Reviewer’s report

Title: A Combination of 3-D Discrete Wavelet Transform and 3-D Local Binary Pattern for Classification of Mild Cognitive Impairment

Version: 0 Date: 30 Jun 2019

Reviewer: Bernd Taschler

Reviewer's report:

The authors propose an interesting classification and feature construction method for MCI/AD converters and cognitively normal controls based on a combination of discrete wavelet transform and local binary patterns for 3D MRI data. Accurate, image-based classification in this area is still a challenge and a method for the extraction of a (small) set of highly informative features would be highly desirable.

However, there are several key issues that need to be addressed to strengthen the present manuscript.

## Major

[background] The abstract and background sections contain some inaccuracies:

- [line 32] "The detection of AD in formative stages especially MCI, makes the treatment effective and competent." This is a strong and as of yet incorrect claim since there is no (effective) treatment for AD.

- [line 67f] The statement of MCI-to-AD conversion should be based on annual rates. (See, e.g., Ward et al. 2013 in Dementia and Geriatric Cognitive Disorders).

- [line 95] "ML based diagnosis of MCI is better than the conventional methods [...]". This is not quite accurate, since there exists no diagnostic ML tool that can substitute for a "conventional" exam by an expert. (Maybe here the authors mean that ML methods are better in classification tasks compared to, e.g., logistic regression?)

- The authors mention the distinction between amnestic and non-amnestic MCI but it is unclear how this relates to their data and analysis.

- [line 88-91] ".. brain imaging techniques have been established for the diagnosis of MCI." This is misleading. As can be seen from the cited references, brain imaging has been used to classify MCI vs AD, MCI vs controls, etc. but not for diagnosis.

- [line 105]. Ref. 32 examines changes in atrophy rates but does not deal with diagnosis of MCI.
There seems to be a confusion about MCI subgroups and what they mean:

- [abstract] The first paragraph mentions "MCI-controls" and "MCI-non-controls". I presume the authors instead mean MCI-to-AD "converters" and "non-converters".

- [line 163] The text mentions subjects "converted to MCI" and "not converted to MCI". Again, I suspect the authors mean conversion from MCI to AD.

[Data selection and pre-processing] These sections would benefit from a more detailed description, especially regarding the pre-processing workflow of MRI data by addressing for example the following points:

- Subject selection: How were subjects selected from the ADNI data base? Why the relatively small numbers? (For example, the cited paper by Chincarini et al. 2011 uses 302 MCI patients and 189 age-matched controls.)

- Image acquisition: What imaging modality, acquisition parameters, etc. were used? (A relevant reference would probably be sufficient here.)

- Standardisation: Affine registration? Registration to the MNI template?

- Was a resampling of voxel size employed to make them isotropic? (e.g. to 1mm^3?)

- Why was a slice time correction performed? This is used in fMRI but is not necessary in sMRI.

- Tissue segmentation: Was this based on voxel based morphometry? What software was used?

[Methods] The description of 3D discrete wavelet transform (3D-DWT) is difficult to follow and lacking explanations:

- What wavelet function and thresholding were used in the present study?

- Most of the terms and variables used in Eqs. (1-15) are not explained. Either a comprehensive technical description or a high-level, verbal explanation (or both) would strengthen this section and make it easier to understand for the reader.

[Methods] 3D variant of local binary patterns (LBP):

- It is unclear to me how the authors arrive at 10 neighboring voxels in 3D. (Taking only faces into account, there are 6 neighbors. Considering also edges would result in 18 neighbors and adding corners would give 26 neighbors.)

[Methods] The section on pre-processing mentions image segmentation into GM, WM and CSF.
- How are these segmentations used in the feature construction?

- Are the proposed methods run on each segmented image separately?

[methods] Cross-validation (CV):

- Since performance measures (acc, sens, spec) are obtained from CV, the reported values should include errorbars (Table 2, Fig. 3).

- I would very strongly suggest the use of a nested CV procedure: Within each fold of the outer loop, another CV loop should be used to estimate model parameters and perform feature reduction (see e.g. Chincarini et al 2011, Fig. 7 for a schematic depiction). The reported maximum accuracy over all possible numbers of features is very likely due to some level of overfitting.

- Additionally, for MCI vs CN classification: Since there are only 45 CN subjects, I would suggest stratified CV-folds to ensure that each fold contains at least a few CN.

[Table 3] Comparison with other studies: There seems to be a mix-up in the numbers reported for other studies. A few examples:

- Gerardin et al 2009 do not classify MCI-C vs MCI-NC but rather MCI vs controls (and AD vs controls).

- The stated sensitivity and specificity numbers for Chincarini et al 2011 are actually reversed in the original paper.

- The numbers reported for Dai et al 2013 do not appear in the original paper. (In fact, Dai et al 2013 don't consider MCI-C vs MCI-NC.)

- The same holds for Khedher et al 2015.

- The numbers reported for Suk et al 2014 do not match the original paper.

- etc.

## Minor

- The title is somewhat misleading since the paper focuses on classification rather than prediction.

- Figure 1 is never referenced or explained in the main text.

- The numbers in Table 1 (male/female) for CN don't add up to n=45.
- The expression in Eq. (20) is incomplete as it does not seem to depend on the feature index i.

- Figures 3 and 4 show the same results, suggesting that one could be omitted.

- Figure 6 is never referenced or explained in the main text.

**Are the methods appropriate and well described?**
If not, please specify what is required in your comments to the authors.

No

**Does the work include the necessary controls?**
If not, please specify which controls are required in your comments to the authors.

Yes

**Are the conclusions drawn adequately supported by the data shown?**
If not, please explain in your comments to the authors.

Yes

**Are you able to assess any statistics in the manuscript or would you recommend an additional statistical review?**
If an additional statistical review is recommended, please specify what aspects require further assessment in your comments to the editors.

I am able to assess the statistics

**Quality of written English**
Please indicate the quality of language in the manuscript:

Needs some language corrections before being published

**Declaration of competing interests**
Please complete a declaration of competing interests, considering the following questions:

1. Have you in the past five years received reimbursements, fees, funding, or salary from an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?

2. Do you hold any stocks or shares in an organisation that may in any way gain or lose financially from the publication of this manuscript, either now or in the future?
3. Do you hold or are you currently applying for any patents relating to the content of the manuscript?

4. Have you received reimbursements, fees, funding, or salary from an organization that holds or has applied for patents relating to the content of the manuscript?

5. Do you have any other financial competing interests?

6. Do you have any non-financial competing interests in relation to this paper?

If you can answer no to all of the above, write 'I declare that I have no competing interests' below. If your reply is yes to any, please give details below.

I declare that I have no competing interests.

I agree to the open peer review policy of the journal. I understand that my name will be included on my report to the authors and, if the manuscript is accepted for publication, my named report including any attachments I upload will be posted on the website along with the authors' responses. I agree for my report to be made available under an Open Access Creative Commons CC-BY license (http://creativecommons.org/licenses/by/4.0/). I understand that any comments which I do not wish to be included in my named report can be included as confidential comments to the editors, which will not be published.

I agree to the open peer review policy of the journal.