Preoperative to Intraoperative Laparoscopy Fusion: Structured description of the challenge design

CHALLENGE ORGANIZATION

Title

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

Preoperative to Intraoperative Laparoscopy Fusion

Challenge acronym

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

P2ILF

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.

Augmented reality (AR) in laparoscopic liver surgery needs key landmark detection in intraoperative 2D laparoscopic images and its registration with the preoperative 3D model created from CT/MRI data. Such AR techniques are vital to surgeons as they enable precise tumor localisation for surgical removal. A full resection of targeted tumor minimises the risk of recurrence. However, the task of automatic anatomical curve segmentation (considered as landmarks), and its registration to 3D models is a non-trivial and complex task. Most developed methods in this domain are built around traditional methodologies in computer vision. This challenge is designed to challenge participants to deploy machine learning methods for two tasks - a) task I: segmentation of key anatomical curves from laparoscopic video images and 3D model, and b) task 2: matching these segmented curves to the 3D liver model from volumetric data (CT/MR). Thus the challenge is aimed at segmenting anatomical curves such as ridges, liver contours and midline of ligament (supervised and semi-supervised), and 2D-3D registration problem (semi-supervised and unsupervised) between the segmented landmarks with the provided dense 3D point cloud of liver. Here, we will assess the quality of registration algorithms using widely used target registration errors while liver landmark segmentation will be evaluated-based on F1-score.

Challenge keywords

List the primary keywords that characterize the challenge.

Liver anatomical curve segmentation; intraoperative and preoperative registration

Year

The challenge will take place in ...

2022

FURTHER INFORMATION FOR MICCAI ORGANIZERS

Workshop

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

EndoVis has agreed to integrate this challenge

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.

50-100

Publication and future plans

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

The results will be announced during the EndoVis Challenge at MICCAI'22. The results and rankings will be made publicly available on the challenge website. The submitted results will be further dissected through more rigorous statistical tests and will be published as a joint-journal paper post challenge at IEEE TMI or medical image analysis.

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.

Projectors, microphones

TASK: Segmentation of anatomical curves in liver

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.

The task of automatic anatomical curve segmentation in both 2D liver laparoscopy data and 3D CT/MRI data is a non-trivial and complex task. Thus, task 1 is aimed at segmentation of anatomical curves (landmarks) from intra-operative laparoscopy video images and preoperative 3D model.

Keywords

List the primary keywords that characterize the task.

Segmentation

ORGANIZATION

Organizers

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

Sharib Ali, Department of Engineering Science, University of Oxford, UK
Yueming Jin, University College London, UK
Yamid Espinel López, Université Clermont Auvergne, France
Emmanuel Buc, Clermont University Hospital, France
Bertrand Le Roy, Saint-Etienne University Hospital, France
Patrick Teoule, Universitätsmedizin Mannheim, Germany
Christoph Reissfelder, Universitätsmedizin Mannheim, Germany
Adam Bailey, Translational Gastroenterology Unit, Oxford University Hospitals NHS Trust, Oxford, UK
Zahir Soonawalla, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
Alex Gordon-Weeks, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
Michael Silva, Nuffield Department of Surgical Sciences,University of Oxford, Oxford, UK
Lena Maier-Hein, DKFZ, Heidelberg, Germany
Adrien Bartoli, Department of Clinical Research and Innovation, Clermont University Hospital, France

b) Provide information on the primary contact person.

Sharib Ali, Department of Engineering Science, University of Oxford, UK

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.

https://grand-challenge.org/challenges/

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

Part of EndoVis challenge (https://endovis.grand-challenge.org/). Sub-challenge site TBD

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.

Pretrained models will be allowed provided used data sources are publicly available and accessible to other members of the challenge

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 participate but not eligible for awards and not listed in leaderboard.

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

There will be cash awards and certificates for the winner and first runner-up in each sub-task. Cash awards will be according to the availability of funds from the sponsors. Contacts will be made for taking sponsors onboard.

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.

The results will be announced during the EndoVis Challenge at MICCAI'22. The results and rankings will be made publicly available on the challenge website. The submitted results will be further dissected through more rigorous statistical tests and will be published as a joint-journal paper post challenge.

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).

All participating teams will be required to submit a report describing their approach. Top 5 teams (based on algorithmic novelty and leaderboard ranking) from each sub-task will be invited for the joint journal writing. Each invited team can nominate 2 authors for the joint paper. The joint journal will be compiled and made available online within 8 months after the challenge (embargo period). The participating teams can publish their methods separately but only after posted arXiv preprint.

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.

Docker container on grand-challenge website will be used. Please refer to the link: https://comic.github.io/grand-challenge.org/evaluation.html

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.

We will provide all metric codes before the challenge that will allow teams to evaluate their methods on validation set of the training dataset. Codes will be provided at https://github.com/sharibox/P2ILF.

Test dataset will be released only two weeks prior to the workshop at MICCAI 2022. During this period, each team will be allowed only two submissions maximum. The most accurate submission will be taken into consideration at the time of closing.

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

Registration and release date of training dataset (subset-I): March 15, 2022
Release of training dataset (subset-II): May 15, 2022
Launch of forum, evaluation github codes, leaderboard setup: June 15, 2022
Registration ends and challenge on leaderboard begins: July 15, 2022
Test data release: August 15, 2022
Participants send results: August 17, 2022
Leaderboard closes: August 17, 2022
Report and presentation slides/videos: September 5th, 2022
Challenge Day: September 18th, 2022 (all test results presented)

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 released data will have patient consented data and ethically approved where applicable at the local centers. All ethical approval via institutional review board will be clearly mentioned on the readme file of the dataset and in all publications (once available if missing). All collected data are currently under regular patient consenting protocol.
Mannheim: DRKS00021748; date: June 17th, 2020; ethics committee template no.: 2020-575N, Medical Ethics Committee II, Medical Faculty Mannheim of the University of Heidelberg
Clermont-Ferrand: IRB00008526-2019-CE58
Oxford: We are working on ethics committee clearance for this center.

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 SA.

Additional comments: The data will be multi-center data released with consent from collaborators for free research and education purposes with the condition of citing the relevant publications linked to the datasets. Currently agreed license type is CC-BY-NC-SA which will be during the challenge. Post challenge we will revisit licensing types together with organisers.

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.

The Github repository will be updated with the evaluation codes. For details on evaluation metric that will be used in this task, we refer to reference [1].

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

Not a requirement but we encourage participants to make the code publicly available.

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.

Conflict of interest: Organisers do not have conflicts of interest.

Sponsors: Currently, we are looking for sponsors. Once we have this in place we will update the MICCAI 2022 challenge team.
Access to the test case labels: Only the organising team members (namely Sharib Ali, Yuemin Jin and Yamid Espinel López) of this challenge will have access to the splitted test case labels.

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

Research, Treatment planning, Assistance, Surgery, Intervention planning, Training.

Task category(ies)

State the task category(ies).

Examples:

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

Anatomical curve segmentation in liver

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.

Target cohort is the same for the challenge cohort. It is intended that developed algorithms can be applied to real clinical scenarios. For this we will include diverse case samples e.g. laparoscopic videos with and without surgical procedures. Additionally, we will also provide phantom cases to increase data variability. Participants will be evaluated on both real patient data and phantom data.

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

Pair of same patient laparoscopic video images and 3D model (generated from preoperative 3D CT or MRI scans from the same patient). Additionally, we will also provide phantom cases acquired by similar procedure to increase data variability. Participants are allowed to train their deep learning model either separately for patient and phantom data or as combined.

Imaging modality(ies)

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

Laparoscopy images; 3D mesh
Modalities: White light endoscopy images; Mesh

Context information

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

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

15 CT volumes and 15 corresponding patient laparoscopic videos will be used for this challenge.

Additionally, 10 mesh and 10 corresponding laparoscopic videos of phantom data will be also be added.

Pixel-level segmentation masks with annotated anatomical curves will be provided in png format.

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

All data will be fully anonymised.

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.

Liver acquired from Karl Stolz laparoscope for video image data. CT was acquired from Siemen's device of the same organ.

The preoperative data has been acquired after providing a contrast agent to the patient to improve the tissue segmentation in the MRI data. The intraoperative videos have been filmed using the laparoscope at minimal zoom and with medium-intensity lightning. Phantom deformations have been generated using the Abaqus software and then 3D printed using polylactic acid (PLA).

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.

Liver anatomical curve segmentation

Both target and challenge cohort will include real patient data and phantom data. For challenge cohort participants will be allowed to use the data as they intend to for algorithmic development, however, for test evaluation we will quantify methods on both real and phantom data.

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)).

Accuracy, Consistency, Precision, Feasibility, Specificity.

Additional points: For anatomical curve segmentation (assessed F1-score that provides a trade-off between precision and recall). Additionally, we will recalculate the TP, FP and FN with 2 % tolerance as detailed in [1]. F1-score and Hausdorff distance will be used. A separate ranking system will be used to avoid any confusion. The final ranking will be the averaged rank of the two scores. The reason behind using two scores is to tackle bias in evaluation for segmentation as studied previously [4].

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).

Karl Storz; Siemens

The videos have been recorded using a PC with a capture card, or directly saved into a USB storage device from the laparoscopic system.

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

For every patient, the 3D CT/MRI images were acquired several days before the surgery (preoperative 3D volumes of liver). Then, the relevant tissues were segmented in these volumes, and the preoperative 3D models were generated by interpolating these segmentations using MITK software. During surgery, an exploration of the intra-abdominal scene is done and the video stream is captured using endoscopy (laparoscopy).

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.

The centers involved for data acquisition are:
[1] Centre Hospitalier Universitaire de Clermont-Ferrand, France
[2] Oxford Radcliffe Hospitals, Oxford, UK
[3] Universitätsmedizin Mannheim, Germany

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).

The data was acquired by experienced laparoscopic surgeons using state-of-the-art endoscopic systems, as part of their routine clinical interventions.

The phantom data has been generated from a real patient liver. This liver has been filled with artificial markers and inner control points, and then deformed using Abacus. The deformations simulate gas pressure and instrument manipulations. 10 different deformations have been generated, from which 10 phantom models were 3D printed using Polylactic Acid (PLA). For each of the phantoms, several images were taken using a medical laparoscope from different viewpoints as done in the surgery room. Ground truth data was generated for 10 images per phantom by aligning the deformed 3D models to the images using the Perspective-n-Point algorithm.

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.

A case refers to both the preoperative 3D model and laparoscopic images from a particular patient or phantom data. The preoperative 3D model will be annotated with the 3D anatomical curves corresponding to their 2D counterparts in the laparoscopic images. In addition to the anatomical curves, the laparoscopic images will be provided with the visible key anatomical curves that correspond to the liver ridge, falciform ligament and silhouette contour.

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

Patient data
Train: 250 frames (total 10 patients)
Val: 50 frames (total 10 patients)
Test: 100 (additional unseen 5 patients)
Each will be linked with points associated to anatomical curves in 3D model

Phantom data
Training: 70 frames (total seven phantom videos and 3D models)
Test: 30 frames (total three phantom videos and three 3D models)

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

It is chosen to balance a good tradeoff in annotation effort and to maintain sufficient visual diversity, e.g., frames from some centres are not diverse so including more might lead to over fitting.

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.

For real data train and test set will have mixed centers (i.e. an equal distribution from all three centers). Participants will be allowed to choose their validation set as per their choice and the details on data will be provided to them.

All these centers have the same hardware for acquiring laparoscopic images (Karl Storz) while 3D models will not make much difference on the acquisition. However, each surgical procedure and patient is unique so providing 10 unique patient data for training and testing on unseen 5 unique patients will be the challenge in terms of how well teams can leverage their method to real-world settings. In addition we will also assess participants on 3 phantom sequence and volumes for which 7 volumes will be provided for training and validation.

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.

Reference annotations are based on visual cues that are available in liver laparoscopy procedure (e.g., ligaments, liver surface, ridges, silhouettes etc.). For 3D models, the falciform ligament can be characterised by expert surgeons, while ridges and liver surface again serve as visual cues.
All generated annotations will be manual that will involve five annotators and at least three consultant surgeon reviews. Also refer to 23 c for more details.

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.

A defined protocol has been set to minimise annotator variability. Additionally, samples with annotations from expert surgeons will also be provided for reference. We define each anatomical landmark as bellow:
Ridges: Lower curvy liver boundaries (curve)
Silhouettes: Upper curvy boundaries of liver (curve)
Falciform ligament: thin, sickle-shaped, fibrous structure that connects the anterior part of the liver to the ventral wall of the abdomen (curve)
Liver surface: entire area of the liver (will be pixel-level segmentation)
For full description on protocol please see Section 3 in the link below:
https://docs.google.com/document/d/10ZUhFPu6McHXPDDNG8tdKEiB0Foou8SRnoWksIcAZkI/edit?usp=sharing)

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.

Below information are provided and coincide with all cases:
Software: Labelbox, https://labelbox.com

Software for 3D CT/MRI annotation: MITK
Type: Manual annotation
Expertise of annotators: One PhD student (4 years working on same field), two post PhD researchers (4-8 years), two surgeon (over 3 years)
Expertise of reviewers: Consultant surgeon (over 10 years)

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.

Each annotation will be performed by an individual annotator, however, reviewed by multiple experts (consultants) in surgery. Annotations will be reviewed by expert surgeons who will determine if the annotation suffices the quality. In case of poor annotation quality, the rejected frame will be reannotated with supervision of two expert surgeons. This protocol follows for all data and has been detailed in the protocol document provided in 23 b.

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.

The following preprocessing steps (where applicable) will be used for all cases in the provided dataset:
a) Distortion correction of images using intrinsic camera parameters (where applicable)
b) Simplification of the 3D models using Meshlab.

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.

Possible errors from annotation:
Poor lighting of the intraoperative scene.
Blurriness in the images.
Presence of smoke/blood/water.
In CT/MRI, tissues with low contrast or blurry images.
The tolerance for error in 2D images will be less than 2% as in ref [1].

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

Poor judgement of boundaries
Poor understanding of the liver anatomy
Annotation tool
Mitigations:
To mitigate above relevant sources of error, we have put in place expert review of annotations. Additionally, we have requested the annotators to use interactive tablets specialised for annotation, e.g., Wacom devices.

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)

F1-score (balance between precision and recall) will only be used to score the teams. Here, TP (true positive), FP (false positive) and FN (false negative) will be calculated as 2% tolerance for the predicted curves (see reference [1]). Additionally to tackle bias in evaluation for segmentation as studied previously [4], we will also use Hausdorff distance as a distance metric between the curve points.

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

F1-score provides a better tradeoff between precision and recall. We want to measure the fraction of relevant instances among the predicted pixel instances taking into account both the false positive predictions and false negative predictions. The motivation behind choosing 2% tolerance of the anatomical curves in segmentation is detailed in [1].

Additionally to tackle bias in evaluation for segmentation as studied previously [4], we will also use Hausdorff distance as a distance metric between the curve points.

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.

A separate ranking will be used each for F1-score (increasing reflect best) and Hausdorff distance (decreasing reflect best) to avoid any confusion. The final ranking will be the averaged rank of the two scores.

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

Not allowed. Such submissions will be considered invalid.

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

While increasing F1-score reflect superior ranking, Hausdorff distance with decreasing values indicate superior ranking. Thus, we have set separate ranking for each first which will be then averaged to get the final ranking.

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.

Each anatomical curve lines will be assessed separately and averaged. Missing landmark curve points will directly affect the average scoring. However, we believe that 2% tolerance for estimating TP, FP and FN [1] will help us

mitigate this issue.
Variability in ranking: Intended post challenge
Statistical approach: Intended post challenge

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

Intended post challenge

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.

Intended post challenge

TASK: 2D-3D registration

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.

Augmented reality (AR) techniques are vital to surgeons as they enable precise tumour localisation for surgical removal, however, there require better alignment of preoperative 3D model with intra-operative 2D surgical laparoscopic liver surgery data. Thus, this task is aimed at registration of extracted 2D curves to the corresponding anatomical curves in 3D liver model created from preoperative volumetric data. Participants are required to provide transformation matrixes for this task for which re-projection error and target registration error will be computed.

Keywords

List the primary keywords that characterize the task.

Intraoperative and preoperative registration

ORGANIZATION

Organizers

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

Yamid Espinel López, Université Clermont Auvergne, France
Sharib Ali, Department of Engineering Science, University of Oxford, UK
Yueming Jin, University College London, UK
Emmanuel Buc, Clermont University Hospital, France
Bertrand Le Roy, Saint-Etienne University Hospital, France
Patrick Teoule, Universitätsmedizin Mannheim, Germany
Christoph Reissfelder, Universitätsmedizin Mannheim, Germany
Adam Bailey, Translational Gastroenterology Unit, Oxford University Hospitals NHS Trust, Oxford, UK
Zahir Soonawalla, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
Alex Gordon-Weeks, Nuffield Department of Surgical Sciences, University of Oxford, Oxford, UK
Michael Silva, Nuffield Department of Surgical Sciences,University of Oxford, Oxford, UK
Lena Maier-Hein, DKFZ, Heidelberg, Germany
Adrien Bartoli, Department of Clinical Research and Innovation, Clermont University Hospital, France

b) Provide information on the primary contact person.

Sharib Ali, Department of Engineering Science, University of Oxford, UK

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.

https://grand-challenge.org/challenges/

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

Part of EndoVis challenge (https://endovis.grand-challenge.org/). Sub-challenge site TBD

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.

Pretrained models will be allowed provided used data sources are publicly available and accessible to other members of the challenge

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 participate but not eligible for awards and not listed in leaderboard.

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

There will be cash awards and certificates for the winner and first runner-up in each sub-task. Cash awards will be according to the availability of funds from the sponsors. Contacts will be made for taking sponsors onboard.

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.

The results will be announced during the EndoVis Challenge at MICCAI'22. The results and rankings will be made

publicly available on the challenge website. The submitted results will be further dissected through more rigorous statistical tests and will be published as a joint-journal paper post challenge.

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).

All participating teams will be required to submit a report describing their approach. Top 5 teams (based on algorithmic novelty and leaderboard ranking) from each sub-task will be invited for the joint journal writing. Each invited team can nominate 2 authors for the joint paper. The joint journal will be compiled and made available online within 8 months after the challenge (embargo period). The participating teams can publish their methods separately but only after posted arXiv preprint.

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.

Docker container on grand-challenge website will be used. Please refer to the link: https://comic.github.io/grand-challenge.org/evaluation.html.

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.

We will provide all metric codes before the challenge that will allow teams to evaluate their methods on validation set of the training dataset. Codes will be provided at https://github.com/sharibox/P2ILF.

Test dataset will be released only two weeks prior to the workshop at MICCAI 2022. During this period, each team will be allowed only two submissions maximum. The most accurate submission will be taken into consideration at the time of closing.

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

Registration and release date of training dataset (subset-I): March 15, 2022
Release of training dataset (subset-II): May 15, 2022
Launch of forum, evaluation github codes, leaderboard setup: June 15, 2022
Registration ends and challenge on leaderboard begins: July 15, 2022
Test data release: August 15, 2022

Participants send results: August 17, 2022
Leaderboard closes: August 17, 2022
Report and presentation slides/videos: September 5th, 2022
Challenge Day: September 18th, 2022 (all test results presented)

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 released data will have patient consented data and ethically approved where applicable at the local centers. All ethical approval via institutional review board will be clearly mentioned on the readme file of the dataset and in all publications (once available if missing). All collected data are currently under regular patient consenting protocol.
Mannheim: DRKS00021748; date: June 17th, 2020; ethics committee template no.: 2020-575N, Medical Ethics Committee II, Medical Faculty Mannheim of the University of Heidelberg
Clermont-Ferrand: IRB00008526-2019-CE58
Oxford: We are working on ethics committee clearance for this center.

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 SA.

Additional comments: The data will be multi-center data released with consent from collaborators for free research and education purposes with the condition of citing the relevant publications linked to the datasets. Currently agreed license type is CC-BY-NC-SA which will be during the challenge. Post challenge we will revisit licensing types together with organisers.

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.

The Github repository will be updated with the evaluation codes. For details on evaluation metric that will be used in this task, we refer to reference [2] for target registration error and [3] for re-projection error computation.

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

Not a requirement but we encourage participants to make the code publicly available.

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.

Conflict of interest: Organisers do not have conflicts of interest.
Sponsors: Currently, we are looking for sponsors. Once we have this in place we will update the MICCAI 2022 challenge team.
Access to the test case labels: Only the organising team members (namely Sharib Ali, Yuemin Jin and Yamid Espinel López) of this challenge will have access to the splitted test case labels.

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, Treatment planning, Assistance, Surgery, Intervention planning, Training.

Task category(ies)

State the task category(ies).

Examples:

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

2D-3D registration

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.

Target cohort is the same for the challenge cohort. It is intended that developed algorithms can be applied to real clinical scenarios. For this we will include diverse case samples e.g. laparoscopic videos with and without surgical procedures. Additionally, we will also provide phantom cases to increase data variability. Participants will be evaluated on both real patient data and phantom data.

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

Pair of same patient laparoscopic video images and 3D model (generated from preoperative 3D CT or MRI scans from the same patient). Additionally, we will also provide phantom cases acquired by similar procedure to increase data variability. Participants are allowed to train their deep learning model either separately for patient and phantom data or as combined.

Imaging modality(ies)

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

Laparoscopy images; 3D mesh
Modalities: White light endoscopy images; Mesh

Context information

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

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

15 CT volumes and 15 corresponding patient laparoscopic videos will be used for this challenge.

Additionally, 10 mesh and 10 corresponding laparoscopic videos of phantom data will be also be added.

3D models of segmented liver volume (dense meshes in .ply format); locations of ground-truth 3D landmark curve locations as .csv files

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

All data will be fully anonymised.

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.

Liver acquired from Karl Stolz laparoscope for video image data. CT was acquired from Siemen's device of the same organ.

Patient data:
The preoperative data has been acquired after providing a contrast agent to the patient to improve the tissue segmentation in the MRI data. The intraoperative videos have been filmed using the laparoscope at minimal zoom and with medium-intensity lightning. Camera intrinsic parameters k of the laparoscope is determined before acquiring the images. In order to measure the reprojection error in the patient data, the registered model is projected into a control view. Then, the 2D distances in pixels between the boundaries of the liver in the image and the projected 3D model are measured [3].

Phantom data:
Phantom deformations have been generated using the Abaqus software and then 3D printed using polylactic acid (PLA). Ground truth data has been generated by tracking the deformed artificial markers and control points, and then finding their positions in every image using PnP [2].

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.

2D-3D registration of preoperative 3D model with intraoperative laparoscopic data

Both target and challenge cohort will include real patient data and phantom data. For challenge cohort participants will be allowed to use the data as they intend to for algorithmic development, however, for test evaluation we will quantify methods on both real and phantom data.

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)).

Reliability, Accuracy, Consistency, Feasibility, Robustness.

Additional points: For 2D-3D registration we will use: 1) Target registration error (TRE) for phantom data, and 2) reprojection error (root mean square error, RMSE) for patient data. For ranking we will use two separate ranks one for phantom data and another for patient data.

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).

Karl Storz; Siemens

The videos have been recorded using a PC with a capture card, or directly saved into a USB storage device from the laparoscopic system. 3D models were generated by interpolating segmentation masks of the liver using MITK software.

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

For every patient, the 3D CT/MRI images were acquired several days before the surgery (preoperative 3D volumes of liver). Then, the relevant tissues were segmented in these volumes, and the preoperative 3D models were generated by interpolating these segmentations using MITK software. During surgery, an exploration of the intra-abdominal scene is done and the video stream is captured using endoscopy (laparoscopy).

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.

The centers involved for data acquisition are:
[1] Centre Hospitalier Universitaire de Clermont-Ferrand, France
[2] Oxford Radcliffe Hospitals, Oxford, UK
[3] Universitätsmedizin Mannheim, Germany

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).

The data was acquired by experienced laparoscopic surgeons using state-of-the-art endoscopic systems, as part of their routine clinical interventions.

The phantom data has been generated from a real patient liver. This liver has been filled with artificial markers and inner control points, and then deformed using Abacus. The deformations simulate gas pressure and instrument manipulations. 10 different deformations have been generated, from which 10 phantom models were 3D printed using Polylactic Acid (PLA). For each of the phantoms, several images were taken using a medical laparoscope from different viewpoints as done in the surgery room. Ground truth data was generated for 10 images per phantom by aligning the deformed 3D models to the images using the Perspective-n-Point algorithm.

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.

A case refers to both the preoperative 3D model and laparoscopic images from a particular patient or phantom data. The preoperative 3D model will be annotated with the 3D anatomical curves corresponding to their 2D counterparts in the laparoscopic images. In addition to the anatomical curves, the laparoscopic images will be provided with the visible key anatomical curves that correspond to the liver ridge, falciform ligament and silhouette contour.

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

Training/validation and test meshes will be the same set as described in the task 1.

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

It is chosen to balance a good tradeoff in annotation effort and to maintain sufficient visual diversity, e.g., frames from some centres are not diverse so including more might lead to over fitting.

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.

All these centers have the same hardware for acquiring laparoscopic images (Karl Storz) while 3D models will not make much difference on the acquisition. However, each surgical procedure and patient is unique so providing 10 unique patient data for training and testing on unseen 5 unique patients will be the challenge in terms of how well teams can leverage their method to real-world settings. In addition we will also assess participants on 3 phantom sequence and volumes for which 7 volumes will be provided for training and validation.

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.

Reference annotations are based on visual cues that are available in liver laparoscopy procedure (e.g., ligaments, liver surface, ridges, silhouettes etc.). For 3D models, the falciform ligament can be characterised by expert surgeons, while ridges and liver surface again serve as visual cues.

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.

For full description on protocol please see Section 3 in the link below:
https://docs.google.com/document/d/10ZUhFPu6McHXPDDNG8tdKEiB0Foou8SRnoWksIcAZkI/edit?usp=sharing

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.

3D contour landmarks will be obtained using in-house software called Hepataug that allows to interactively visualise and mark ridges (curvature at the bottom side of the liver, ligament (division between the right and left lobes) and silhouette.

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.

Each annotation will be performed by an individual annotator, however, reviewed by multiple experts (consultants) in surgery. Annotations will be reviewed by expert surgeons who will determine if the annotation suffices the quality.

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.

The following preprocessing steps (where applicable) will be used for all cases in the provided dataset:
a) Distortion correction of images using intrinsic camera parameters (where applicable)
b) Simplification of the 3D models using Meshlab.
c) 3D contour landmarks will be obtained using in-house software called Hepataug

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.

Possible errors from annotation:
Poor lighting of the intraoperative scene.
Blurriness in the images.
Presence of smoke/blood/water.
In CT/MRI, tissues with low contrast or blurry images.
The tolerance for error in 2D images will be less than 2% as in ref [1].

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

Poor judgement of boundaries
Poor understanding of the liver anatomy
Annotation tool
Mitigations: To mitigate above relevant sources of error, we have put in place expert review of annotations. Additionally, we have requested the annotators to use interactive tablets specialised for annotation, e.g., Wacom devices.

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)

Target Registration Error (TRE) will be computed on artificial markers and control points added to the phantom datasets. TRE is measured as the average 3D distance between the registered and the groundtruth markers and control points. TRE also includes the standard deviation of the distances across all the images [2]. However, in absence of such markers in real patient data we will compute reprojection error. For this, camera intrinsic parameters k of the laparoscope is determined before acquiring the images. In order to measure the reprojection error, the registered model is projected into a control view. Then, the 2D distances in pixels between the boundaries of the liver in the image and the projected 3D model are measured [3].

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

Target Registration Error (TRE) can be established in the phantom data and will provide a more accurate quantitative evaluation than reprojection error. However, in terms of absence of landmarks in the real data, we will have to rely on classically used reprojection error [3].

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.

Two separate ranking will be conducted - 1) for phantom, a lower mean TRE will be expected while 2) for patient data, a lower mean reprojection error will be desired. In terms of ties, participants with least deviation will be announced winner.

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

Not allowed. Such submissions will be considered invalid.

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

TRE is one of the accurate way of assessing registration algorithms. This is possible in our phantom data. However, reprojection error will be used in patient data due to absence of these landmarks. The ranking cannot be combined as these are two different approaches. Thus, we will provide a separate ranking for this scheme. Two winners will be declared (one for phantom and other for real patient test data) for this task.

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.

Variability in ranking: Intended post challenge
Statistical approach: Intended post challenge

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

Intended post challenge

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.

Intended post challenge

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.

[1] François T, Calvet L, Madad Zadeh S, Saboul D, Gasparini S, Samarakoon P, Bourdel N, Bartoli A. Detecting the occluding contours of the uterus to automatise augmented laparoscopy: score, loss, dataset, evaluation and user study. Int J Comput Assist Radiol Surg. 2020 Jul;15(7):1177-1186. DOI: https://doi.org/10.1007/s11548-020-02151-w
[2] Espinel, Y., Calvet, L., Botros, K., Buc, E., Tilmant, C., & Bartoli, A. (2021). Using Multiple Images and Contours for Deformable 3D-2D Registration of a Preoperative CT in Laparoscopic Liver Surgery. MICCAI ; Plantefève R, Peterlik I, Haouchine N, Cotin S. Patient-specific Biomechanical Modeling for Guidance during Minimally-invasive Hepatic Surgery. Annals of Biomedical Engineering, Springer Verlag, 2015. DOI: https://doi.org/10.1007/978-3-030-87202-1_63
[3] Espinel Y, Özgür E, Calvet L, Le Roy B, Buc E, Bartoli A. Combining Visual Cues with Interactions for 3D-2D Registration in Liver Laparoscopy. Ann Biomed Eng. 2020;48(6):1712-1727. DOI:10.1007/s10439-020-02479-z.
[4] Reinke et al. (2018). How to exploit weaknesses in biomedical challenge design and organization. Medical Image Computing and Computer Assisted Intervention - MICCAI 2018 (pp. 388-395). Springer International Publishing.