

Identifying MRI sequence type from pixel data

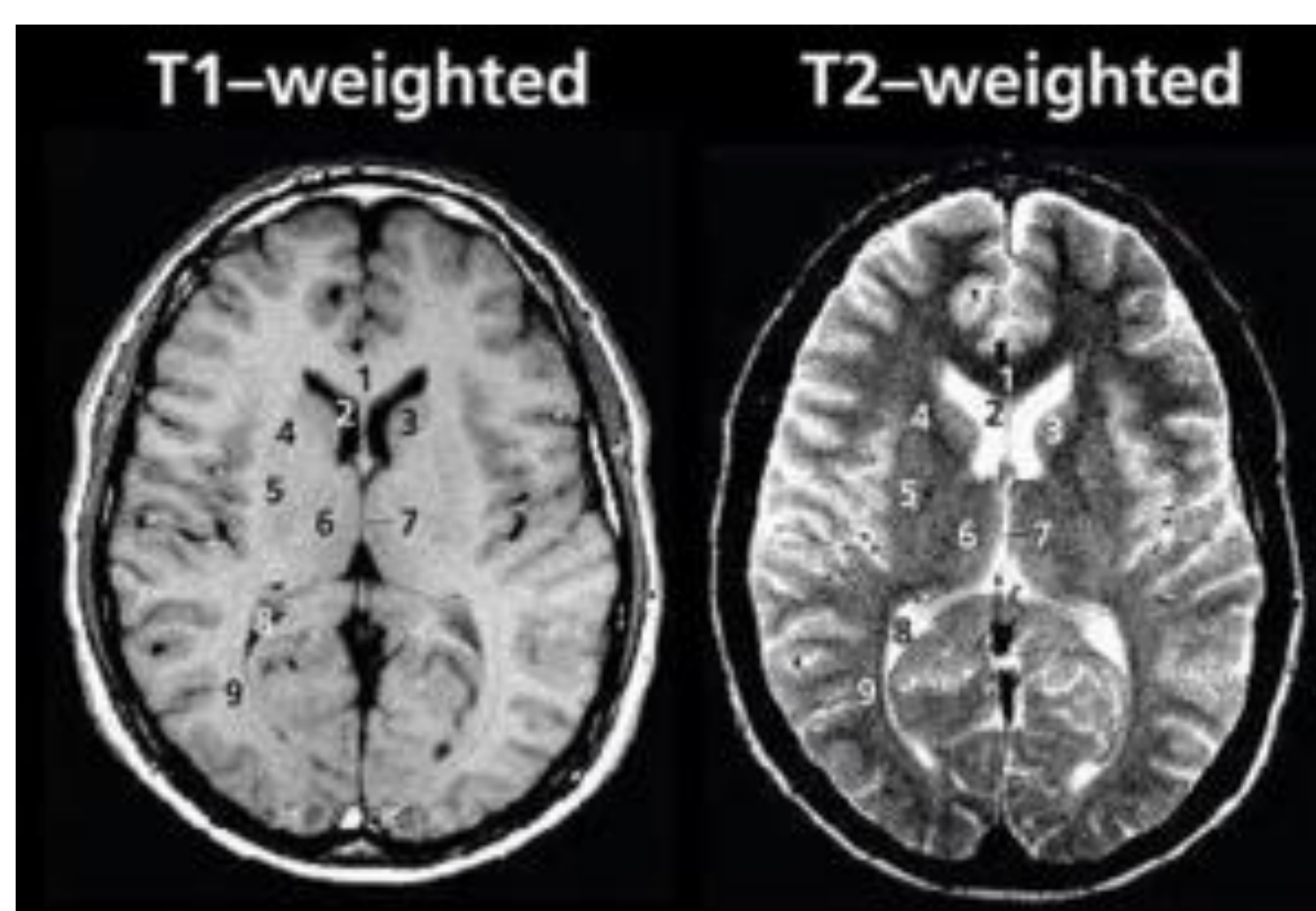
Smarti Reel¹, Esma Mansouri-Benssassi¹, Kara Moraw², Susan Krueger¹, Emily Jefferson^{1,3}¹Health Informatics Centre (University of Dundee), ²EPCC (Edinburgh University), ³HDR UK

Aim

To develop a robust, flexible and reusable pipeline for identifying brain magnetic resonance imaging (MRI) sequence type (T1, T2) by applying Machine Learning techniques to extract information from the pixel data in routinely collected images.

1. Background

- Building cohorts of medical images from routinely collected scans is a challenging process.
- Usually, a clinician uses image data to identify T1- and T2-weighted images by visually analysing the Cerebrospinal fluid (CSF). A dark CSF corresponds to T1-weighted imaging and bright to T2-weighted imaging (as shown below [1]).



- The Digital Imaging and Communications in Medicine (DICOM) standard provides a wide range of metadata consisting of tags providing additional information about an image.
- Tags in DICOM standard which indicate the sequence type to build meaningful cohorts are often missing/incomplete in routinely collected images.

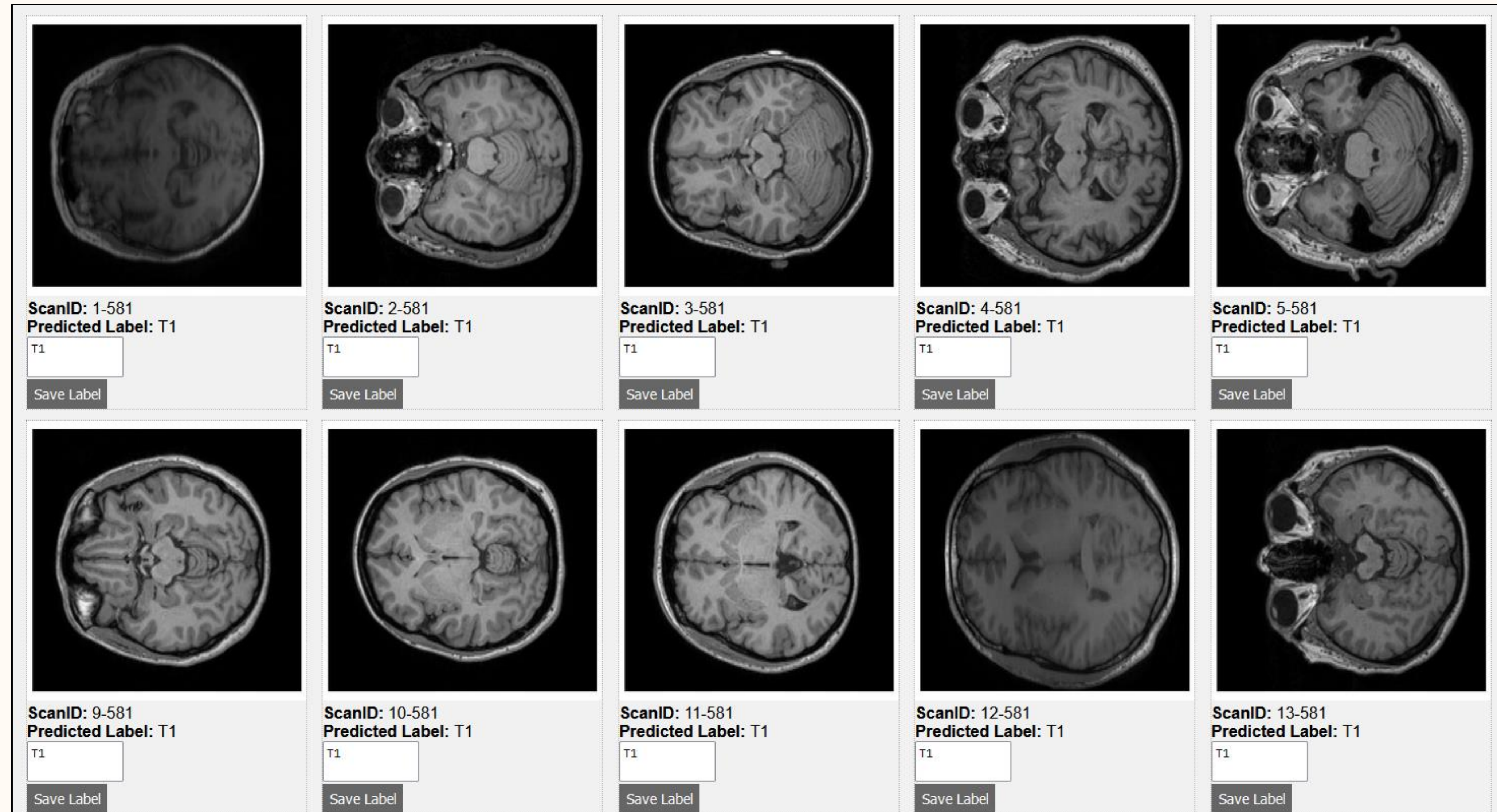
2. Motivation

- For cohort building, a more technical solution is required to facilitate and automate the process.
- Missing information, inconsistent tags and variability across different scanners make it highly challenging to query DICOM images solely using metadata [2–3].
- Using pixel data information can enhance the ability to build accurate and reliable cohorts.
- A robust method is required that uses pixel images from routinely collected brain magnetic resonance imaging (MRI) scans for identifying their sequence type: T1 or T2 for cohort building.

Acknowledgements

University of Dundee, University of Edinburgh, Public Health Scotland (eDRIS)

PICTURES project – This work was supported by the Medical Research Council (MRC) grant No. MR/M501633/1 and the Wellcome Trust grant No. WT086113 through the Scottish Health Informatics Programme (SHIP). This project has also been supported by MRC and EPSRC (grant No. MR/S010351/1) and by the Scottish Government through the “Imaging AI” grant award.



3. Model Architecture

The model architecture consists of two parts:

- An autoencoder to derive numerical features from the pixel data
 - Both the encoder and decoder components are composed of two convolutional layers.
 - The networks were first trained to reconstruct the input image with minimum error.
 - The lower-dimensional embeddings calculated by the encoder network were then stored as numerical features of the input images.
- A Support Vector Machine (SVM) for sequence type classification.
 - The classifier was trained on the extracted features instead of an entire image. This improved the performance and memory efficiency.

4. Model Training & Testing

Dataset: The models were trained on the open-source IXI dataset [4].

- It contains nearly 600 MR images from normal, healthy subjects from three hospitals in London.
- This work focused on the most common MRI sequences: T1-weighted and T2-weighted scans.

Preprocessing: MRI brain scans are composed of a varied number of 2D slices.

- Initially, a single axial mid-slice was extracted from each sample in the dataset.
- These 2D images were resized and fed into the autoencoder for feature extraction.
- To counter partial brain scans and varied image orientations in routinely collected data: 10 slices were sampled from each of the three axes per sample in the IXI dataset.

Results: The experiments show an overall accuracy of 97% for the MRI sequence type classification when evaluated on the IXI dataset.

5. Application GUI

- Pre-trained model can be applied to routinely collected brain MRI images.
- An intuitive graphical user interface (see figure above) allows an expert user to correct any misclassifications.
- The corrected labels can be fed back to retrain the classifier and refine its prediction performance.

Conclusions & Future work

An automated MRI sequence type classification combined with a human-in-the-loop approach is proposed.

- It enables users such as research coordinators, researchers, or data managers to share, annotate and extract relevant cohorts when DICOM tags metadata is unavailable or is insufficient in identifying datasets of interest.
- Our method is modular and thus reusable for similar applications.

Future work will aim at:

- examining the robustness and reusability of the model when applied on routinely collected MRI scans.
- reusing the model architecture and processing pipeline to identify body parts from pixel data (in the wider scope of the PICTURES project).

References

- Jefferson, E., Krueger, S., Macleod, R., Sutherland, J., Nind, T., Mudie, R., Prodan, B., Brooks, A., Wallace, R., Morris, C., Caldwell, J., Baxter, R., Parsons, M.: An architecture for building cohorts of images from real-world clinical data from the whole Scottish population supporting research and AI development. *Int J Popul Data Sci.* 7, 1916. <https://doi.org/10.23889/ijpds.v7i3.1916>.
- Park, C., You, S.C., Jeon, H., Jeong, C.W., Choi, J.W., Park, R.W.: Development and Validation of the Radiology Common Data Model (R-CDM) for the International Standardization of Medical Imaging Data. *Yonsei Med J.* 63, S74–S83 (2022). <https://doi.org/10.3349/ymj.2022.63.S74>.
- Willeminck, M.J., Koszek, W.A., Hardell, C., Wu, J., Fleischmann, D., Harvey, H., Folio, L.R., Summers, R.M., Rubin, D.L., Lungren, M.P.: Preparing Medical Imaging Data for Machine Learning. *Radiology.* 295, 4–15 (2020). <https://doi.org/10.1148/radiol.2020192224>.
- IXI Dataset – Brain Development, <https://brain-development.org/ixi-dataset/>, last accessed 2023/05/04.