SAMson: an automated brain extraction tool for rodents using SAM

Daniel Panadero Soler[1], Mohamed Kotb Selim[1], Emma Muñoz-Moreno[2], Pedro Ramos-Cabrer[3], Pilar López-Larrubia[4], Antonio Pertusa[5], Silvia De Santis[1] and Santiago Canals[1]

[1] Instituto de Neurociencias, Consejo Superior de Investigaciones Científicas - Universidad Miguel Hernández (CSIC-UMH), Sant Joan d'Alacant, Spain
[2] Institut d'Investigacions Biomèdiques August Pi i Sunyer (IDIBAPS), Barcelona, Spain
[3] Center for Cooperative Research in Biomaterials (CIC biomaGUNE), Basque Research and Technology Alliance (BRTA), Donostia–San Sebastián, Spain
[4] Instituto de Investigaciones Biomédicas Sols-Morreale, Consejo Superior de Investigaciones Científicas - Universidad Autónoma de Madrid (CSIC-UAM), Madrid, Spain
[5] University Institute for Computer Research, University of Alicante, Alicante, Spain
scanals@umh.es

Abstract. Brain extraction, i.e. the precise removal of MRI signal outside the brain boundaries, is a key step in MRI preprocessing pipelines, typically achieved via masks delineating the region of interest (ROI). Existing automated tools often lack accuracy for rodent MRI due to resolution limitations, so large manual editing efforts are required. This work introduces SAMson, a high-precision automated mask generator built on Meta AI's Segment-Anything Model (SAM). SAM's adaptability to diverse tasks, akin to other foundation models (Chat-GPT), is harnessed to address the scarcity of training data in this domain. SAMson is a Python-based tool that integrates SAM's capabilities with the requirements and characteristics of multislice MRI data. SAMson offers two prompt generation approaches: Semi-Auto, enabling manual prompt selection for precise control, and Full-Auto, with automated prompt generation. Evaluation against gold-standard masks extracted by an experienced experimenter demonstrated excellent performance of SAMson, and benchmarking against established methods (BET from FSL) demonstrated SAMson's superiority. Robustness assessments on datasets obtained from an external MRI facility, confirmed versatility across scanner setups and image resolutions. Our tool's adaptability and performance underscore its potential for widespread adoption in rodent MRI research, supported by open-source availability for the scientific community.

Keywords: Brain extraction, Segmentation, Rodent MRI, MRI analysis, Automatic pipeline, AI.

1 Introduction

Magnetic resonance imaging (MRI) is a fundamental component of preclinical research in neuroscience due to its capacity to provide both functional and structural information

about the brain in longitudinal experiments. MRI-based preclinical results have demonstrated great translational validity (i.e. [1,2]). Brain extraction, i.e. the precise elimination of MRI signal from voxels outside the boundaries of the brain, is a key step in MRI preprocessing pipelines. Typically, this is achieved by creating masks that outline the region of interest (ROI), from which all MRI measurements are ultimately computed. Masks are binary versions of MRI images wherein voxel assignments are set to 1 to denote brain inclusion and to 0 to indicate absence of brain region. Several automated segmentation tools are available for rodent MRI, but unlike in human MRI, the accuracy of these methods is compromised due to the low image resolution relative to the brain-skull distance and the complex anatomical structure [3], leading to large manual editing efforts. Here, we present a high-precision automated mask generator built on the most powerful segmentation model to date: Meta AI's Segment-Anything Model (SAM), whose performance is often comparable to fully supervised results [4]. Our aim was to achieve operator-independent mask extraction of rodent MRI brain images. We demonstrate that our tool provides highly accurate segmentation outcomes, showing excellent agreement with masks extracted manually by an experienced experimenter. This advancement holds significant implications for efficient brain extraction in rodent MRI, streamlining preclinical research workflows and facilitating translational research.

2 Methods

2.1 Segmentation model

SAM is a promptable image segmentation model that can generalise to tasks and data distributions beyond those encountered during training [5], solving segmentation problems for which there is scarce training data. We have developed SAMson, a Python-based tool that exploits this potential, by integrating SAM's capabilities with the requirements and characteristics of multi-slice MRI data. We use the NIfTI (Neuroimaging Informatics Technology Initiative) image format, a widely adopted standard for storing neuroimaging data, for both the input data fed into SAMson and the segmentation outcomes. For each slice in a brain image, SAMson produces an independent mask prediction, in response to input prompts. These prompts consist of markers, dots or bounding boxes, approximately indicating the ROI to be segmented.

2.2 SAMson's Semi-Auto & Full-Auto approaches

We implemented two strategies to generate the prompts required for brain segmentation. Firstly, the "Semi-Auto workflow" involves manual prompt selection for each slice, leveraging the intuitive application interface of SAM (executed locally as a React JavaScript application), affording precise control over the extraction result. In this case, SAMson copes with the application processing PNG-formatted images and generating cut-outs of the selected ROIs instead of the intended masks. The Semi-Auto procedure starts with the generation of PNG image files of all brain slices, enabling users to extract them sequentially within the application environment and save the outputs likewise as

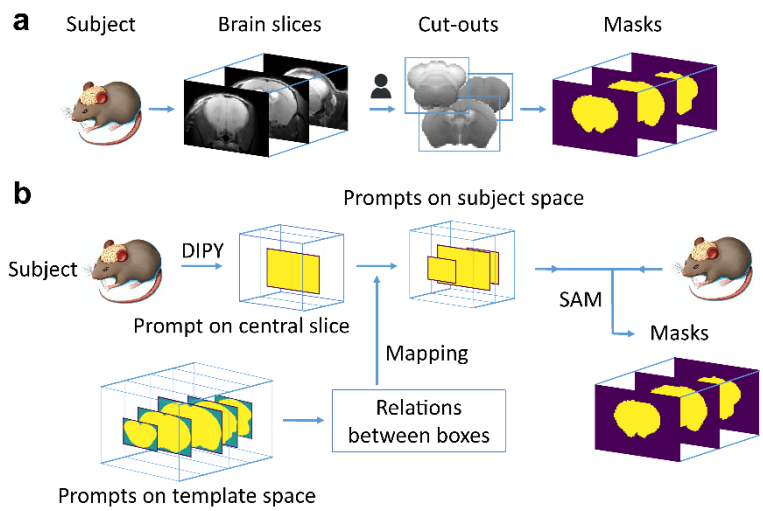

Fig. 1. Schematic representation of the semi-automatic (a) and fully automatic (b) approaches for mask generation using SAMson. The Semi-automatic approach involves user interaction in selecting ROIs from PNG-formatted brain slice images, followed by SAMson's determination of cut-out positions within the complete brain slice images to generate masks. Conversely, the fully automatic approach entails SAMson autonomously generating masks directly from input images without user intervention.

PNG files. SAMson subsequently identifies the position of the cut-outs within the brain slice images, ultimately producing the desired masks.

Secondly, we developed a "Full-Auto pipeline" that generates the prompts automatically, adapting to the displacement and resizing of the ROI across the brain slices, to compensate interindividual variability and variations in subject placement within the MRI scanner. Box input prompts are used to enhance prediction stability, as they offer less ambiguity in indicating ROI size. This adaptability is achieved by determining the placement of the input box for the central slice of the brain using the median-Otsu method [6] implemented in DIPY (Diffusion Imaging in Python library) [7] and interpolating its boundary positions to the rest of the slices through the following mapping technique. Input prompts for a template, enclosing the labeled brain region in each slice, are automatically generated, facilitating the determination of spatial relations between the central slice bounding box and those of the remaining slices in the template space. These relations are transposed into the subject space, wherein the allocation of subject slices within the anteroposterior range of the template is considered. This allocation is guided by user-defined specification of the template's slices that best correspond to the subject's most rostral and caudal slices. Afterwards, they are applied on the previously produced central slice box to obtain the complete set of input boxes Subsequently, following the generation of prompts for a specific subject, SAMson can provide three distinct mask predictions for each slice, empowering the user to select the optimal mask formulation. These outputs do not directly represent SAM's predictions,

but rather stem from their modification through morphological image processing techniques. Such operations are common in the field of computer vision and facilitate systematic refinement of masks.

2.3 Performance analysis

The Semi-Auto workflow is a completely supervised method that leaves no margin for brain extraction errors beyond those assumed by the user. The quality of the masks generated using the Full-Auto approach was evaluated through comparison with those extracted by an expert experimenter. Three different metrics were used: the accuracy, the Jaccard index and the Dice coefficient. More specifically, the accuracy is computed as (1- ((FP+FN)/AP))·100%, where FP (FN) ≡ number of pixels incorrectly predicted as (not) belonging to the brain and AP ≡ number of pixels comprising the brain. The Jaccard index and the Dice coefficient, both common metrics in sample similarity analysis, are based on the ratio of the intersection over union of the two masks. For this performance assessment, a cohort of N=53 mice subjects across different age groups (3, 4.5, 6, and 12 months) was used. Each subject comprised 16 brain slices, resulting in 848 automatically generated masks. Additionally, the performance of the tool was benchmarked against a prevalent method, BET from FSL, across all N=53 subjects. Finally, SAMson's robustness was evaluated by applying the same methodology to repository images obtained from external MRI services, acquired employing different scanner setups and image resolutions. This external image batch encompassed N=10 animals, with 20 slices each (200 masks). In this case, the anteroposterior range of the image acquisition exceeded the extent of the brain, so the first and last two slices were excluded from the evaluation.

3 Results

Both the Semi-Auto and Full-Auto SAMson versions yielded highly accurate segmentation outcomes for rodent MRI brain images, showing excellent agreement with the masks extracted manually by an experienced experimenter. Changes in the MRI modality, resolution or number of brain slices of the images did not affect performance nor required code adjustments. In particular, SAMson's Semi-Auto workflow made "manual" brain extraction much faster and efficient than with current analogous methods, as only few manual selections are required for each mask extraction. The results of our Full-Auto pipeline were outstanding on average among all our mice subjects (rat analysis ongoing): accuracy 92.88 ± 1.34 %; Dice coefficient 0.9643 ± 0.0064; Jaccard index 0.9320 ± 0.0115 (Fig.2a). Additionally, the robustness of the tool tested with external data exhibited promising results: accuracy 96.01 ± 0.52%; Dice coefficient 0.9798 ± 0.0027; Jaccard index 0.9607 ± 0.0051 (Fig.2b).

To asses SAMson's performance in more detail, we next compared the performance in the anteroposterior axis, across slices (excluding external data due to variances in the antero-posterior range). This approach enables a more accurate depiction of our tool's

segmentation capabilities, which could be undermined in Fig.2 owing to averaging effects. Fig.3a shows representative median masks provided by SAMson in different antero-posterior slices, together with the quantified performance per slice (Fig. 4b). Manual editing efforts are contingent not only upon the average error but also on its distribution across slices. The concentration of errors observed in a few slices facilitates manual correction, if deemed necessary.

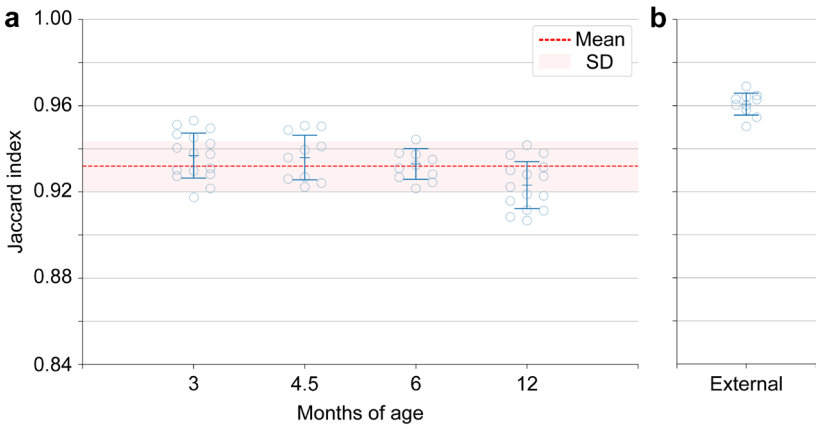

Fig. 2. Brain extraction performance with SAMson. a) Jaccard index computed among N=53 subjects (mice), i.e. 848 generated masks, with respect to manual masks defined by an expert. Each data point represents the average performance for each subject, computed across slices. Global mean and standard deviation in red. b) Extension of (a) to include performance analysis on external data. Subjects in the external batch had an average age of 6 months.

To underscore the significance of the developed automated pipeline, we conducted a comparative analysis of its performance against one of the most established brain extraction methods: FSL's Brain Extraction Tool (BET). We employed a set of parameters in BET optimised for mouse MRI data. We compared the performance of SAMson and BET in generating brain masks in our mouse dataset (Fig. 4). Analysing both the mean performance per subject (Fig, 4a) or per slice (Fig. 4b), we found that SAMson clearly outperforms BET.

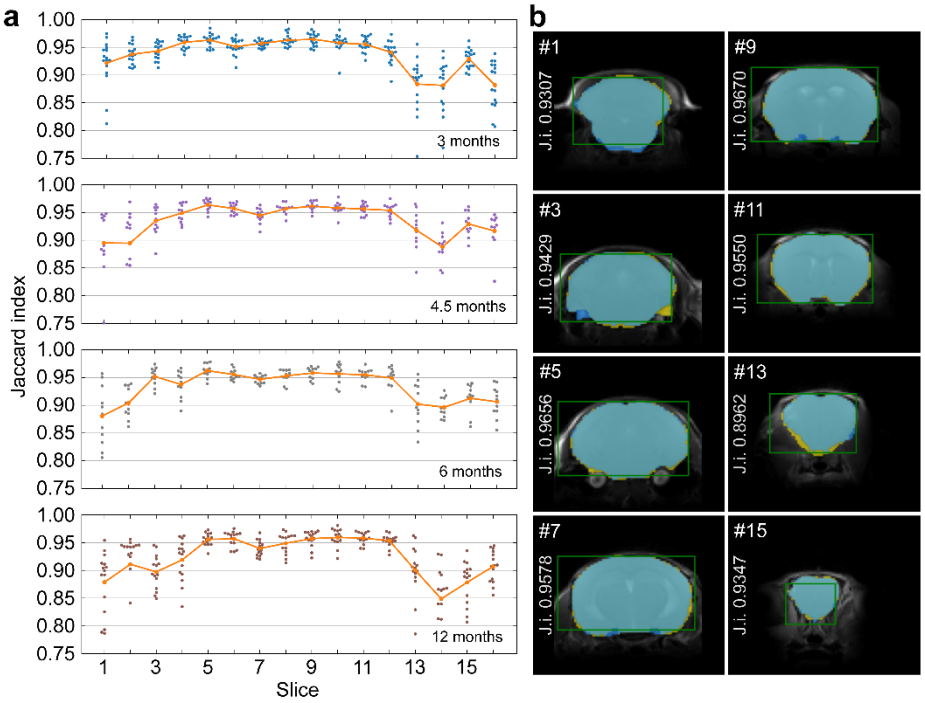

Fig. 3. SAMson performance in the antero-posterior axis. a) Jaccard index computed among N=53 subjects (mice), i.e. 848 generated masks, with respect to manual masks. Each data point represents the performance score of a single slice mask. Orange lines connect the mean values computed for each slice group within their respective age group. b) Comparison of manually created masks (yellow) and SAMson's predictions (blue) overlaid on the T2-weeighted MRI image, alongside automatically generated bounding boxes (green) used as input prompts for SAM segmentation. Shown are 8 selected, out of 16 total, median masks of slice groups from more caudal (1) to more rostral (15), corresponding to 3-month-old mice.

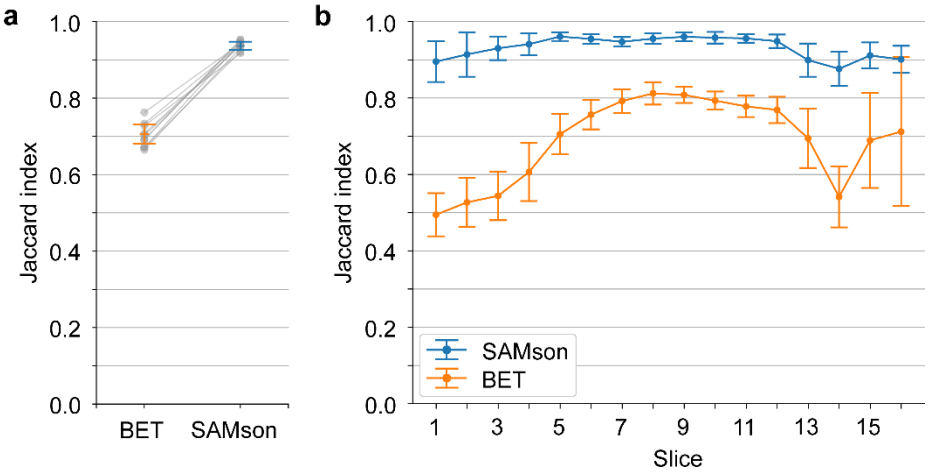

Fig. 4. Comparison between optimized BET and SAMson performance. a) Performance per subject averaging across slices, in the 3-month age group (n=16 mice, 256 masks) in which BET exhibited its most favorable performance. Data points corresponding to the same subject are connected (grey). b) Performance per slice averaging across all mice (n=53, 848 masks). Symbols represent mean ± SD.

4 Discussion

SAMson has demonstrated remarkable accuracy in segmenting rodent MRI brain images, superior than currently available methods, and proves robust when applied to images from external MRI services. The versatility of our tool is underscored by the complementary nature of its two approaches, which serve to a wide spectrum of user requirements. For instance, the Semi-Auto mode is ideal for scenarios demanding utmost accuracy, specific brain region or tumor segmentation, aged subjects, or cases where the convenience of the Full-Auto implementation is not applicable.

Automatically generated masks by SAMson rarely exhibited recurrent errors across multiple slices, which can be attributed to two primary factors. Firstly, SAM's tolerance to input prompt inaccuracies enables it to adapt its prediction strategy to encompass regions beyond the specified bounding box, as exemplified in Fig.4. Secondly, the application of morphological transformations on SAM's mask predictions consistently improved accuracy by removing false positive pixels and regions marginally connected to the main cluster of brain pixels. This refinement process aims to enhance SAM's efficacy, aligning it more closely to the nuances of our specific segmentation problem.

It is noted that SAMson's errors tend to be localized within specific slices, facilitating mask refinement when required, thereby enhancing operational efficiency. This phenomenon can be attributed to SAM's segmentation of two-dimensional images, resulting in independent predictions for each slice of a subject. In contrast, other brain extraction tools, e.g. BET, usually generate volumetric predictions, where masks for different image slices are not extracted independently. In case of errors, this results in

propagation across slices, a feature absent in our approach. When errors in the automatic segmentation arose, they consistently clustered within the most rostral or caudal slices. In these regions, anatomical changes manifest more abruptly, and partial volume effects result in the voxels at the edges containing a greater mixture of brain and extracerebral tissue, thereby posing a challenge to mask prediction. Higher performance in the external dataset is likely due to the higher resolution of the images, which will translate into smaller partial volume effects, and better definition of tissue borders. However, the decrease in performance in the more rostral and caudal masks, where the masks are smaller, can also be attributed to a size bias. This bias is due to the fact that the Jaccard index is based on relative overlap, which is influenced by mask size differences. In the smaller rostral and caudal masks, any deviation from the ground truth translates into a larger relative difference in overlap, leading to lower scores.

To conclude, we have introduced a brain extraction tool with potential for widespread adoption, enhancing the analysis of rodent brain imaging data and offering a significant advantage for preclinical neuroscience. The versatility of our software renders SAMson highly efficient and adaptable across a diverse array of rodent experiments, scanner resolutions and MRI modalities. Our ongoing commitment involves refining the tool in future iterations, with a focus on optimizing SAM for MRI rodent brain image segmentation. Accordingly, we are actively pursuing robust automated brain extraction for rat subjects. Our ultimate objective is to foster accessibility to the scientific community by providing open-source code, thereby empowering researchers to utilize and further refine the tool.

Acknowledgments. Spanish Ministerio de Ciencia e Innovación, Agencia Estatal de Investigación (PID2021-128158NB-C21 and PID2021-128909NA-I00) and Programs for Centres of Excellence in R&D Severo Ochoa (CEX2021-001165-S). The Generalitat Valenciana Government (Prometeo excellence program CIPROM/2022/15, and Valencian Consellería de Innovación INVESTIGO 2022 program INVEST/2022/396, fellowship code 2022/PER/00004). "la Caixa" Foundation (ID 18100010434), fellowship code LCF/BQ/DI18/11660067 and Marie Skłodowska-Curie-COFUND agreement (Grant No. 713673).

Disclosure of Interests. The authors have no competing interests to declare that are relevant to the content of this article.

References

1. De Santis S, Bach P, Pérez-Cervera L, et al. Microstructural White Matter Alterations in Men With Alcohol Use Disorder and Rats With Excessive Alcohol Consumption During Early Abstinence. JAMA Psychiatry. 2019;76(7):749–758. doi:10.1001/jamapsychiatry.2019.0318
2. Sommer WH, Canals S, Bifone A, Heilig M, Hyytiä P. From a systems view to spotting a hidden island: A narrative review implicating insula function in alcoholism. Neuropharmacology. 2022 May 15;209:108989. doi: 10.1016/j.neuropharm.2022.108989. Epub 2022 Feb 22. PMID: 35217032.
3. HSU, Li-Ming, et al. Frontiers in neuroscience, 2020, vol. 14, p. 568614
4. KIRILLOV, Alexander, et al. “Segment Anything” 2023 IEEE/CVF International Conference on Computer Vision (ICCV) (2023): 3992-4003
5. BOMMASANI, Rishi, et al. arXiv preprint, 2021, arXiv:2108.07258
6. Otsu, Nobuyuki. "A threshold selection method from gray-level histograms." IEEE transactions on systems, man, and cybernetics 9.1 (1979): 62-66.
7. Garyfallidis E, Brett M, Amirbekian B, Rokem A, van der Walt S, Descoteaux M, Nimmo-Smith I; Dipy Contributors. Dipy, a library for the analysis of diffusion MRI data. Front Neuroinform. 2014 Feb 21;8:8. doi: 10.3389/fninf.2014.00008. PMID: 24600385; PMCID: PMC3931231.