Baishideng Publishing Group

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

PEER-REVIEW REPORT

Name of journal: *World Journal of Gastroenterology*

Manuscript NO: 70291

Title: Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach

Provenance and peer review: Unsolicited manuscript; Externally peer reviewed

Peer-review model: Single blind

Reviewer's code: 05140714
Position: Peer Reviewer
Academic degree: PhD
Professional title: Postdoctoral Fellow, Research Fellow, Teaching Assistant
Reviewer's Country/Territory: Italy
Author's Country/Territory: South Korea

Manuscript submission date: 2021-07-29

Reviewer chosen by: AI Technique

Reviewer accepted review: 2021-07-29 07:06

Reviewer performed review: 2021-08-02 12:32

Review time: 4 Days and 5 Hours

Scientific quality	[ ] Grade A: Excellent [Y] Grade B: Very good [ ] Grade C: Good [ ] Grade D: Fair [ ] Grade E: Do not publish
Language quality	[Y] Grade A: Priority publishing [ ] Grade B: Minor language polishing [ ] Grade C: A great deal of language polishing [ ] Grade D: Rejection
Conclusion	[ ] Accept (High priority) [ ] Accept (General priority) [Y] Minor revision [ ] Major revision [ ] Rejection
Re-review	[Y] Yes [ ] No

Baishideng Publishing Group

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

Peer-reviewer statements	Peer-Review: [Y] Anonymous [ ] Onymous Conflicts-of-Interest: [ ] Yes [Y] No

SPECIFIC COMMENTS TO AUTHORS

The authors present a DL-based method for predicting mutations in gastric cancer using histopathological images. The manuscript is well written and results are sound. Statistical analysis is appropriate. In my opinion, deep learning could represent a time- and cost-efficient tool for mutation detection in pathology. My major comments are listed below: 1) Dataset composition: the dataset is highly unbalanced, so the authors randomly selected patients without mutations in order to balance the dataset. Personally, I would have selected all patients without mutations (perhaps using fewer patches from each patient) so as to train CNN on a more heterogeneous dataset. 2) Network training - Page 12, Line 7: ". The same label for all tumor tissue patches in a WSI as either 'wild-type' or 'mutated' were assigned based on the mutational status of the patient." Slide-level classification is very different from patch-level classification. Even if a WSI is labeled as "mutated," it is not certain that all its tumor patches contain features related to the gene mutation. This means that the network may accept patches that are labeled as "mutated" (because they come from a "mutated" WSI) but do not actually contain any alterations. This may represent a bias during network training. Other comments: - Page 11, Line 3: "We divided a WSI into non-overlapping patches of 360×360 pixel tissue images at 20× magnification to detect mutational status. " How were these parameters chosen? - Page 11,Line 13: "Morphologic features reflecting mutations in specific genes might be expressed mainly in tumor tissues rather than normal tissues." Please add at least one or two reference sfor this sentence. - more details about the training should be provided (optimiser, transfer learning yes/no, number of epochs, early stopping criteria, etc.) - Page 12, Line 19: "Color normalization was applied to the tissue patches to avoid

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

the effect of stain differences.". Recent studies have shown that stain normalization is an effective preprocessing step to build reliable deep learning frameworks in digital pathology (doi: 10.1016/j.compbiomed.2020.104129, doi:10.1038/s41598-020-71420-0). At least the one reference is needed for this sentence. - What is the overall accuracy of the cancer detection system? Were the same images from this study used to train the classifier? (normal/tumor classifier)

Baishideng Publishing Group

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

PEER-REVIEW REPORT

Name of journal: *World Journal of Gastroenterology*

Manuscript NO: 70291

Title: Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach

Provenance and peer review: Unsolicited Manuscript; Externally peer reviewed

Peer-review model: Single blind

Reviewer's code: 05735339
Position: Peer Reviewer
Academic degree: PhD

Professional title: Assistant Professor

Reviewer's Country/Territory: India

Author's Country/Territory: South Korea

Manuscript submission date: 2021-07-29

Reviewer chosen by: AI Technique

Reviewer accepted review: 2021-07-29 04:45

Reviewer performed review: 2021-08-06 10:29

Review time: 8 Days and 5 Hours

Scientific quality	[ ] Grade A: Excellent [Y] Grade B: Very good [ ] Grade C: Good [ ] Grade D: Fair [ ] Grade E: Do not publish
Language quality	[ ] Grade A: Priority publishing [Y] Grade B: Minor language polishing [ ] Grade C: A great deal of language polishing [ ] Grade D: Rejection
Conclusion	[ ] Accept (High priority) [ ] Accept (General priority) [Y] Minor revision [ ] Major revision [ ] Rejection
Re-review	[Y] Yes [ ] No

Baishideng Publishing Group

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

Peer-reviewer statements	Peer-Review: [Y] Anonymous [ ] Onymous Conflicts-of-Interest: [ ] Yes [Y] No

SPECIFIC COMMENTS TO AUTHORS

This study demonstrated that the CDH1, ERBB2, KRAS, PIK3CA, and TP53 mutations can be predicted from GC H&E WSIs using DL-based classifiers. The ROC curves and their AUCs were presented to analyze the performance of developed model. The AUCs for ROC curves ranged from 0.727 to 0.862 for the TCGA frozen WSIs and from 0.661 to 0.858 for the TCGA FFPE WSIs shows the effectiveness of extensive work. The work is to be accepted for the publication after the revision of the following comments: a. Author's has to highlights the major contributions of the manuscript in introduction section. Also briefly describe the flow of manuscript for the improvement in readability of the article. b. Why you guys are selecting Deep learning model rather than Machine learning model. You have to also explain the reason to choose Inception model in place of other pre-trained model. c. In this work Authors are simply importing the pre-trained CNN model. It shows that the novelty of the work is missing in case of model development. d. As we know that the deep learning model perform better for large dataset. In this case you are using very less amount of data. In this case model suffers from under fitting. So author's has requested to justify that your model is not suffering from under fitting. e. For the performance measuring, ROC is not only the sufficient approach. Specially for medical science research, you have to also perform statistical so that the significance of the designed model can be verified properly. f. Authors have not listed the social impact of the study. You have to also mention it. g. How this work can be extended further? h. Authors can improve the literature work by adding some quality work like • Echle, Amelie, Niklas Timon Rindtorff, Titus Josef Brinker, Tom Luedde, Alexander Thomas Pearson, and Jakob Nikolas Kather. "Deep learning in cancer pathology: a new

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

generation of clinical biomarkers." British journal of cancer 124, no. 4 (2021): 686-696. • Calderaro, Julien, and Jakob Nikolas Kather. "Artificial intelligence-based pathology for gastrointestinal and hepatobiliary cancers." Gut 70, no. 6 (2021): 1183-1193. • Coudray, Nicolas, and Aristotelis Tsirigos. "Deep learning links histology, molecular signatures and prognosis in cancer." Nature Cancer 1, no. 8 (2020): 755-757. • Bhatt, Chandradeep, Indrajeet Kumar, V. Vijayakumar, Kamred Udham Singh, and Abhishek Kumar. "The state of the art of deep learning models in medical science and their challenges." Multimedia Systems (2020): 1-15.

Baishideng Publishing Group

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

PEER-REVIEW REPORT

Name of journal: *World Journal of Gastroenterology*

Manuscript NO: 70291

Title: Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach

Provenance and peer review: Unsolicited Manuscript; Externally peer reviewed

Peer-review model: Single blind

Reviewer's code: 05128846
Position: Peer Reviewer
Academic degree: MD

Professional title: Doctor

Reviewer's Country/Territory: United Kingdom

Author's Country/Territory: South Korea

Manuscript submission date: 2021-07-29

Reviewer chosen by: AI Technique

Reviewer accepted review: 2021-08-02 00:30

Reviewer performed review: 2021-08-11 02:09

Review time: 9 Days and 1 Hour

Scientific quality	[ ] Grade A: Excellent [ ] Grade B: Very good [Y] Grade C: Good [ ] Grade D: Fair [ ] Grade E: Do not publish
Language quality	[ ] Grade A: Priority publishing [ ] Grade B: Minor language polishing [Y] Grade C: A great deal of language polishing [ ] Grade D: Rejection
Conclusion	[ ] Accept (High priority) [ ] Accept (General priority) [ ] Minor revision [Y] Major revision [ ] Rejection
Re-review	[Y] Yes [ ] No

Baishideng Publishing Group

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

Peer-reviewer statements	Peer-Review: [Y] Anonymous [ ] Onymous Conflicts-of-Interest: [ ] Yes [Y] No

SPECIFIC COMMENTS TO AUTHORS

In the study entitled "Prediction of genetic alterations from gastric cancer histopathology images using a fully automated deep learning approach" Jang and colleagues assess the potential of deep learning for the automated prediction of genetic alterations in H&E images from gastric cancer. The use of AI in pathology is a hot topic at the moment. They used images from the TCGA dataset to develop an algorithm that they then test on new unseen cases from Seoul St. Mary's Hospital. Comments: Although the problem and significance of this study are clearly exposed the manuscript requires significant language editing which would enable the reader to understand better the study. The number of patient samples assessed are limited which as the authors suggest affected the generalizability of their developed classifiers. Although I agree that heterogeneity exists within cancer tissues from different countries, hospitals etc. and hence a mixed classifier which has trained on both the TCGS and the SSMH datasets would be more appropriate, the authors should have kept a small number of cases as their testing or validation cohort. I recommend either increasing the number of cases included in the study by specifically assessing their developed classifiers in an unseen dataset or that they divide their current dataset into a training and test/validation set. Page 11 "Deep learning model": "For the tumor or mutation classifiers described below, only proper tissue patches were analyzed" what do the authors mean by proper? This sentence should be edited to make sure others can reproduce the work if they wanted to. Page 10 "After reviewing the quality of WSIs from the GC dataset of TCGA (TCGA-STAD), we selected slides from 25, 19, 34, 64, and 160 patients, which were confirmed to have mutations in CDH1, ERBB2, KRAS, PIK3CA, and TP53 genes,

7041 Koll Center Parkway, Suite 160, Pleasanton, CA 94566, USA
Telephone: +1-925-399-1568
E-mail: bpgoffice@wjgnet.com
https://www.wjgnet.com

respectively. “Do these numbers represent patient numbers as well? Or are there slides which are from the same patients, just different blocks? Details regarding the selection criteria of their images should be provided and the logic behind their selections should be discussed. Page 12 “Mutation classifiers were trained separately for the selected tumor patches for frozen and FFPE tissues”, details regarding the two classifiers (frozen and FFPE) should be discussed (ie. number of patients, slides, training regions etc). How many patches per image were they assessed on average? This information would be good to be added. How long were the classifiers trained for? Further details regarding their training process is crucial for reproduction. Overall, although this study and its topic of high importance, substantial edits need to be performed with major modifications of their materials and methods section.