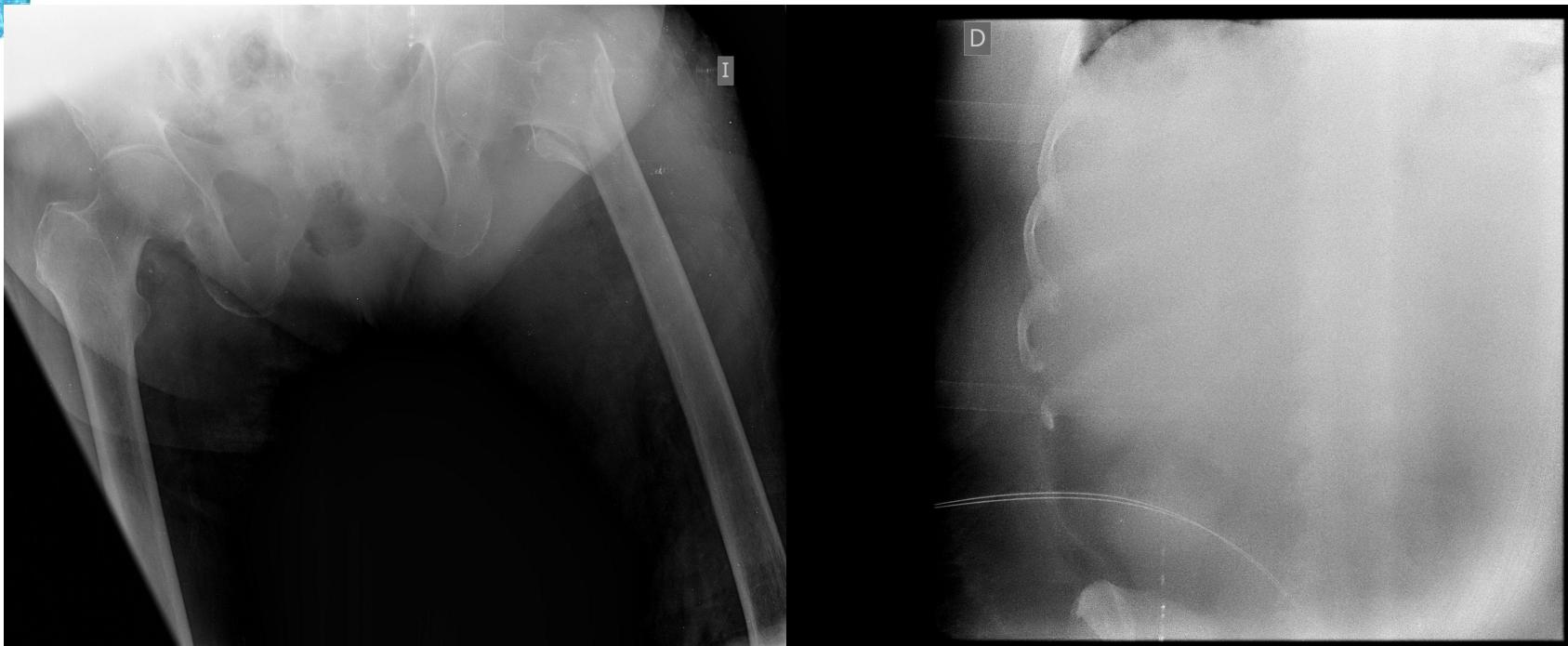


The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



Introduction

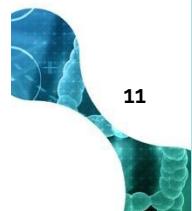
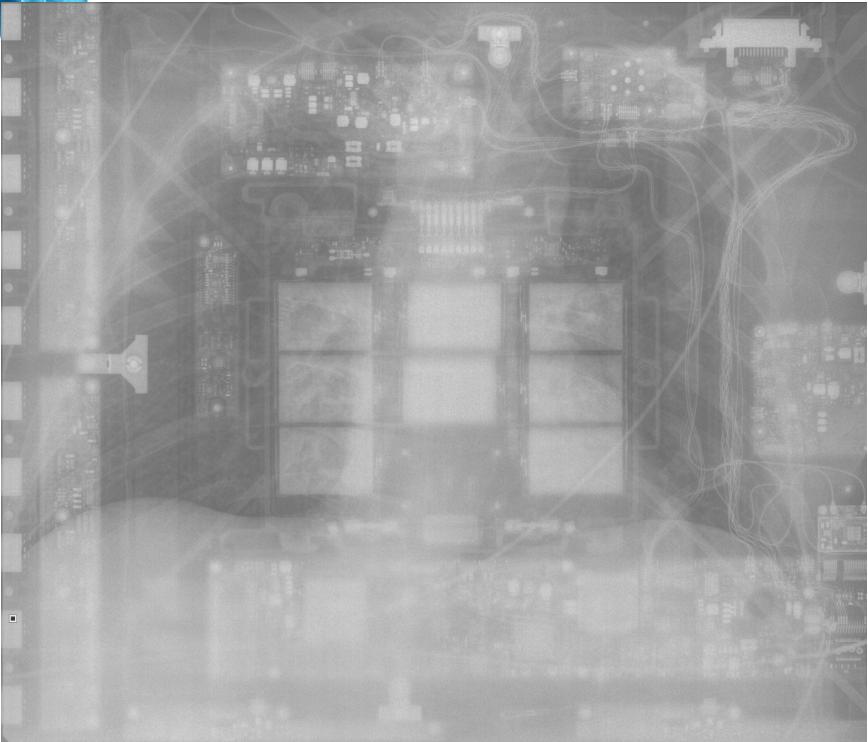
unable images





Introduction

unable images





Introduction

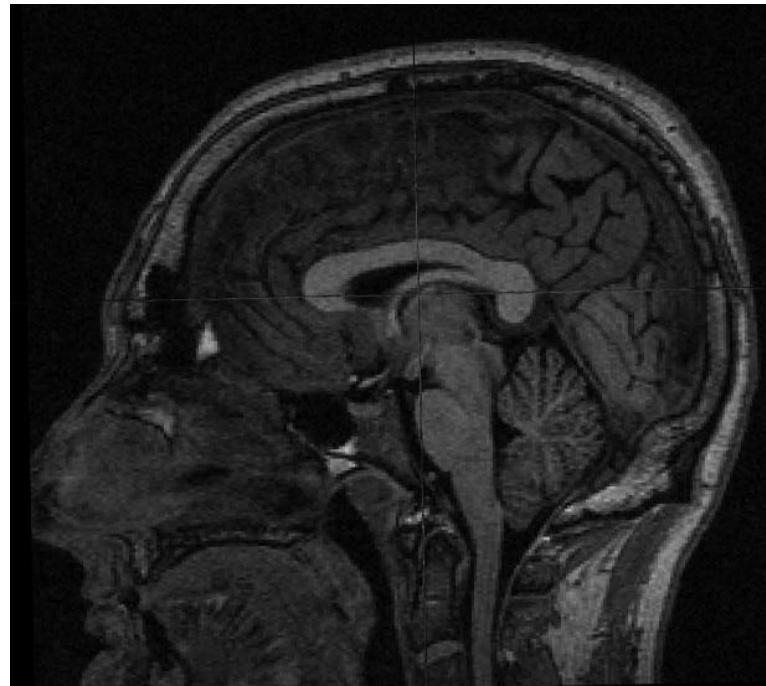
Datasets



Therefore these datasets require a thorough cleaning of the images that are not useful for the study.

But this is not the only problem...

Do patients have to be asked for permission to use their data?





Introduction

Datasets



The last dataset we set up on covid19+ contains +2000 subjects.

Asking for consent from so many patients would be impossible.

If the data is anonymised correctly, an exception from informed consent can be requested for one dataset.



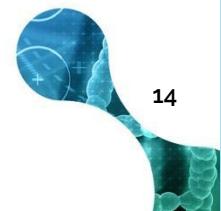
Introduction

Datasets

How much can we say that data are anonymised?

There has been talks of anonymisation of metadata and radiology reports

Can they already be shared?





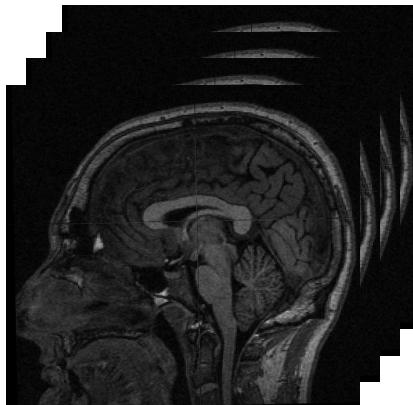
DEEPHEALTH

Facial information contained in medical images



Facial information

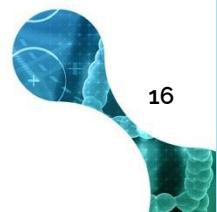
3D Images



16



The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



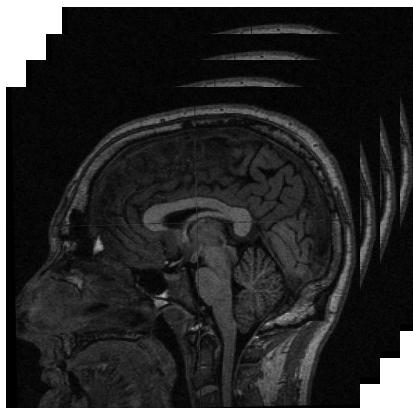
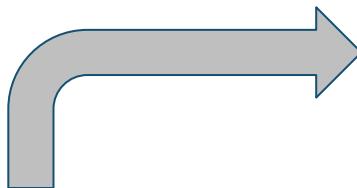


DEEPHEALTH

Facial information

3D Images

Reconstruction

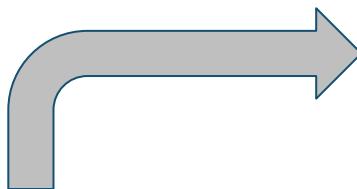




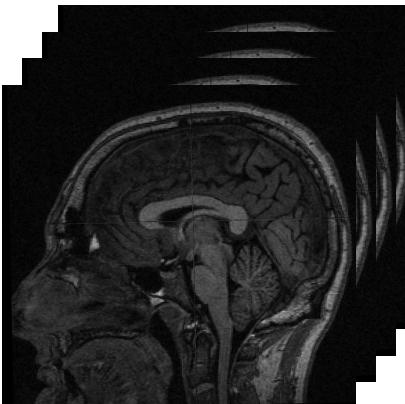
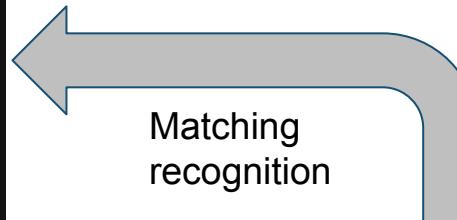
Facial information

3D Images

Reconstruction



Matching
recognition

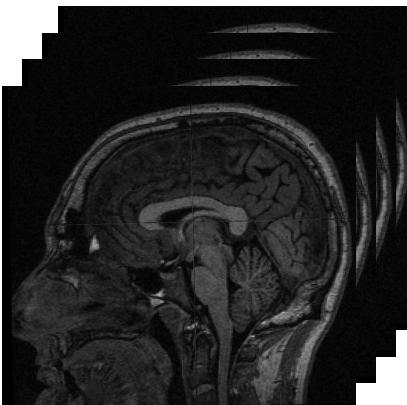




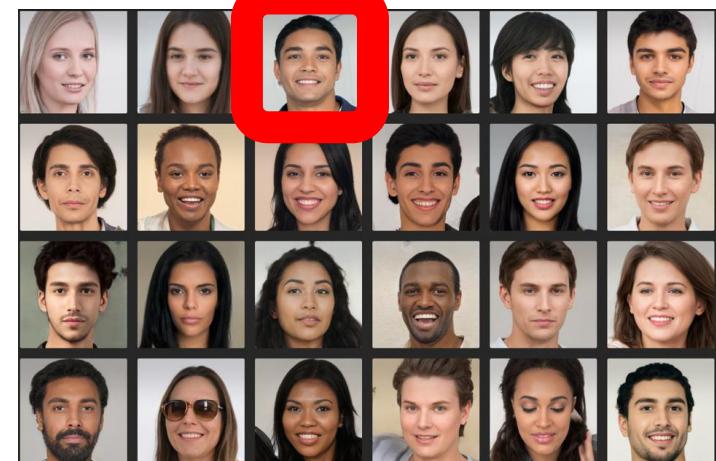
Facial information

3D Images

Reconstruction



Matching
recognition



19



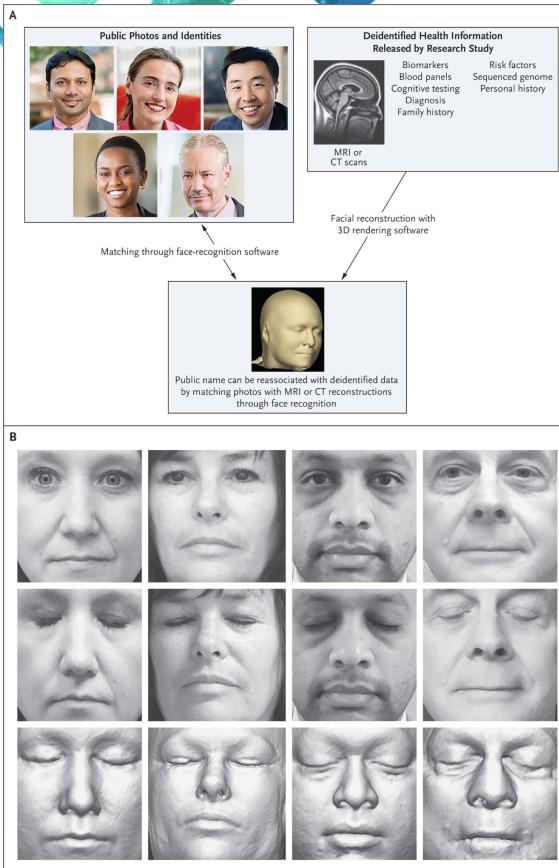
The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



DEEPHEALTH

Facial information

3D Images



In the article: [Identification of Anonymous MRI Research Participants with Face-Recognition Software](#)
(October 24, 2019)

previous studies	Accuracy
Human visual raters	0.4
automated face-recognition software developed in 2008	0.275

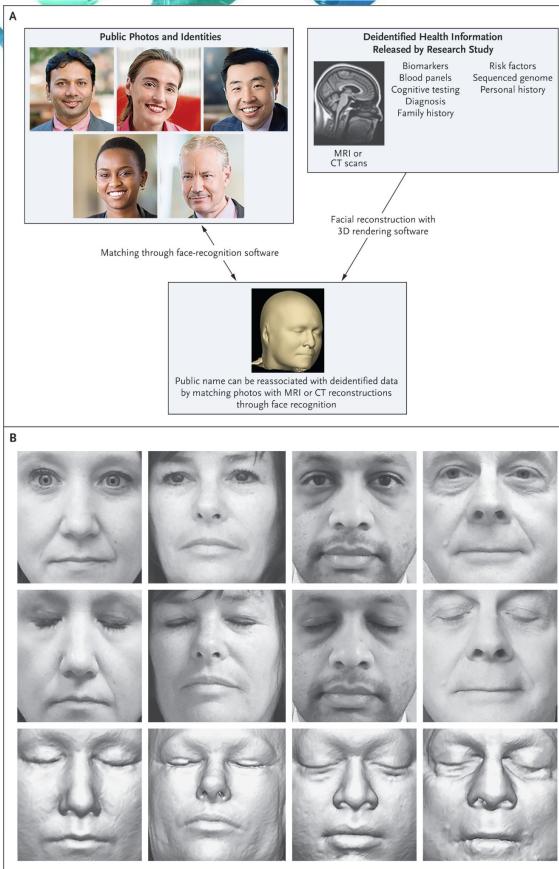




DEEPHEALTH

Facial information

3D Images



In the article: [Identification of Anonymous MRI Research Participants with Face-Recognition Software](#)
(October 24, 2019)

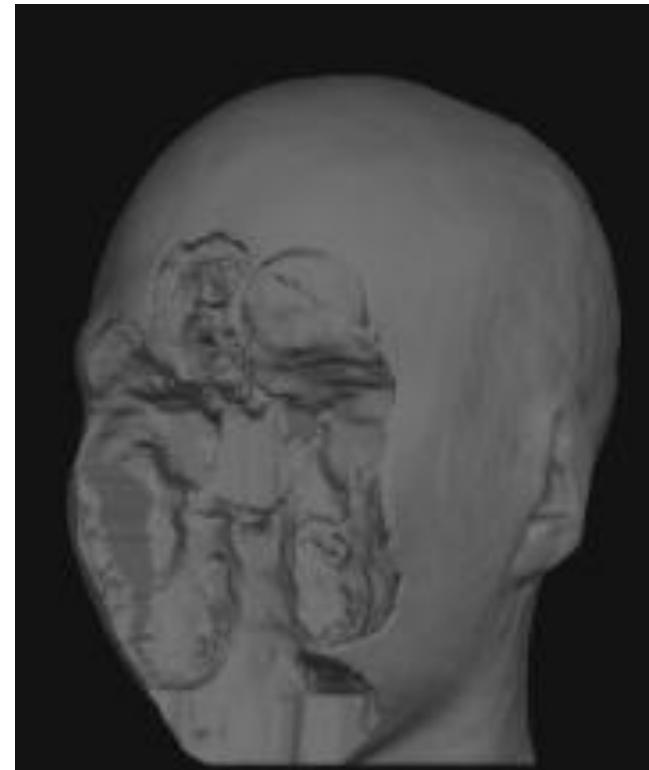
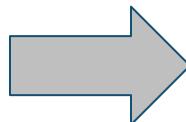
	Top 1 Accuracy	Top 5 Accuracy
automated face-recognition software (Microsoft Azure)	0.83	0.95





Facial information

Objective





DEEPHEALTH

Applications



The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



Software



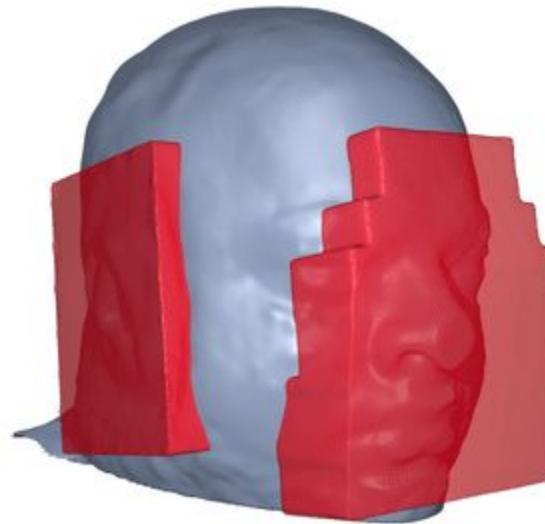
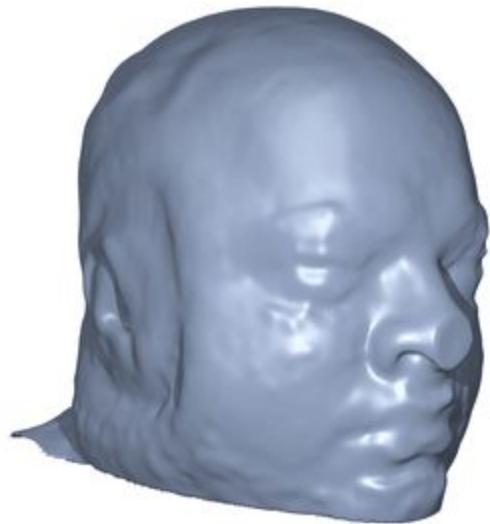
- [MRI Deface](#): Software implemented in Freesurfer. Can be used separately.
- [Face Masking](#): Software implemented in XNAT. Can be used separately.
- [Defacer](#): De-Identification of Facial Features in Magnetic Resonance Images: Software Development Using Deep Learning Technology





Software

Our method: MRI Deface





Software

Our method: MRI Deface



Native image



Reference image



Reference mask

0.58	0.60	0.33
0.95	-0.62	0.67
-0.44	0.47	-0.56
1.00	1.00	1.00

Affinity matrix

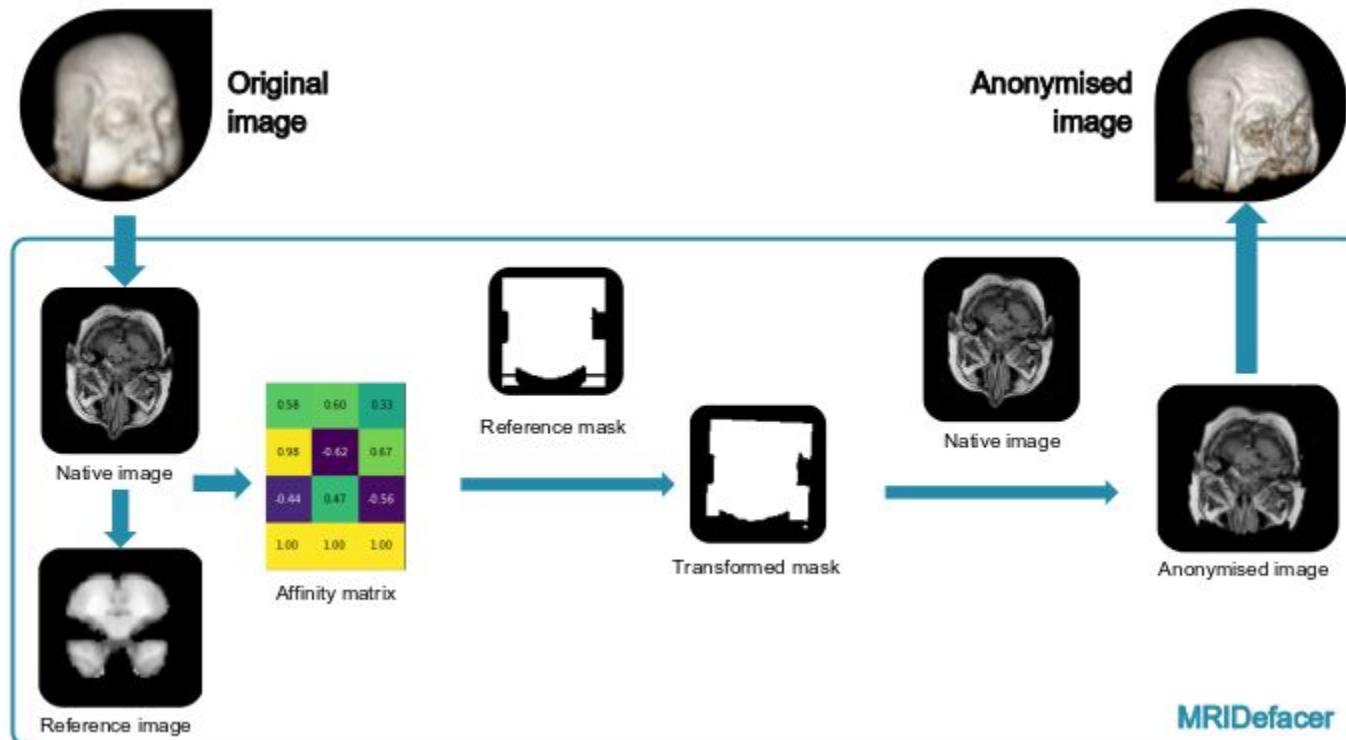




DEEPHEALTH

Software

Our method: MRI Deface

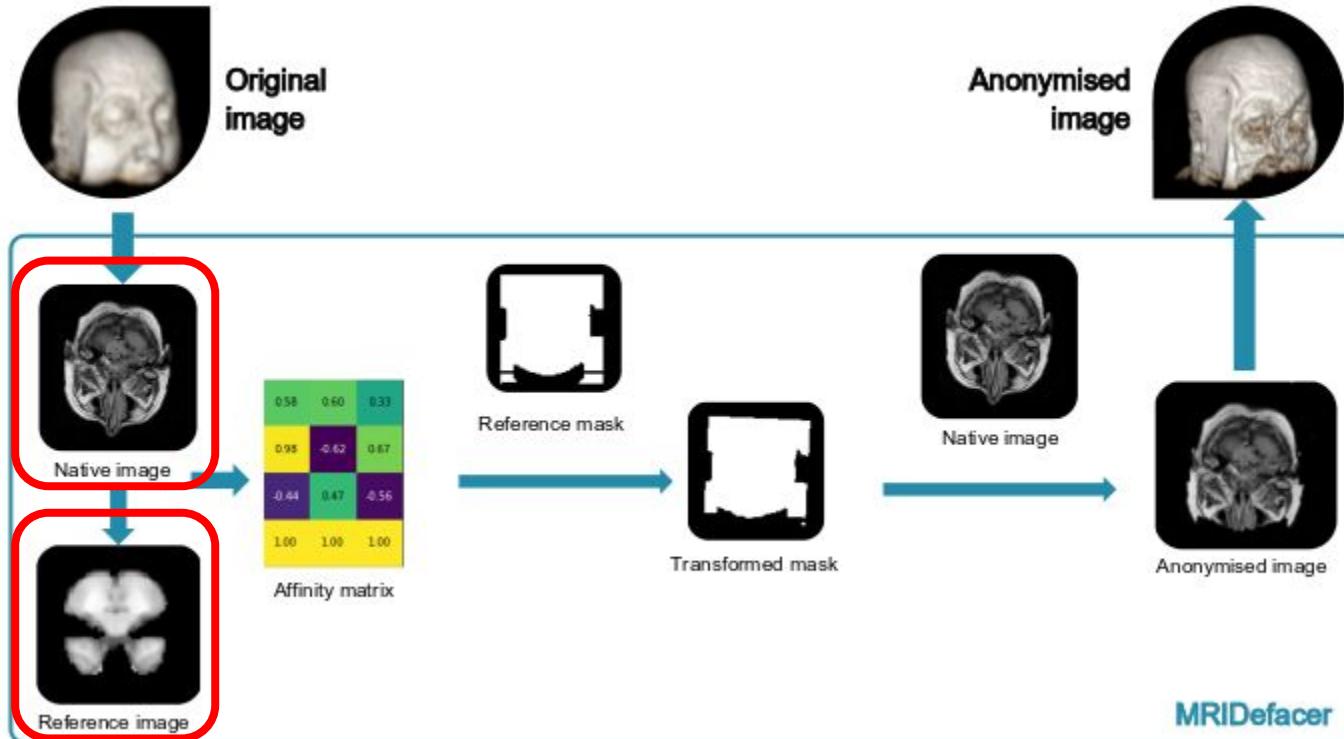




DEEPHEALTH

Software

Our method: MRI Deface



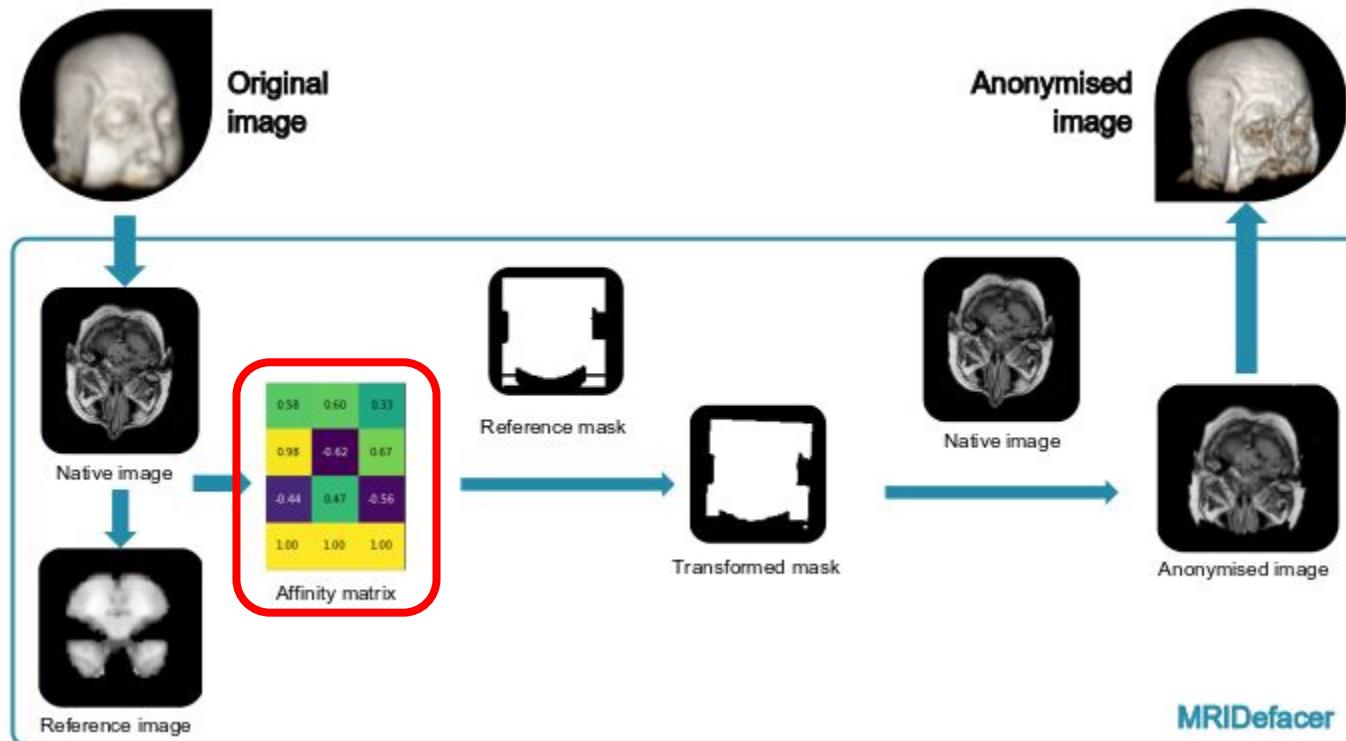
The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



DEEPHEALTH

Software

Our method: MRI Deface

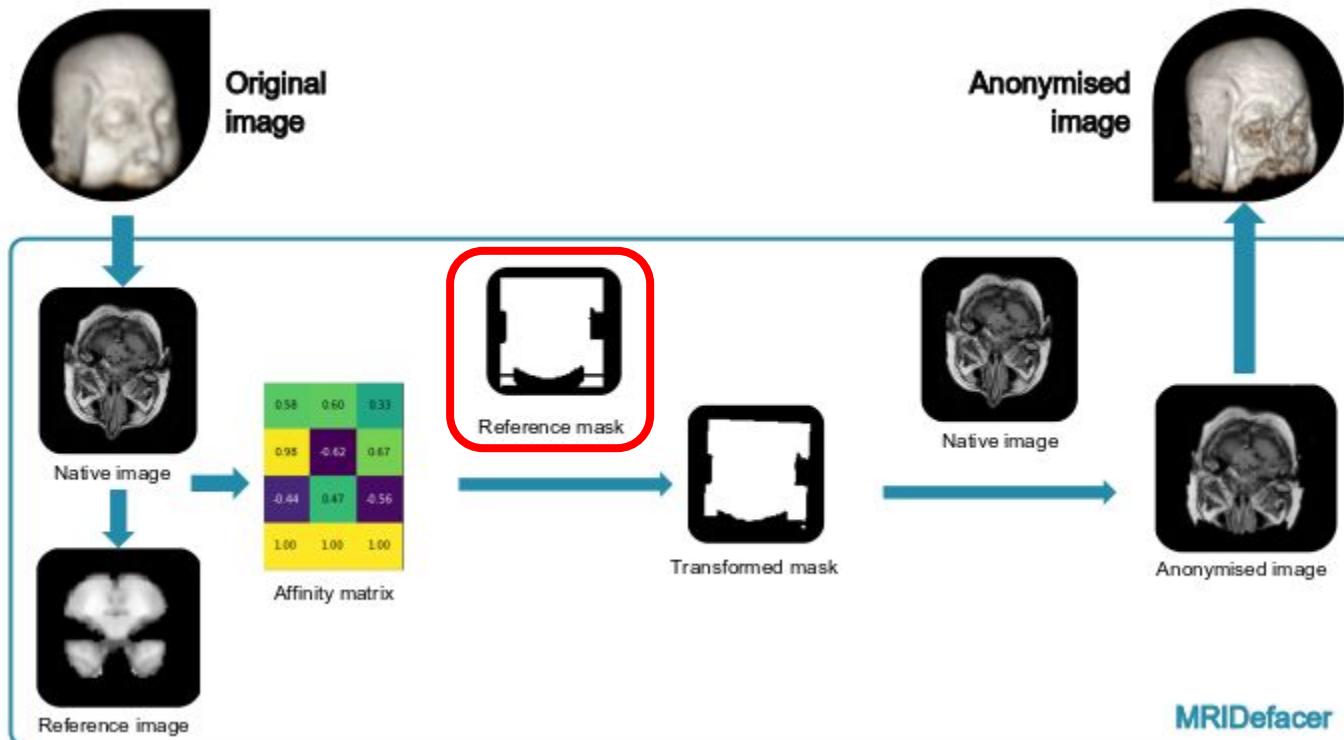




DEEPHEALTH

Software

Our method: MRI Deface

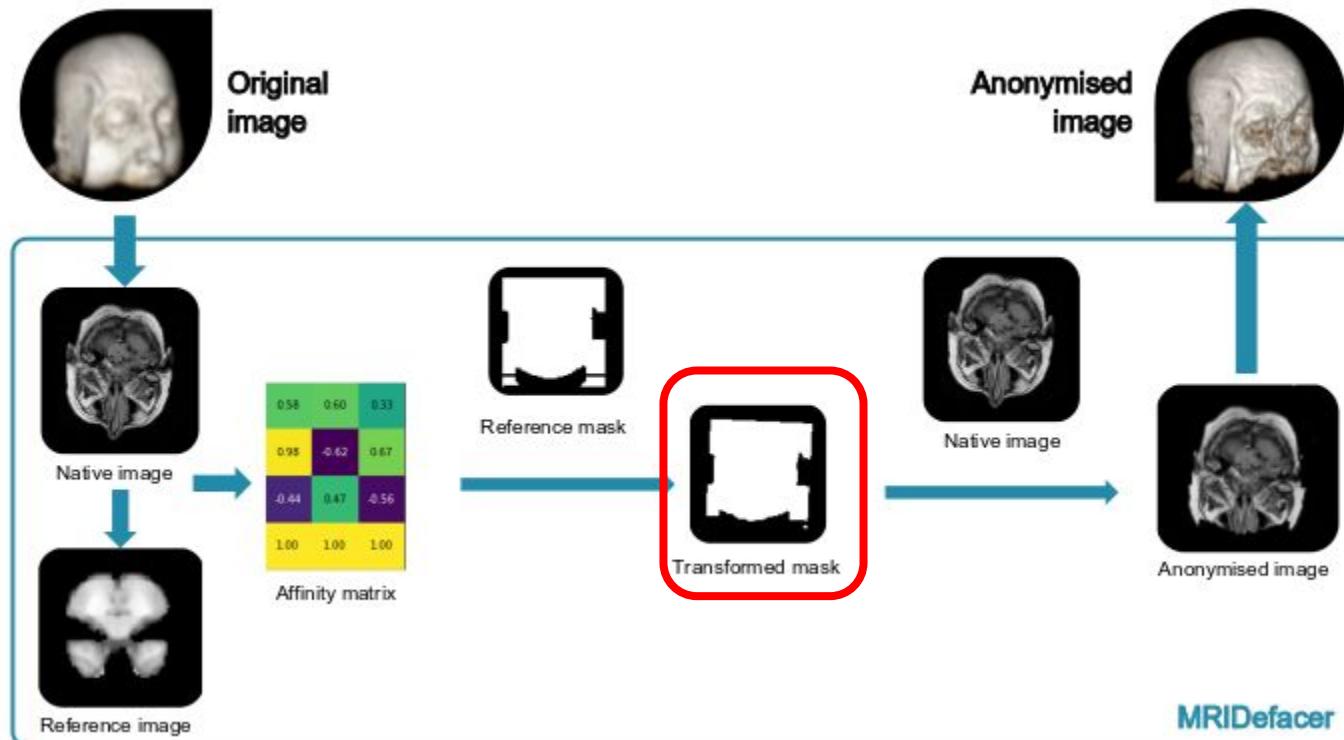




DEEPHEALTH

Software

Our method: MRI Deface

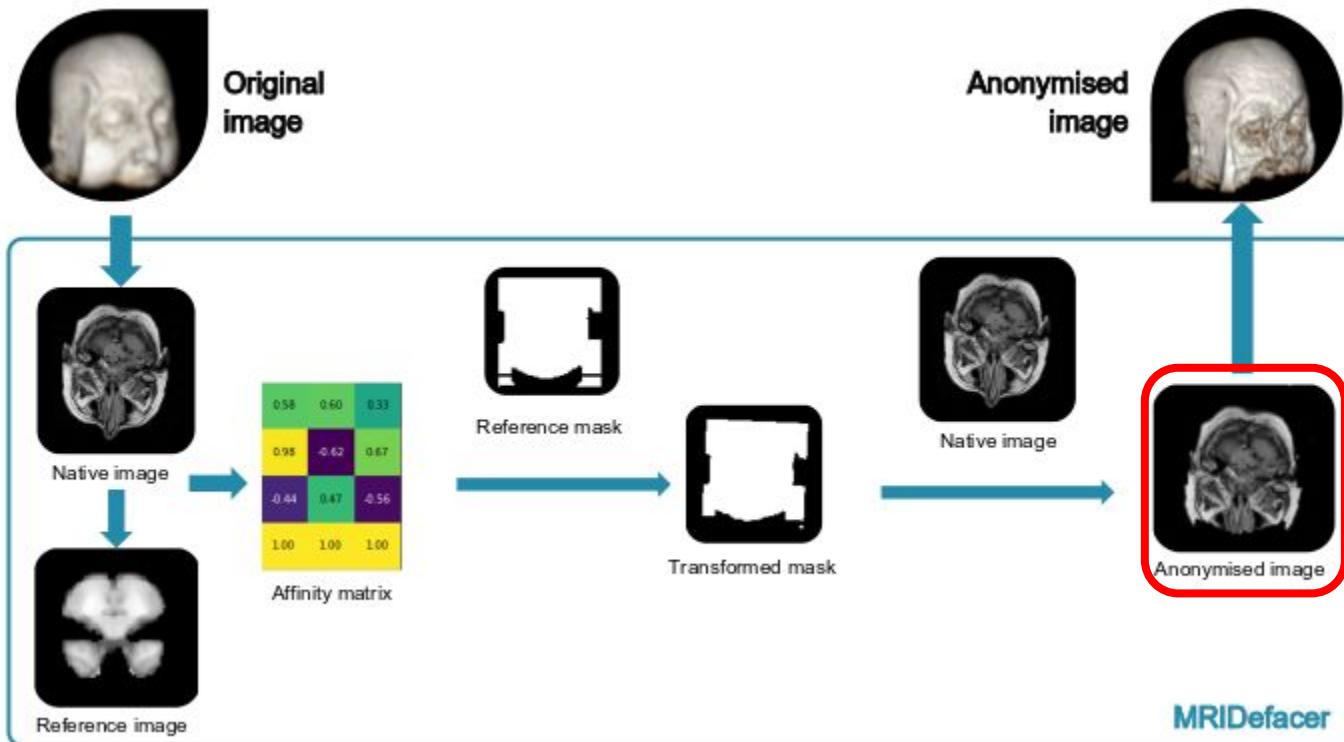




DEEPHEALTH

Software

Our method: MRI Deface



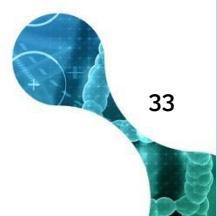


Software

Our method: MRI Deface



- Advantages:
 - The image is completely anonymised including the ears.
 - It works with both high and low resolution images.
- Disadvantages:
 - This software does not work on all head images and miscalculation of the affine matrix occurs, especially in clinical imaging.
 - It does not keep internal structures such as eyeballs, upper respiratory tract, skull near the ears, etc.





Software

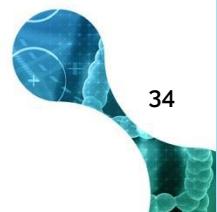
Face Masking



34



The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.





Software

Face Masking



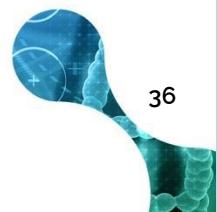


Software

Face Masking



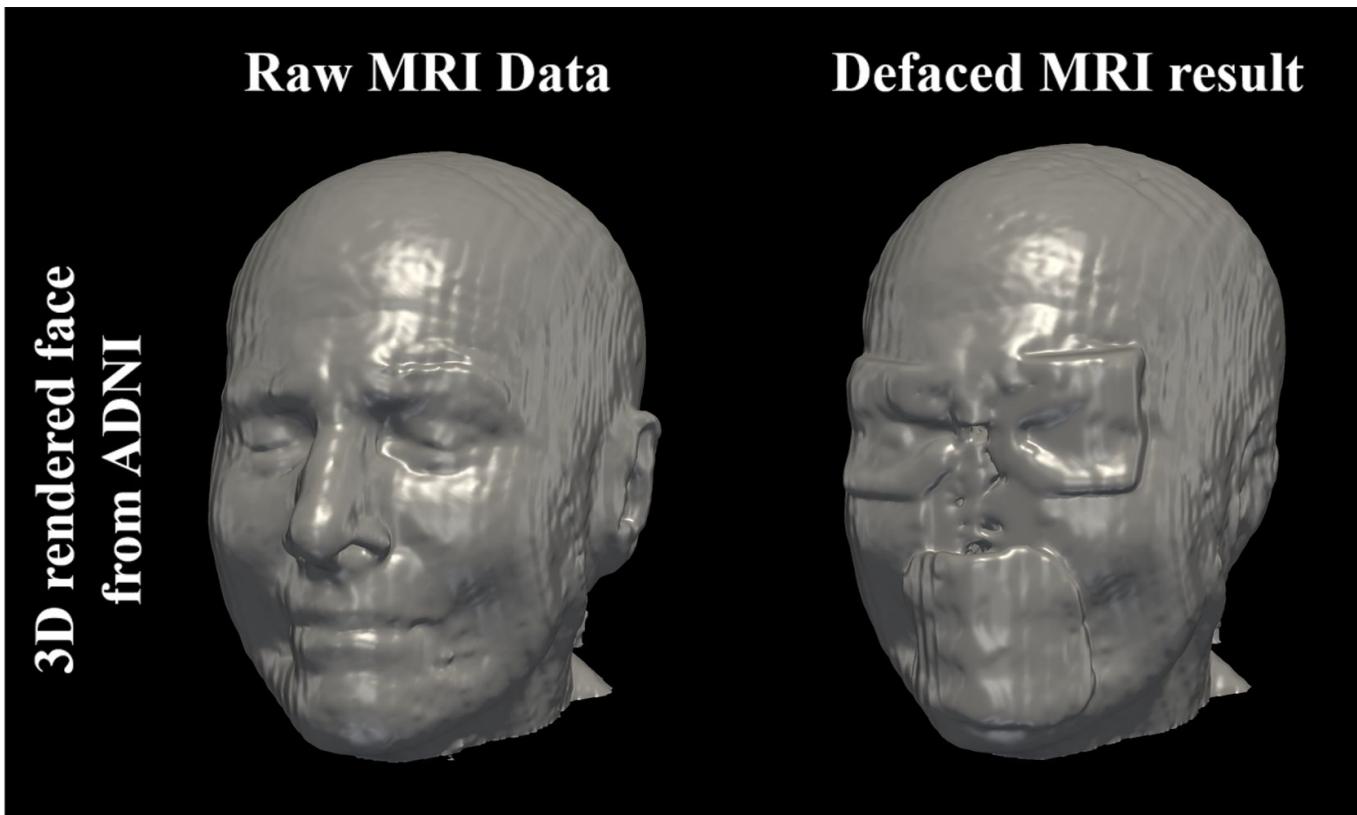
- Advantages:
 - The face is completely anonymised.
 - It keeps internal structures.
- Disadvantages:
 - The ears are not anonymised.





Software

Defacer



3D rendered face
from ADNI

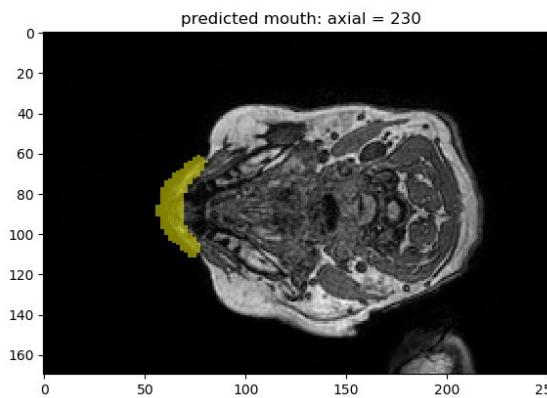
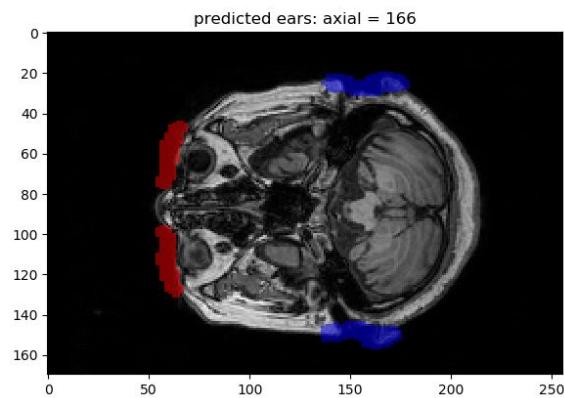
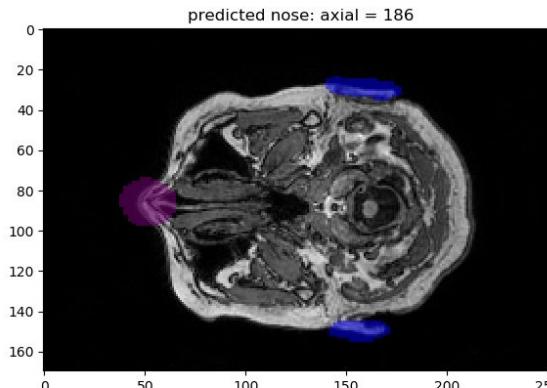
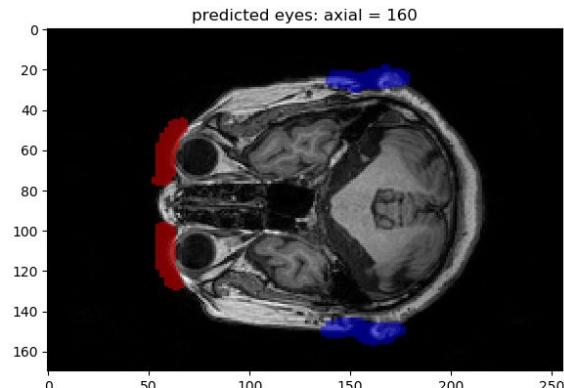


The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



Software

Defacer





Software

Defacer



- Advantages:
 - Only anonymised ears, eyes, nose and mouth.
 - It keeps internal structures.
- Disadvantages:
 - Among the facial features, wrinkles or around the mouth, there can be identifiers.
 - All training images are obtained in ADNI and OASIS datasets with optimal resolution. We do not know if it works in low resolution images.





DEEPHEALTH

Image annotation with text



The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



Some examples

Sensitive information

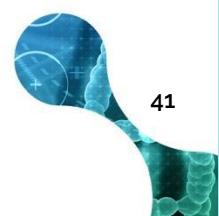
Name Surname1 Surname2	Genre	Birthday	Clinical history number	Other data

L: 0 W: 5770

Acquisition image Date

More data information

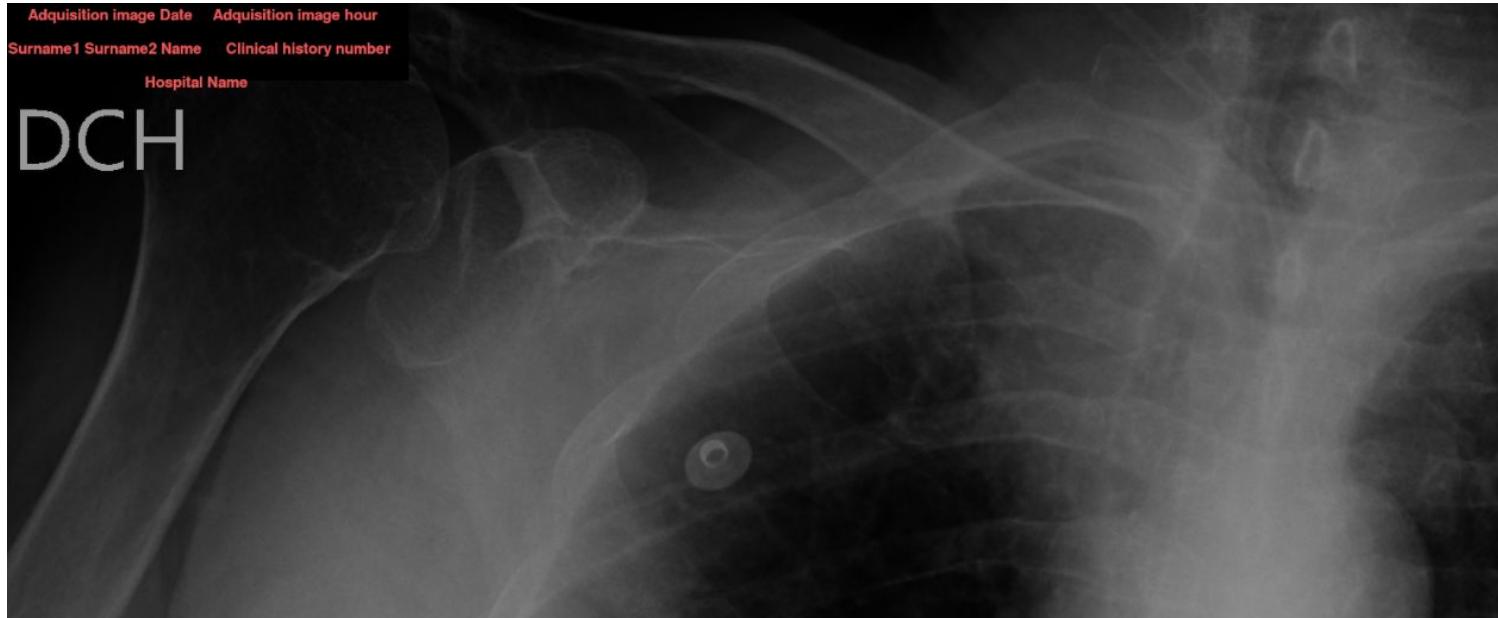
Image 1/1

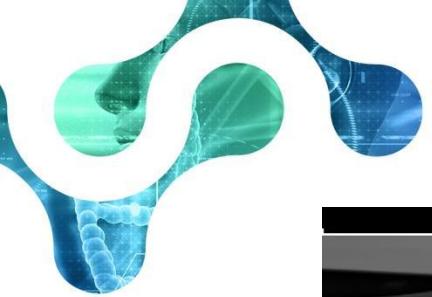




Some examples

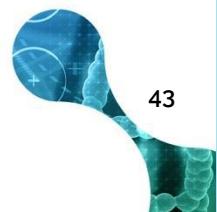
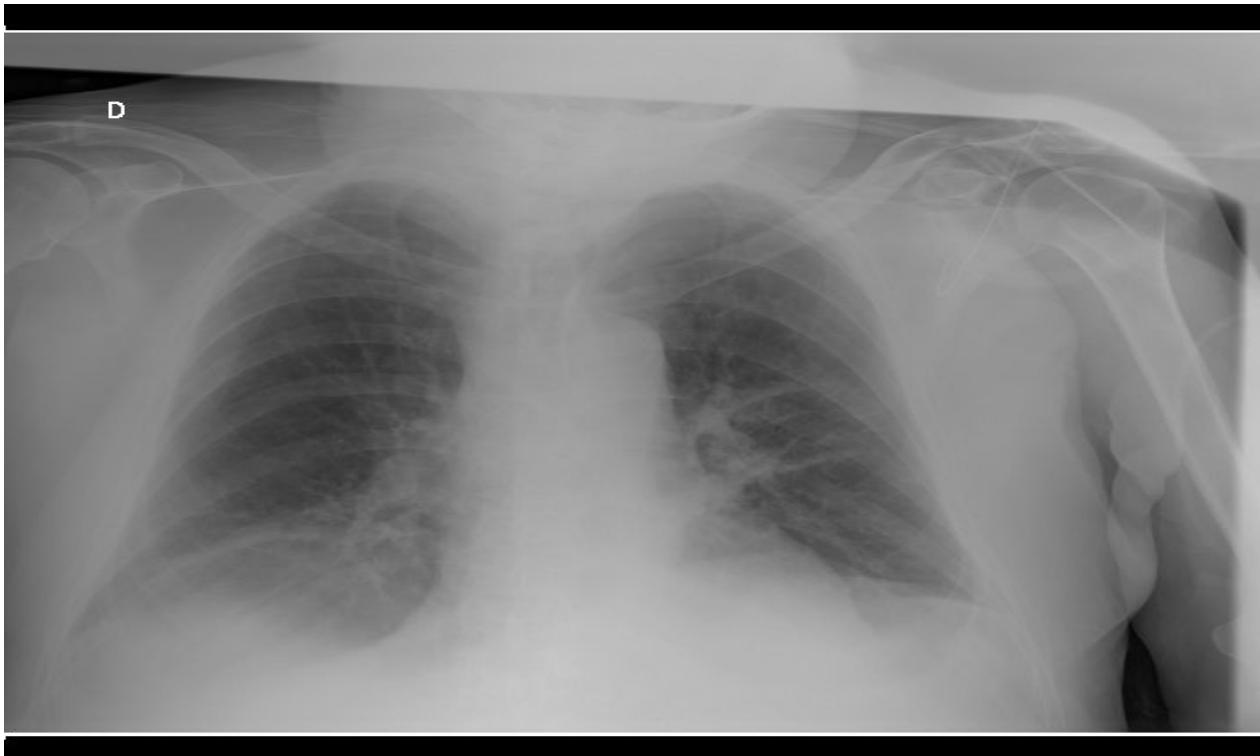
Sensitive information





Some examples

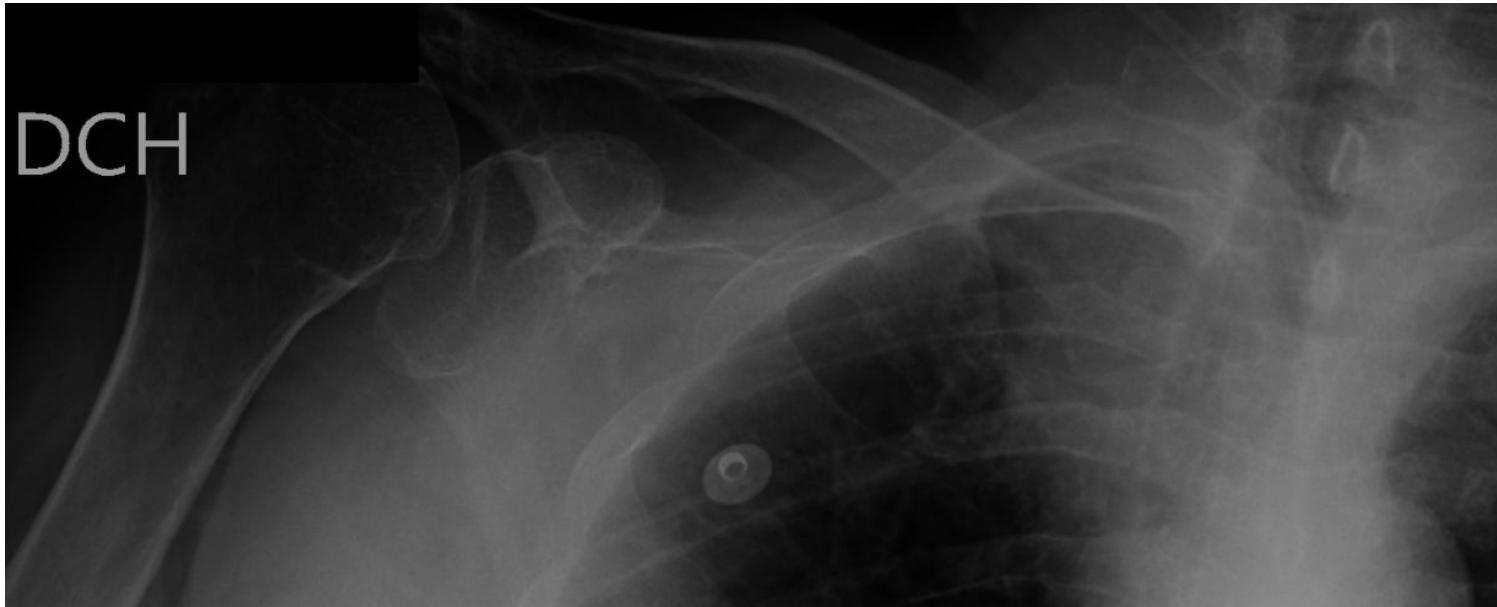
Objective





Some examples

Objective





DEEPHEALTH

Text removal methods

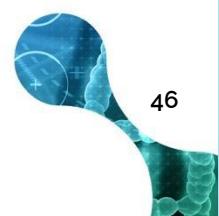


The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.



Mask delete

Our method: manual labeling and apply mask





Mask delete

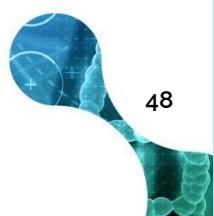
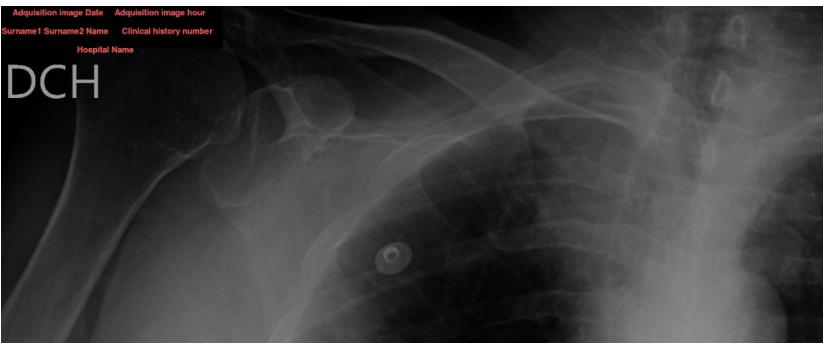
Our method: manual labeling and apply mask





Mask delete

Our method: manual labeling and apply mask





Mask delete

Problems



Carrie Bradshaw made this type of personalised chain fashionable during the Sex and the City series and it became fashionable again in 2018.





Mask delete

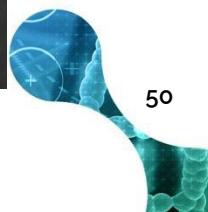
Problems



50



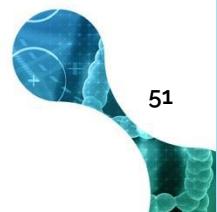
The project has received funding from the European Union's Horizon 2020 research and innovation programme under grant agreement No 825111.





Mask delete

Problems



Other automatic techniques

- OCR pattern recognition: **Pytesseract o OpenCV**
- Deep learning for image recognition: **EDDL + ECVL, Tensorflow, Pytorch**



Other automatic techniques

- Advantages:
 - Saves time by removing the manual labelling process.
- Disadvantages:
 - Can produce false positives and false negatives, therefore manual display is required for relabelling.
 - Does not distinguish (except in the case of Deep Learning) sensitive text from text that does not need to be anonymised.





DEEPHEALTH

Thanks for your
attention