PAIP2021: Perineural Invasion in Multiple Organ Cancer (Colon, Prostate, and Pancreatobiliary tract): Structured description of the challenge design

CHALLENGE ORGANIZATION

Title

Use the title to convey the essential information on the challenge mission.

PAIP2021: Perineural Invasion in Multiple Organ Cancer (Colon, Prostate, and Pancreatobiliary tract)

Challenge acronym

Preferable, provide a short acronym of the challenge (if any).

PAIP2021

Challenge abstract

Provide a summary of the challenge purpose. This should include a general introduction in the topic from both a biomedical as well as from a technical point of view and clearly state the envisioned technical and/or biomedical impact of the challenge.

A distinct feature of malignant tumors compared to normal tissue or benign tumors is the invasiveness of transformed cells. Malignant tumor cells can invade vessels, nerves, stroma (through the basement membrane), lymph nodes, and other distant organs. Metastasis to lymph nodes and distant organs is a unique feature of malignant tumors, and it affects patient survival. Nerves are normal structures which connect the brain to peripheral organs, and transmit electrical signals which are responsible for voluntary and non-voluntary motor control and sensory perception. Nerves and blood vessels are the main pathways through which malignant cells can move from a primary site to other organs. Perineural invasion by malignant tumor cells has been reported as an independent indicator of poor prognosis in cancers. The College of American Pathologists recommended the inclusion of perineural invasion in pathology reports for resected specimens of malignant tumors.
Detection of perineural invasion in small nerves on glass slides is a labor-intensive task. Histologically, perineural invasion is composed of nerve and tumor cells attached to the nerve, and the nerve may be completely enclosed within the tumor mass. The morphology of nerves is the same in all organs, but tumor cells have various forms according to the histologic type and organs they are present in, thus modelling cancer in every organ is difficult. In the case of similar histology among the different cancers, it is challenging to determine whether an algorithm can be made for common histologic elements even if the background organs are different.
This challenge aims to promote the development of a common algorithm for automatic detection of perineural invasion in resected specimens of multi-organ cancers. This challenge will have a technical impact in the following fields: detection of composite target (nerve and tumor) and common modeling for target images in multiple backgrounds. This challenge will provide a good opportunity to overcome the limitations of current disease-organ-specific modeling and develop a technological approach to the universality of histology in multiple organs.

Challenge keywords

List the primary keywords that characterize the challenge.

Perineural invasion, Multi-organ cancer, Carcinoma, Detection, Digital pathology

Year

The challenge will take place in ...

2021

FURTHER INFORMATION FOR MICCAI ORGANIZERS

Workshop

If the challenge is part of a workshop, please indicate the workshop.

None.

Duration

How long does the challenge take?

Half day.

Expected number of participants

Please explain the basis of your estimate (e.g. numbers from previous challenges) and/or provide a list of potential participants and indicate if they have already confirmed their willingness to contribute.

Over 400 of participants is expected based on the number of previous PAIP challenge (PAIP2020). Please refer to the following website and find this year's participants. https://paip2020.grand-challenge.org/rank/

Publication and future plans

Please indicate if you plan to coordinate a publication of the challenge results.

The organizers will be making available channels of scientific paper publication. For example, high ranked teams will be suggested to submit papers to Health Informatics Research(HIR, www.e-hir.org) which is a SCOPUS citation journal. The submission is optional, but the HIR journal will provide the fast track review.
Moreover, as a result of this challenge, organizers will publish a joint paper compiling the methodology of the top 10 participants.

Space and hardware requirements

Organizers of on-site challenges must provide a fair computing environment for all participants. For instance, algorithms should run on the same computing platform provided to all.

The Grand Challenge platform will be used for the submission of solutions. On the day of the challenge, a projector, a computer, two microphones, and maximum 50 pax capacity conference room will be needed in order to announce the outcomes of the challenge and have the participants present their proposed solutions.

TASK: Detection of perineural invasion in three organ cancers

SUMMARY

Abstract

Provide a summary of the challenge purpose. This should include a general introduction in the topic from both a biomedical as well as from a technical point of view and clearly state the envisioned technical and/or biomedical impact of the challenge.

A distinct feature of malignant tumors compared to normal tissue or benign tumors is the invasiveness of transformed cells. Malignant tumor cells can invade vessels, nerves, stroma (through the basement membrane), lymph nodes, and other distant organs. Metastasis to lymph nodes and distant organs is a unique feature of malignant tumors, and it affects patient survival. Nerves are normal structures which connect the brain to peripheral organs, and transmit electrical signals which are responsible for voluntary and non-voluntary motor control and sensory perception. Nerves and blood vessels are the main pathways through which malignant cells can move from a primary site to other organs. Perineural invasion by malignant tumor cells has been reported as an independent indicator of poor prognosis in cancers. The College of American Pathologists recommended the inclusion of perineural invasion in pathology reports for resected specimens of malignant tumors.
Detection of perineural invasion in small nerves on glass slides is a labor-intensive task. Histologically, perineural invasion is composed of nerve and tumor cells attached to the nerve, and the nerve may be completely enclosed within the tumor mass. The morphology of nerves is the same in all organs, but tumor cells have various forms according to the histologic type and organs they are present in, thus modelling cancer in every organ is difficult. In the case of similar histology among the different cancers, it is challenging to determine whether an algorithm can be made for common histologic elements even if the background organs are different.
This challenge aims to promote the development of a common algorithm for automatic detection of perineural invasion in resected specimens of multi-organ cancers. This challenge will have a technical impact in the following fields: detection of composite target (nerve and tumor) and common modeling for target images in multiple backgrounds. This challenge will provide a good opportunity to overcome the limitations of current disease-organ-specific modeling and develop a technological approach to the universality of histology in multiple organs.

Keywords

List the primary keywords that characterize the task.

Perineural invasion, Multi-organ cancer, Carcinoma, Detection, Digital pathology

ORGANIZATION

Organizers

a) Provide information on the organizing team (names and affiliations).

Jinwook Choi (Seoul National University Hospital, Seoul, South Korea)
Kyoungbun Lee, MD (Seoul National University Hospital, Seoul, South Korea)
Won-Ki Jeong (Korea University, Seoul, South Korea)
Se Young Chun (Ulsan National Institute of Science and Technology, Ulsan, South Korea)

b) Provide information on the primary contact person.

Jinwook Choi, jinchoi@snu.ac.kr

Life cycle type

Define the intended submission cycle of the challenge. Include information on whether/how the challenge will be continued after the challenge has taken place.Not every challenge closes after the submission deadline (one-time event). Sometimes it is possible to submit results after the deadline (open call) or the challenge is repeated with some modifications (repeated event).

Examples:

- One-time event with fixed conference submission deadline
- Open call (challenge opens for new submissions after conference deadline)
- Repeated event with annual fixed conference submission deadline

Repeated event with fixed submission deadline.

Challenge venue and platform

a) Report the event (e.g. conference) that is associated with the challenge (if any).

MICCAI.

b) Report the platform (e.g. grand-challenge.org) used to run the challenge.

grand-challenge.org

c) Provide the URL for the challenge website (if any).

TBA

Participation policies

a) Define the allowed user interaction of the algorithms assessed (e.g. only (semi-) automatic methods allowed).

Fully automatic.

b) Define the policy on the usage of training data. The data used to train algorithms may, for example, be restricted to the data provided by the challenge or to publicly available data including (open) pre-trained nets.

No additional data allowed.

c) Define the participation policy for members of the organizers' institutes. For example, members of the organizers' institutes may participate in the challenge but are not eligible for awards.

May not participate.

d) Define the award policy. In particular, provide details with respect to challenge prizes.

1st $1,000
2nd $500
Final TOP 10 participants will be announced and have a corresponding certificates will be awarded at the workshop.

e) Define the policy for result announcement.

Examples:

- Top 3 performing methods will be announced publicly.
- Participating teams can choose whether the performance results will be made public.

All submitted results will be evaluated and their ranks based on evaluation metrics published on the website. And the top 10 methods will be announced publicly, and they will be presented at the MICCAI2021 conference.

f) Define the publication policy. In particular, provide details on ...

- ... who of the participating teams/the participating teams' members qualifies as author
- ... whether the participating teams may publish their own results separately, and (if so)
- ... whether an embargo time is defined (so that challenge organizers can publish a challenge paper first).

From the participating teams, the first and last authors will qualify as author for the resulting publication. Participating teams are free to publish their own results whenever they want, however, they need to acknowledge with the grand number as follows. "De-identified pathology images and annotations used in this research were prepared and provided by the Seoul National University Hospital by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI18C0316)."

Submission method

a) Describe the method used for result submission. Preferably, provide a link to the submission instructions.

Examples:

- Docker container on the Synapse platform. Link to submission instructions: <URL>
- Algorithm output was sent to organizers via e-mail. Submission instructions were sent by e-mail.

Participants send the algorithm output to the organizers via the evaluation platform.
The participants are required to submit single channel files in tiff file format of same size as corresponding test case images. Each output file should be consisted of binary values where 1 indicates the output line and 0 indicates the background. The list of the final top 10 contestants will be asked to provide a short description or an abstract of their learning model. Multiple submissions are allowed once every day during the test phase.

b) Provide information on the possibility for participating teams to evaluate their algorithms before submitting final results. For example, many challenges allow submission of multiple results, and only the last run is officially counted to compute challenge results.

During the validation phase, the participants will be allowed to submit their validation results 10 times a day. They will be evaluated at the leaderboard in real time before their final submission.

Challenge schedule

Provide a timetable for the challenge. Preferably, this should include

- the release date(s) of the training cases (if any)
- the registration date/period
- the release date(s) of the test cases and validation cases (if any)
- the submission date(s)
- associated workshop days (if any)
- the release date(s) of the results

•Challenge start and release of training dataset (Mid-April)
•Validation dataset release and start of validation submission (June)
•Team merger deadline and end of validation phase (Early July)
•Test dataset release and start of final submission (Mid-July)
•End of submission (End of July)

•Top10 notification and invitation for workshop (Mid-August)
•Challenge Workshop in conjunction with MICCAI 2021: Results announced

Ethics approval

Indicate whether ethics approval is necessary for the data. If yes, provide details on the ethics approval, preferably institutional review board, location, date and number of the ethics approval (if applicable). Add the URL or a reference to the document of the ethics approval (if available).

All datasets were approved by Seoul National University Hospital Institutional Review Board (IRB) (IRB No.H-1808-035-964).

Data usage agreement

Clarify how the data can be used and distributed by the teams that participate in the challenge and by others during and after the challenge. This should include the explicit listing of the license applied.

Examples:

- CC BY (Attribution)
- CC BY-SA (Attribution-ShareAlike)
- CC BY-ND (Attribution-NoDerivs)
- CC BY-NC (Attribution-NonCommercial)
- CC BY-NC-SA (Attribution-NonCommercial-ShareAlike)
- CC BY-NC-ND (Attribution-NonCommercial-NoDerivs)

CC BY NC.

Code availability

a) Provide information on the accessibility of the organizers' evaluation software (e.g. code to produce rankings). Preferably, provide a link to the code and add information on the supported platforms.

Only the pre-processing code will be available for access through github.

b) In an analogous manner, provide information on the accessibility of the participating teams' code.

Participating teams' code will not be released for public access.

Conflicts of interest

Provide information related to conflicts of interest. In particular provide information related to sponsoring/funding of the challenge. Also, state explicitly who had/will have access to the test case labels and when.

This challenge was supported by a grant of the Korea Health Technology R&D Project through the Korea Health Industry Development Institute (KHIDI), funded by the Ministry of Health & Welfare, Republic of Korea (grant number: HI18C0316).
VUNO Inc. has accessed pancreatobiliary tract cases.

MISSION OF THE CHALLENGE

Field(s) of application

State the main field(s) of application that the participating algorithms target.

Examples:

- Diagnosis
- Education
- Intervention assistance
- Intervention follow-up
- Intervention planning
- Prognosis
- Research
- Screening
- Training
- Cross-phase

Decision support, Assistance, Research, Medical data management, Diagnosis, Treatment planning.

Task category(ies)

State the task category(ies).

Examples:

- Classification
- Detection
- Localization
- Modeling
- Prediction
- Reconstruction
- Registration
- Retrieval
- Segmentation
- Tracking

Detection.

Cohorts

We distinguish between the target cohort and the challenge cohort. For example, a challenge could be designed around the task of medical instrument tracking in robotic kidney surgery. While the challenge could be based on ex vivo data obtained from a laparoscopic training environment with porcine organs (challenge cohort), the final biomedical application (i.e. robotic kidney surgery) would be targeted on real patients with certain characteristics defined by inclusion criteria such as restrictions regarding sex or age (target cohort).

a) Describe the target cohort, i.e. the subjects/objects from whom/which the data would be acquired in the final biomedical application.

Hematoxylin eosin stained slides of resected specimen for colorectal, prostate and pancreatobiliary tract cancer.

b) Describe the challenge cohort, i.e. the subject(s)/object(s) from whom/which the challenge data was acquired.

Patients who were histologically diagnosed with ductal adenocarcinoma or adenocarcinoma of colorectum, prostate and pancreatobiliary tract. Scanned image data of resected tumor tissues of the colon, prostate, and pancreas diagnosed at SNUH, SNUBH, and SMG-SNU BMC from January 2005 to June 2019 will be provided. And all personal labels in scanned images were removed in order to protect the privacy.

Imaging modality(ies)

Specify the imaging technique(s) applied in the challenge.

All whole slide images were stained by hematoxylin and eosin and scanned by the Aperio AT2 at 20X magnification.

Context information

Provide additional information given along with the images. The information may correspond ...

a) ... directly to the image data (e.g. tumor volume).

Case number: randomly applied number after removing the labeling of the original specimen.
Pathological information: organ (colon, prostate, and pancreas), histology (adenocarcinoma)

b) ... to the patient in general (e.g. sex, medical history).

No additional information will be given along with the images.

Target entity(ies)

a) Describe the data origin, i.e. the region(s)/part(s) of subject(s)/object(s) from whom/which the image data would be acquired in the final biomedical application (e.g. brain shown in computed tomography (CT) data, abdomen shown in laparoscopic video data, operating room shown in video data, thorax shown in fluoroscopy video). If necessary, differentiate between target and challenge cohort.

Representative one scanned virtual slide in each case will be presented. They were histologically diagnosed as ductal adenocarcinoma or adenocarcinoma in resected specimen of colorectum, prostate, pancreatobiliary tract.

b) Describe the algorithm target, i.e. the structure(s)/subject(s)/object(s)/component(s) that the participating algorithms have been designed to focus on (e.g. tumor in the brain, tip of a medical instrument, nurse in an operating theater, catheter in a fluoroscopy scan). If necessary, differentiate between target and challenge cohort.

Detection of perinerual invasion in whole slide images of resected specimen. Definition of perineural invasion is nerve infiltrated or enclosed by tumor cells. Because whole nerve and infiltrating pattern of tumor cells are so diverse, final target at image level is boundary line between the nerve and contacting tumor.

Assessment aim(s)

Identify the property(ies) of the algorithms to be optimized to perform well in the challenge. If multiple properties are assessed, prioritize them (if appropriate). The properties should then be reflected in the metrics applied (see below, parameter metric(s)), and the priorities should be reflected in the ranking when combining multiple metrics that assess different properties.

- Example 1: Find highly accurate liver segmentation algorithm for CT images.
- Example 2: Find lung tumor detection algorithm with high sensitivity and specificity for mammography images.

Corresponding metrics are listed below (parameter metric(s)).

Precision, User satisfaction, Accuracy.

Additional points: Draw the touching lines where the perineural invasion has occurred.

DATA SETS

Data source(s)

a) Specify the device(s) used to acquire the challenge data. This includes details on the device(s) used to acquire the imaging data (e.g. manufacturer) as well as information on additional devices used for performance assessment (e.g. tracking system used in a surgical setting).

One representative whole slide images scanned by Leica Aperio AT2- Digital whole slide scanner at x20.

b) Describe relevant details on the imaging process/data acquisition for each acquisition device (e.g. image acquisition protocol(s)).

All whole slide images were stained by hematoxylin and eosin and scanned by the Aperio AT2 at 20X magnification.

c) Specify the center(s)/institute(s) in which the data was acquired and/or the data providing platform/source (e.g. previous challenge). If this information is not provided (e.g. for anonymization reasons), specify why.

Seoul National University Hospital (SNUH), Seoul National University Bundang Hospital (SNUBH), and SMG-SNU Boramae Medical Center (BRMH).

d) Describe relevant characteristics (e.g. level of expertise) of the subjects (e.g. surgeon)/objects (e.g. robot) involved in the data acquisition process (if any).

Expert pathologists in GI tract and urogenital tract pathology involved in annotation of boundary line of perineural invasion, normal nerve, tumor and non-tumor area.

Training and test case characteristics

a) State what is meant by one case in this challenge. A case encompasses all data that is processed to produce one result that is compared to the corresponding reference result (i.e. the desired algorithm output).

Examples:

- Training and test cases both represent a CT image of a human brain. Training cases have a weak annotation (tumor present or not and tumor volume (if any)) while the test cases are annotated with the tumor contour (if any).
- A case refers to all information that is available for one particular patient in a specific study. This information always includes the image information as specified in data source(s) (see above) and may include context information (see above). Both training and test cases are annotated with survival (binary) 5 years after (first) image was taken.

One WSI is a representative one section of several blocks of resected specimen holding tumor, nerve, normal tissue. Training data is provided with labelling perineural invasion (boundary line of nerve and tumor cells), tumor, non-tumor area, and nerve without tumor cell. Each case of test set is also one representative WSI obtained from one patients.

b) State the total number of training, validation and test cases.

The training dataset contains 150 WSIs (50 colorectum, 50 prostate, 50 pancreatobiliary tract).
The validation dataset contains 30 WSIs. (10 colorectum, 10 prostate, 10 pancreatobiliary tract).
The test dataset contains 60 WSIs. (20 colorectum, 20 prostate, 20 pancreatobiliary tract).

c) Explain why a total number of cases and the specific proportion of training, validation and test cases was chosen.

The number of training cases was decided based on previous PAIP 2019 liver cancer segmentation challenge. 60 cases including validation set was sufficient for developing segmentation model for cancer. A total of 180 cases were selected, consisting of 60 cases per three organs.

d) Mention further important characteristics of the training, validation and test cases (e.g. class distribution in classification tasks chosen according to real-world distribution vs. equal class distribution) and justify the choice.

In the training group, each case has more than one perineural invasion sites but test dataset is composed of regardless of existence of perineural invasions. Because the incidence of perineural invasion is different among organs, consecutive cases could not be selected. Cases without tumor were also included in the test dataset because routine glass slides of resected tumor sample in real-world practice include non-tumor tissue for histologic mapping or resection margin, or etc. None of datasets is overlapped with PAIP2020 Challenge and the validation and test datasets would be annotated respectively by 3-4 pathologists.

Annotation characteristics

a) Describe the method for determining the reference annotation, i.e. the desired algorithm output. Provide the information separately for the training, validation and test cases if necessary. Possible methods include manual image annotation, in silico ground truth generation and annotation by automatic methods.

If human annotation was involved, state the number of annotators.

Three to four expert pathologists in GI tract and genitourinary tract pathology manually annotate tumor, non-tumor, perineural invasion and nerve by Aperio image scope Ver.12.4.0.5.43.
To help understanding, we have included two figures by link. Please check images of hyperlink.

- Example: Fig 1 (https://drive.google.com/file/d/1bu-lFZe7_RpqvmKIT5xzFp4Mg2pcNrJs/view?usp=sharing)
- Example: Fig 2 (https://drive.google.com/file/d/1WshK-GXtZcYZcykX1lmKIRhJ22X86RAG/view?usp=sharing)

The annotation details are as follows:
- There are three box layers and one boundary line layer.
- Red line (layer 1) indicates perineural invasion.
- Yellow box (layer 2) indicates tumor free nerve tissue area.
- Green box (layer 3) indicates tumor area.
- Black box (layer 4) indicates no tumor area.

Training Cases
- Layer 1 : perineural invasion
Perineural invasion can be displayed as a line or a closed circle. A line shows a contact boundary between tumor and nerve tissue. A circle indicates tumor cells inside a nerve tissue or the other way around. The boundary between the tumor and the nerve tissue is depicted as a red line, which is the perineural invasion of tumor. (When you zoom in, you can see two red lines. Sorry for the small image!)
- Layer 2 : tumor free nerve tissue area
Nerve tissues not infiltrated by tumor cells are depicted in closed yellow lines. (See Fig. 2e)
- Layer 3 : tumor
Representative tumor area, not whole tumor area in slides.
More than one square box holding tumor cells with less than 5% of normal tissue or nerve.
- Layer 4: non-tumor

More than one square box not holding tumor cells, representative area for non-tumorous tissue, such as normal colon, prostate, pancreas, bile duct, liver around bile duct, or mesenteric or peripancreatic adipose tissue
cf. Layer 3 and 4 can be partially or completely overlapped in some cases.

Validation & Test Cases
Whole slide images without annotation will be provided for validation & test sets.
Detail information for the submission, please refer to submission method section.

b) Provide the instructions given to the annotators (if any) prior to the annotation. This may include description of a training phase with the software. Provide the information separately for the training, validation and test cases if necessary. Preferably, provide a link to the annotation protocol.

At least 10 normal nerves were included in Layer 2(normal nerve) and Layer 3(tumor area) was selected from viable tumor cells excluding necrosis, hemorrhage or inflammation.

c) Provide details on the subject(s)/algorithm(s) that annotated the cases (e.g. information on level of expertise such as number of years of professional experience, medically-trained or not). Provide the information separately for the training, validation and test cases if necessary.

Pathologists of 13 years and 18 years of experience in GI tract pathology and urogenital tract pathology.

d) Describe the method(s) used to merge multiple annotations for one case (if any). Provide the information separately for the training, validation and test cases if necessary.

None.

Data pre-processing method(s)

Describe the method(s) used for pre-processing the raw training data before it is provided to the participating teams. Provide the information separately for the training, validation and test cases if necessary.

Each image of training dataset is composed of one whole slide image (.svs) and annotation file (.xml) and test dataset images are single WSIs without annotation.

Sources of error

a) Describe the most relevant possible error sources related to the image annotation. If possible, estimate the magnitude (range) of these errors, using inter-and intra-annotator variability, for example. Provide the information separately for the training, validation and test cases, if necessary.

The length of boundary line between tumor and nerve (tissue) can be slightly different between human and algorithms. For example, tissue crack can be made between tumor cell and nerve due to tissue artifact or mucin, but human pathologist usually recognize it as perineural invasion.

b) In an analogous manner, describe and quantify other relevant sources of error.

None.

ASSESSMENT METHODS

Metric(s)

a) Define the metric(s) to assess a property of an algorithm. These metrics should reflect the desired algorithm properties described in assessment aim(s) (see above). State which metric(s) were used to compute the ranking(s) (if any).

- Example 1: Dice Similarity Coefficient (DSC)
- Example 2: Area under curve (AUC)

Hausdorff distance with cut-off threshold (curved line detection)
Detection F1 (detection)

Each curve segment will be represented using a bounding box. If two bounding boxes (one from GT, one from prediction) overlap more than 50%, compute curve-to-curve distance using a Hausdorff distance metric. If the curves are close enough (Hausdorff distance is less than the threshold), marked it as detected. Otherwise not detected. Based on this, we compute F1 score, which is used to determine the tentative ranking.

b) Justify why the metric(s) was/were chosen, preferably with reference to the biomedical application.

When a tumor region and a nerve region touch each other in a close distance, the corresponding location is identified as the perineural invasion. The perineural invasion is labeled as curved line in between tumor and nerve regions in close contact. The task is to find the perineural invasion regions by detecting such lines. To evaluate the output, each output line is to be matched with the closest label line. The degree of proximity is determined in terms of the distance between two lines. The output line of distance value less than a cut-off threshold value is decided as true positive and others over the threshold value as false positive. Then, true positive and false positive lines are counted for quantitative evaluation by calculating the detection score.
For better comparison among the top ranked solutions of similar detection counts, mean distance-based score of the true positive lines is computed to evaluate detection accuracy. To confirm the pathological significance, the pathologists will perform qualitative analysis on the solutions.

Ranking method(s)

a) Describe the method used to compute a performance rank for all submitted algorithms based on the generated metric results on the test cases. Typically the text will describe how results obtained per case and metric are aggregated to arrive at a final score/ranking.

For each submitted solutions, detection score and distance-based measures will be calculated by comparing the solutions on the test cases with the ground truth labels. The proposed evaluation metrics, which represents the performance of each algorithm, will be used to rank all the submitted algorithms. Then, among the top 10 ranked participants, the assigned experts will conduct qualitative analysis to verify the pathological significance, also working as tiebreakers among algorithms with similar quantitative measures.

b) Describe the method(s) used to manage submissions with missing results on test cases.

For submission results that missed the existing labels, the missing cases will be counted as false negative and will be considered in calculating the detection score.

c) Justify why the described ranking scheme(s) was/were used.

This ranking scheme effectively evaluates performance in the descending significance order of counting measure, quantitative detection score, and detection quality.

Statistical analyses

a) Provide details for the statistical methods used in the scope of the challenge analysis. This may include

- description of the missing data handling,
- details about the assessment of variability of rankings,
- description of any method used to assess whether the data met the assumptions, required for the particular statistical approach, or
- indication of any software product that was used for all data analysis methods.

Expert satisfaction rating
Distance measure based on cut-off threshold
Imagescope, ImageJ, ASAP

b) Justify why the described statistical method(s) was/were used.

The expert satisfaction rating will be used as measures for qualitative evaluation. Distance measure based on cut-off threshold is computed to compare the detection quality in quantitative measures. The following software products were used to view the input whole slide images in tiff file format.

Further analyses

Present further analyses to be performed (if applicable), e.g. related to

- combining algorithms via ensembling,
- inter-algorithm variability,
- common problems/biases of the submitted methods, or
- ranking variability.

Further analyses will include ranking variability in terms of the results on the test cases. The top 10 ranked algorithms will be qualitatively analyzed.

ADDITIONAL POINTS

References

Please include any reference important for the challenge design, for example publications on the data, the annotation process or the chosen metrics as well as DOIs referring to data or code.

None.

Further comments

Further comments from the organizers.

This challenge will be an opportunity developer to experience histology consisting of two or more cellular objects and to overcome the limitations of current disease-organ-specific modeling and think about a technological approach to university of histology in multiple organs.