ISLES 2022: A multi-center magnetic resonance imaging stroke segmentation dataset.

Hernandez Petzsche MR[1], de la Rosa E[2,3], Hanning U[4], Wiest R[5], Valenzuela Pinilla WE[5], Reyes M[6], Meyer MI[2], Liew SL[7], Kofler F[3], Ezhov I[3], Robben D[2], Friedrich T[1], Zarth T[1], Bürkle J[1], Baran T[1], Menze B[3,8], Broocks G[4], Meyer L[4], Zimmer C[1], Boeckh-Behrens T[1], Berndt M[1], Wiestler B[1], Kirschke JS[1]

[1]Department of Diagnostic and Interventional Neuroradiology, Klinikum rechts der Isar, School of Medicine, Technical University of Munich, Germany

[2]icometrix, Leuven, Belgium

[3]Department of Computer Science, Technical University of Munich, Munich, Germany

[4]Department of Diagnostic and Interventional Neuroradiology, University Medical Center Hamburg-Eppendorf, Germany.

[5]Institute of Diagnostic and Interventional Neuroradiology, University of Bern, Bern, Switzerland.

[6]ARTORG Center for Biomedical Engineering Research, Univ. of Bern, Switzerland

[7]Chan Division of Occupational Science and Occupational Therapy, University of Southern California, Los Angeles, CA, USA

[8]Department of Quantitative Biomedicine, University of Zurich, Zurich, Switzerland

Dataset overview: This dataset comprises 400 multi-vendor magnetic resonance imaging (MRI) stroke cases with high variability in lesion size, quantity and location. It is split into a training dataset of n=250 and a test dataset of n=150. The test dataset will be used for model testing only and will not be released to the public.

For each case, an expert level annotation of the stroke lesions is included along with the following three imaging sequences: Fluid attenuated inversion recovery (FLAIR), diffusion weighted imaging (DWI, b=1000) and its corresponding apparent diffusion coefficient (ADC) map. All imaging data and annotations are released in the Neuroimaging Informatics Technology Initiative (NIfTI) format (https://nifti.nimh.nih.gov/nifti-1), according to the BIDS convention. All imaging data are released in the native space without prior registration. Prior to release, skull-stripping was performed to de-identify patients.

Image acquisition was performed on one of the following devices: 3T Philips MRI scanners (Achieva, Ingenia), 3T Siemens MRI scanner (Verio) or 1.5T Siemens MAGNETOM MRI scanners (Avanto, Aera). All images were obtained by healthcare professionals as part of the clinical imaging routine for stroke patients at three different stroke centers and imaging data was collected retrospectively for different clinical studies. For this challenge, a subset of fully anonymized images of those studies was selected. Computer-readable scanner metadata from the Digital Imaging and Communications in Medicine (DICOM) header in the JSON file format is provided with the datasets if available.

MRIs from the following centers are included: Center #1: University Hospital of the Technical University Munich, Munich, Germany. Center #2: University Hospital of Bern, Bern, Switzerland. Center #3: University Medical Center Hamburg-Eppendorf, Hamburg, Germany. The training set comprises sub-acute, post-treatment stroke imaging data from centers #1 and #2. The test set comprises data from all the three centers: sub-acute, post-treatment stroke imaging data from centers #1 and #3 and acute, pre-treatment stroke imaging data from center #2. All three participating centers received approval from their respective ethics commission (TUM 2021-274_1-S-SR).

Lesion annotation: This dataset includes acute and subacute stroke lesions. These lesions characteristically have high signal on DWI and FLAIR and low signal on ADC. In rare cases of hyper-acute stroke, FLAIR signal may be normal. In such cases, lesions are defined only based on DWI and ADC signal characteristics. Lesions were pre-segmented by a locally trained neural network and saved in the same imaging space as the DWI. Task-instructed medical students then checked and corrected the lesion predictions from the net. Lesion masks were then reviewed and edited by an experienced neuroradiology fellow with special interest in stroke imaging. Overall quality control of

the annotations was performed by one of three senior neuroradiologists with >10 years of stroke imaging experience.

Challenge task: Prediction of MRI stroke lesions in the test dataset.