Reviewer’s report

Title: Comparing radiomic classifiers and classifier ensembles for detection of peripheral zone prostate tumors on T2-weighted MRI: A multi-site study

Version: 0 Date: 09 Oct 2018

Reviewer: Daniel Hausmann

Reviewer's report:

Review:

Manuscript Number: BMIM-D-18-00141

Title:

Comparing radiomic classifiers and classifier ensembles for detection of peripheral zone prostate tumors on T2-weighted MRI: A multi-site study

Summary:

The authors performed this study to identify the optimal classifiers for a computer-aided diagnosis (CAD) of prostate cancer. The authors compared the performance of 12 ((4 classifier families (QDA, Bayesian learners, Decision Trees, and Support Vector Machines)) different supervised classier schemes in terms of classification accuracy as well as execution time. The authors conclude that simpler classifiers (such as QDA and its ensemble variants) may be more robust, accurate, and efficient for prostate cancer CAD problems, especially in the context of multi-site validation.

General:

The content of the manuscript is generally interesting as use of artificial intelligence gains increasing interest, especially in the context of prostate MRI. Choice of classifier to provide an accurate diagnosis is crucial and is still widely discussed. Please check spelling throughout the manuscript.
Keywords:
Adequate

Abstract:
Generally well-written.

Comment 1:
Please try to comprehensively and precisely state your purpose. Description of purpose/methods could be significantly shortened. Consider organizing your abstract as usual (Purpose, Methods, Results and Conclusion)

Comment 2:
Several misspellings were observed (e.g.: significantly"; "identied"; "classier").

Introduction:
 Interesting, relevant literature cited. Again, I believe the introduction could be shortened (1 page).

Comment 1:
"Both the choice of classier as well as classier trends in different CAD problems using medical imaging cohorts [3, 4, 19] tend to mirror conclusions arrived at from previous large-scale comparison studies involving natural and synthetic data"

Please check spelling and grammar.

Comment 2:
"In this work we aim to compare classier performance in the specific context of voxel-wise detection of tumors in the peripheral zone (PZ) of the prostate via "radiomic" texture-based features derived from T2w MRI."
Did you consider using other components of the multiparametric protocol (e.g. DWI) for Pca detection, taking into account that T2w is not even relevant to classify peripheral zone lesions according to PI-RADSv2 due to large overlap of morphologic features with chronic prostatitis? Please discuss.

Comment 3:

"They achieved relatively high AUCs for identifying which expert-delineated lesions within the PZ had a Gleason score of at least 7, using a linear mixed model (per-site AUC of 0.85 and 0.90)".

In this context, I believe that low-grade tumors (<=7a) are morphologically almost indistinguishable from chronic prostatitis, which is present in most patient at a certain age.

Comment 4:

"Multi-site data suffers from differences in scanners, acquisition protocols, and resolution differences - all of which can affect classifier performance, and hence choice of classifier, significantly."

Will this be a problem in the future? Scanner equipment and sequences will always differ between vendors in the closer future, so a robust classifier to overcome these differences is desirable. Please explain.

Experimental Design:

Comment 1:

"A total of 86 patient datasets were considered in this study."

The collective still appears relatively small for your purpose. Please discuss.

Comment 2:

"For all 86 datasets considered, the central gland and the PZ were annotated on the the axial endorectal T2w MRI image by a radiologist (a different expert annotated data from each institution)."
What do you mean by endorectal T2w? Was an endorectal coil used? How old is your data? Use of an endorectal coil is not considered state-of-the-art anymore, since MR-spectroscopy was omitted.

Comment 3: "For the purposes of this study, only PZ regions within the midgland region of the prostate were considered in all 86 datasets" 

How many tumors were present in the peripheral zone within the midgland region in your small collective? How was the distribution of Gleason Scores?

Results and Discussion: 

Comment 1: „voxel-wize"

Please check spelling.

Comment 2: 

"We identified the most consistently performing method across all 3 sites as the boosted QDA classifier. It had a relatively high AUC (= 0.735) in the training cohort as well as in both validation cohorts (average AUCs of 0.683 and 0.768, respectively)."

AUCs still appear quite low compared to conventional reporting using multiparametric data (T2, MRI perfusion/diffusion). Please discuss.

Tables:

Adequate

Figures:

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

Yes

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 recommend additional statistical review

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