ANSWERING REVIEWERS

December 18, 2018

Dear Editor,

Please find enclosed the edited manuscript in Word format (file name: WJG 44701 revision.doc).

Title: Artificial intelligence in medical imaging of liver

Author: Li-Qiang Zhou, Jia-Yu Wang, Song-Yuan Yu, Ge-Ge Wu, Qi Wei, You-Bin Deng, Xing-Long Wu, Xin-Wu Cui, Christoph F. Dietrich

Name of Journal: *World Journal of Gastroenterology*

ESPS Manuscript NO: 44701

We are very grateful about the reviewers' valuable comments. We believe that these comments have made our manuscript more comprehensible. We have made all the changes and marked them in the main text as required.

Reviewer 00646357:

-Discuss other types of deep learning other than CNN;

Reply: There are many models for deep learning, such as RNN, ANN, FCN, etc. CNN is currently the most widely used model. At present, most of the applications we have retrieved for the imaging diagnosis and evaluation of liver diseases are CNN models, and also ANN model, which was re-described in the last paragraph of the 'Focal liver lesion (FLL) detection' section. (Page 7, line 8)

In addition, we added the application of automatic segmentation of liver tumors from multiphase contrast-enhanced CT images based on FCNs in the 'Segmentation' section. (Page 8, line 37-39)

-Discuss role of deep learning in diffuse liver disease;

Reply: The role of deep learning in the detection of fatty liver, availability has been discussed in the section of 'Focal liver lesion (FLL) detection', and the role of deep learning in the evaluation of diffuse liver steatosis has been discussed in paragraph 3 of 'Focal liver lesion evaluation' section. We have further separated the part of deep learning for staging of liver fibrosis diseases and re-titled as 'diffuse liver disease staging'. (Page 7, line 14-15)

-Discuss role of deep learning in quantitative assessment of DWI parameters using these references:

Razek AAKA, Abdalla A, Barakat T, El-Taher H, Ali K. Assessment of the liver and spleen in children with Gaucher disease type I with diffusion-weighted MR imaging. Blood Cells Mol Dis 2018; 68:139-142.

Razek AA, Massoud SM, Azziz MR, El-Bendary MM, Zalata K, Motawea EM. Prediction of esophageal varices in cirrhotic patients with apparent diffusion coefficient of the spleen. Abdom Imaging 2015; 40:1465-9.

Razek AA, Khashaba M, Abdalla A, Bayomy M, Barakat T. Apparent diffusion coefficient value of hepatic fibrosis and inflammation in children with chronic hepatitis. Radiol Med 2014; 119:903-9.

Razek AA, Abdalla A, Omran E, Fathy A, Zalata K. Diagnosis and quantification of hepatic fibrosis in children with diffusion weighted MR imaging. Eur J Radiol 2011; 78:129-34.

Reply: We are very grateful for the detailed literatures you recommended. We have carefully read the given references, but found that they are not associated with deep learning. We have also further searched for the application of deep learning in quantitative assessment of DWI parameters in liver disease, and did not retrieve relevant literatures.

Reviewer 00039518:

The paper “Artificial intelligence in medical imaging of liver” is well written and provides an exhaustive review of the present state, ways of development, gray areas and limits of the application of artificial intelligence and convolutional neural networks in the field of liver imaging.

Reply: Thank you very much.

Only minor points need to be clarified:

- Input data and teaching data paragraph: explain the meaning of the acronym RGB.

Reply: We corrected it. We delete the “RGB” and changed the description to “R(red)、G(green)、B(blue)”. (Page 5, line 3)

- Input data and teaching data paragraph: write as follows: … the risk of the overfitting problem, because the slight differences in position may lead to the inconsistency between examinations.

Reply: We corrected it as you suggested. (Page 5, line 8-9)

-Focal liver lesion evaluation paragraph: this paragraph concerns not only focal liver lesions evaluation but also the application of CNN in the staging of diffuse liver disease. Please clarify and change the title of the paragraph.

Reply: We have separated the part of deep learning for staging of liver fibrosis diseases and re-titled as ‘diffuse liver disease staging’. (Page 7, line 14-15)

Reviewer 00503601:

This is a review article looking at the sphere of AI in medical imaging of the liver and its potential applications. Overall, the manuscript is difficult to read and follow as the authors have primarily taken findings from various studies and published works and aggregated these. The technical aspects tend to be fairly complex and may make little sense to the general reader.

Reply: At present, physicians usually detect, characterize and monitor diseases by assessing liver medical images visually, and such visual assessment which is based on expertise and experience may be personal and inaccurate. This problem does exist in clinical work. As an emerging technology, recently AI has developed rapidly in the field of imaging applications. It can make a quantitative assessment to assist physicians to make more accurate and reproductive imaging diagnosis. This review aims to show that AI brings great convenience to clinical work from the published liver imaging literatures and call for physicians embrace AI. The technical aspects of AI are indeed complicated, but introducing the part of basic principles and processes of AI work may make the readers better understand and further learn to use AI. We also polished our manuscript again to make it easier to read.

It will be useful for the authors to make mention of how AI compares to the current practiced model of clinician interpretation of the images but there is no data nor description of this comparison. The various clinical uses tend to be fairly generic and there is no specific detailing of where AI really has superiority over conventional clinical reporting.

Reply: The superiority of AI over current practiced model of clinician interpretation of the images is mainly the higher diagnosis accuracy and we descripted it in several paragraphs of the ‘clinical applications’ section. We rewrote and added a more detailed description of this part, as you suggested.

“They found that the deep learning method achieved an overall accuracy 97.2% compared with the accuracy of multi-SVM, KNN, and Naive Bayes, which are 96.5, 93.6, and 95.2%, respectively.” (Page 7, line 6-7)

“In the training cohort, AUCs of DLRE for F4, $\geqslant$ F3, $\geqslant$ F2 were 1.00 (0.99 -1.00), 0.99 (0.97-1.00) and 0.99 (0.97- 1.00), respectively, which were 0.13, 0.18 and 0.25 higher than these of 2D-SWE.” (Page 7, line 31-32)

“Comparing with the hepatorenal index and the gray-level co-occurrence matrix algorithm, which the accuracy is

90.9% and 85.4%, the CNN-based approach achieved significantly better results, with the AUC 0.977, sensitivity 100%, specificity 88.2%, accuracy 96.3%." (Page 8, line 10-12)
"According to Vivantil et al. by using deep learning models based on longitudinal liver CT studies new liver tumors could be detected automatically with the true positive rate of 86%, while the stand-alone detection rate was only 72% and this method achieved a precision of 87%, an improvement of 39% over the traditional SVM mode." (Page 6, line 32-33)
"Compared with two kinds of low level feature extraction method histogram of oriented gradients (HOG) and local binary pattern (LBP), which the mean accuracy is 83.6% and 81.4%, the deep learning method achieved a better classification accuracy of 86.9%." (Page 6, line 43 and Page 7, line 1-2)
"DCCA-MKL achieved 17.31%, 10.45%, 24.00%, 34.44%, 24.00% and 10.45% improvements over A-P-SVM, on classification accuracy, sensitivity, specificity, Youden index, false positive rate and false negative rate, respectively." (Page 8, line 20-21)
In addition, AI's superiority is reflected in the ability to accurately predict treatment response, which is difficult for doctors to accomplish, availability has been discussed in the section of 'Treatment response prediction'.

The conclusion also is somewhat bold in stating that AI is rapidly becoming state of the art without having any evidence eo back this up as the published works reviewed tend to be in the experimental context still, except perhaps for liver segmentation which is currently often automated or semi-automated.
Reply: Deep learning algorithms have made great breakthroughs in accuracy, efficiency and scope of application. This review introduced the superiority of AI by taking the application of deep learning in liver disease as an example. With the continuous promotion of deep neural network methods and its deepening in clinical research, we believe that artificial intelligence will be an indispensable helper for doctors in the future. As Professor Curtis Langlotz, vice president of the Radiological Society of North America (RSNA), said, "Artificial intelligence does not replace radiologists, but radiologists who embrace artificial intelligence will inevitably replace radiologists who resist artificial intelligence. ". Our previous statement may be too absolute, and we modified 'becoming the state of the art algorithms' to 'becoming an extremely promising aid'. Thank you very much for your comments.

The tables presented have no contextual relevance as there is no comparison to current conventional methods.
Reply: The table summarized the good performance of AI in the field of liver imaging according to different applications and the corresponding comparisons to current conventional methods have been discussed in the original text, therefore, we did not show them in the table.

Thank you very much again.

Sincerely yours,

Prof. Xin-Wu Cui