Peer Review File

Article information: http://dx.doi.org/10.21037/atm-22-3712

Reviewer A

Interesting technique though I question the utility of automatic labeling and how much time is actually saved in this semi-supervised labeling process. Would like to see a comparison of workflow timing using DELLO vs manual labeling for an experienced researcher.

Reply: Thank you for your interest on this proposed method. It's reasonable to question the utility of this tool solely based on text description and limited figure illustrations. Therefore, we added a screen capture video demonstrating the basic workflow of electrode localization and labeling. We attached the video file in the supplementary material, so that anyone who is interested in this package could evaluate the codes following the procedures using the example data provided along with the open-source codes.

The electrode localization and labeling procedure is fundamental to SEEG related analysis both from clinical and neuroscience perspective. For experience anatomist or specialist with medical background such as neurosurgeons, the labeling procedure is not very difficult, although time consuming, but for other researchers without much experience on neuroanatomy, it would be quite demanding. Driven by the willingness to facilitate other people's research based on SEEG, we build this DELLO tool and hope to share this technique to the research community. The DELLO tool is flexible to incorporate different volume-based atlas. In this study, we included AAL and YEO7 atlases, but other atlases could be applied with minor modification. We will maintain and update this tool if we could hear more advices. Currently, we are using this tool as a basic method for several follow-up SEEG papers.

As for the comparison of time consuming using DELLO vs manual labeling, the DELLO labeling is fully automatic and the labeling section cost roughly less than 1min computational time for 100 contacts in a mainstream laptop after the semi-automatic contacts coordinates acquisition. In the additional analysis suggested by reviewer B, we invited Dr. KZ and CZ to label all contacts blind to the DELLO results according to the localization images, and the average time for gray/white matter and AAL based anatomical labeling for 100 contacts were ~30 min in total.

Changes in the text:
Method line 251-254:
To further avoid the potential bias introduced by the DELLO anatomical labeling, we invited two senior neurosurgeons (KZ and CZ) to label each contact according to the AAL atlas definitions independently and blind to the algorithm results. Discrepancies were resolved by discussion.

Results line 318-319:
The gray and AAL manual labeling for 100 contacts took ~30min for an experienced specialist.

Reviewer B

In this manuscript the authors describe an open-source algorithm for localization of depth electrodes. They build upon registration, normalization, and parcellation algorithms in SPM to semi-automatically localize contacts. They apply the algorithm to a cohort of 7 patients / 80 electrodes to assess localization accuracy.

Localization of electrode contacts is important for meaningful interpretation of SEEG, and as the authors state, software used to do so often has a high barrier to entry with software complexity and/or significant user input required. Automation of the localization process makes this process easier and more clinically feasible. In particular, automated application of anatomic labels could streamline clinical interpretation.

Some comments:

1. A significant proportion of the introduction and discussion focus on localization of subdural electrodes which is a much different problem with not much bearing on the problem at hand. There is an existing literature regarding algorithms for depth electrode localization (e.g. see Davis et al. 2021 Frontiers in Neuroscience) and similar algorithms have been built into clinical software packages (e.g. Curry by Compumedics). Greater discussion should be directed towards the relevant literature including differences from previously described algorithms.

Reply 1: Thank you for your professional suggestions on the introduction and discussion sections. Indeed, the localization of subdural electrodes is a different problem, we reorganized the introduction and discussion sections and the subdural electrode localization content is largely reduced. In the discussion section, a paragraph introducing other existed tools and describing the DELLO featured functions and has been added.

Changes in the text:
Introduction
Delete some subdural related wording.

Discussion:
Delete:
An additional step for localizing the subdural electrode requires projection of the grids onto the brain surface to correct the brain shift after craniotomy. However, for depth electrodes, brain shift correction is unnecessary.

line 373 – 391 added:
Several inspiring works have been made regarding algorithms for depth electrode localization. For example, Tyler S. Davis et al. proposed locate electrodes graphical user interface (LeGUI)

toolbox containing user friendly interface and its anatomical labeling accuracy reached 94% (26); another tool integrated into 3D Slicer platform SEEG Assistant (27) developed by Massimo Narizzano et al. gave nice 3D visualization of the depth electrodes; fast automated stereo-EEG electrode contact identification and labeling ensemble (FASCILE) written in Python claimed less than 1 mm localization error and ~ 10 min manual identifying, sorting, and labeling for electrode contacts (28). Other techniques such as EpiTools (29) and toolbox by Liang Wang et al. (14) adopted similar workflow for depth electrode reconstruction. In addition, there is also commercially available multimodal neuroimaging software CURRY® (Compumedics Limited, Victoria, Australia) which include the intracranial EEG analysis module. Compared with the abovementioned works, DELLO focused on the localizing the depth electrode with optimized efficient workflow. To achieve this goal, additional CT high density sampling and clustering were employed to further elevate the automatic degree of the algorithm. As a result, the reviewer manual operating time is minimized as possible and the efficiency was further boosted without compromising accuracy.

2. The processing time and required user input/experience seem underestimated. The described method requires file format conversion and use of a separate research-oriented MATLAB package (SPM), in addition to manual input. These do not preclude development of a minimal user input pipeline but as is these limitations should be mentioned.

Reply 2: Thank you for pointing it out. The file format conversion includes converting the MRI and post operative CT DICOM images to NIfTI format, which could be performed using other tools like dcm2nii or dcm2niix. We took advantage and integrated the coregistration and segmentation functions from the SPM package, so that the users typically do not need to call the SPM itself in most cases, unless other procedures are needed, such as reorienting the origin of the images. In the revised manuscript, we also included a screen capture video demonstrating the manual procedures to enhance it utility. We added those points in the limitation parts.

Changes in the text:
Limitations:
Line 426 - 429
Last but not least, additional steps including neuroimage format conversion and other possible data correction were not integrated in the pipeline. These drawbacks will be addressed in future updated versions.

3. It sounds like the neurosurgeon WH was given algorithm results for each contact and asked to classify each result as correct or incorrect. A more robust method would be for a researcher who is blinded to the algorithm results to independently classify the locations of each contact and compare the manual and semi-automated results. The former method introduces bias and uncertainty (when was accuracy "good enough" to be correct?).

Reply 3: Thank you for recommending the better analysis strategy. Admittedly, the review and judge approach could be biased by the output of DELLO labeling results, therefore, we added an additional blind evaluation in the method and results.

Changes in the text:
Method line 251 - 254:
To further avoid the potential bias introduced by the DELLO anatomical labeling, we invited two senior neurosurgeons (KZ and CZ) to label each contact according to the AAL atlas definitions independently and blind to the algorithm results. Discrepancies were resolved by discussion.

Results line 314 - 319:
The match rate between automatic labeling and independent blind Gray/AAL labeling for each patient were: 84.85%/75.00%, 81.94%/79.86%, 85.06%/75.97%, 79.41%/78.68%, 87.07%/81.90%, 91.67%/88.54% and 83.97%/80.13% and the average match rates were 85.15% and 80.39%, preserving acceptable accuracy. The gray and AAL manual labeling for 100 contacts took ~30min for an experienced specialist.

4. It's not clear to me why the gray matter mask wasn't applied to the anatomic atlas based contact localization from the start. AAL labels gray and white matter, but labels like "middle frontal gyrus white matter" are probably not more meaningful that just "white matter." The accuracy for atlas gray matter was almost identical to grey vs. white matter and significantly better than overall AAL labeling. I would suggest emphasizing gray matter localiation accuracy.

Reply 4: Thank you for your suggestion. In current design, the gray and white matter (G/W) labeling and anatomical labeling was in parallel. Indeed, we found the anatomical labeling results were more favorable after excluding the white matter electrodes. However, the gray and white matter labeling was not perfect, we choose to also output the G/W results to give more freedom of manual adjustment for the users. In rare cases, filtering the gray matter contacts from the beginning may lead to the neglecting of important contacts. Taking advantage of the G/W labels, the users could easily filter out the contacts of interest. We emphasized this strategy in the results section.

Changes in the text:
Results line 330-332:
Taking advantage of the output gray and white matter labels, the users could easily filter out the contacts of interest. We recommend filtering AAL labels using gray matter masks.

5. The straight electrode assumption is probably unnecessary and could be problematic in rare cases e.g. with pneumocephalus. Other algorithms have projected to contacts to nearby maxima on CT (can be done after estimating contact locations assuming perfectly straight electrode) which could reduce the effects of the linearity assumption.

Reply 5: Thank you for your professional suggestion. As you noted, the electrode bending condition is actually quite clinically. The local maxima algorithm would be affected by the quality of CT images, since the metal contact often introduce heavy artifact on the CT images. In addition, the single contact from the depth electrodes was not necessarily discernible from the CT images when the electrode posed at certain angle to the CT scanning slices. Therefore, we think the local maxima algorithm is probably more appropriate for the conventional subdural electrodes, of which the contacts and the inter-contact space is often larger. Besides, the linearity assumption could avoid unnecessary jitters and better fitting the realistic contact distribution. We made a brief explanation in the discussion section.

Changes in the text:
Discussion line 361-362
Besides, the linearity assumption could avoid unnecessary jitters and better fitting the realistic contact distribution.

6. "Cluster thresholding" was confusing to me - it sounds like the number of clusters was simply pre-specified as the known number of electrodes implanted.

Reply 6: Thank your points it out. The CT thresholding means the intensity thresholding and number of clusters was pre-defined. We clarified the wording and rephrased this part according to your suggestion.

Changes in the text:
CT thresholding was changed to CT intensity thresholding

Method line 196-198:
Then we constructed clusters from the hierarchical cluster tree, starting from the root until the number of total clusters equal to the number of electrode shafts (figure 2).

7. Page 9, line 279, "not all the clustering was perfect" - can you provide more information about clustering accuracy?

Reply 7: Thank you for asking. Imperfect clustering could occur as the samples from one depth electrodes be assigned to two different clusters depending on the spatial distribution and intensity quality from the CT images, we further clarified the issue in the results section. And this is not problematic since only two points acting as target and entry in one electrode shaft would be needed. In addition, we provided manual correction to cope with rare complicated situations.

Changes in the text:
Results line 281-286
Since not all the clustering were perfect, for example, the samples from one depth electrodes were assigned to two different clusters in rare cases, we also designed manual correction mechanism allowing for sample points correction with large freedom, and alternatively, the users could also choose other samples in certain electrode shaft as the point or entry points.