1st PSCC-IPPMeD Challenge
on "AI for Lung Cancer Imaging"

The Institut Polytechnique de Paris is organizing the first PSCC-IPParis Medical Data (**IPPMeD**) Challenge on "AI for Lung Cancer Imaging", under the Patronage of PSCC Connect, and Telecom Paris Alumni. It will take place remotely from Nov. 21, 2023 to Jan 31, 2024.

Registration - under conditions detailed below - is open on
https://forms.gle/MeCZZUDajvdUr3wdA from November 14th 2023 until November 20th 2023 Noon. Registration criteria will be checked by the organisers before individual official invitations to participate are sent.

Introduction
The PSCC is a Biocluster funded by France2030 to foster innovation in Oncology. IPParis is a founding member of the PSCC, along with Sanofi, Gustave Roussy, Université Paris Saclay and INSERM.

The theme of the **IPPMeD** challenge this year is "AI for Lung Cancer Imaging" with a specific task on **lung cancer diagnosis on CT scan images**.
Lung cancer is the leading cause of cancer-related deaths worldwide, accounting for the highest mortality rates among both men and women, with an estimated 1.8 million deaths (18%) in 2020. The challenge focuses on automated lung tumor contouring, also known as segmentation, from Computer Tomography (CT) thoracic scans.

Lung tumor segmentation is a key element for diagnosing and then guiding therapy, either with resection (surgery) or radiotherapy. In particular, it is necessary to plan the radiation dose and know which area to target in radiotherapy. Hence many approaches to automate lung tumor segmentation from CT scans have been proposed in the past years, and they constitute a clinically meaningful application of AI.

Ranking of submitted AI solutions will involve segmentation quality criteria as well as frugality criteria for minimal computational footprint (see below for details).

Goal of the challenge

The goal of this challenge is to automatically segment one primary lung tumor in thoracic CT scans. Given a 3D CT scan as input (a volume of grayscale slice images), the algorithm should return a binary mask with the same dimensions, with voxel value = 0 or 1 depending on the structure it belongs to:

1 = the pixel is part of the primary lung tumor

0 = the pixel does not belong to the primary lung tumor

The Challenge will only accept **end-to-end** machine learning (e.g. radiomics) or deep learning (e.g CNN) approaches that are able to learn - from the provided training data only - to predict correctly the tumor contours in new scans.

Rules to participate

To take part in the challenge, participants need to form a **team** of 2 to 3 people. All participants must be registered at IPParis or one of its school as either a 3rd/4th year engineering student or a Master 2 student.

Each team should select a "**team name**" (any single word/acronym without blank space), and each team **member** should **register** indicating their team name and their IPParis email address via this link https://forms.gle/MeCZZUDajvdUr3wdA . Individual registrations are mandatory from all team members.

As we strive to design our future event in a way that fosters gender equality and inclusivity within the context of deep learning, the data provided voluntarily during the registration will be used to publish statistics at the end of the challenge.

Prizes

(Pending approval for implementation)

The top-3 teams will be awarded a cash prize at the official Challenge Conclusion Event on Jan 31st 2024. The prizes are sponsored this year by **PSCC Connect** and **Telecom Paris Alumni**.

Prizes per team:

- 1st place: 1.5 K€
- 2nd place: 1 K€
- 3rd place: 500€

-     2 other finalists: 250€ each

Prizes will be equally distributed among members of the winning teams.

Data

At training time:

You will be provided - at the start of the Challenge - with a cohort of N=291 **training CT scans,** along with their corresponding ground-truth (gt) primary tumor segmentations.

Each CT scan is a 3D volume showing the lungs and adjacent thoracic structures.

Each gt mask provides the segmentation labels as follows:

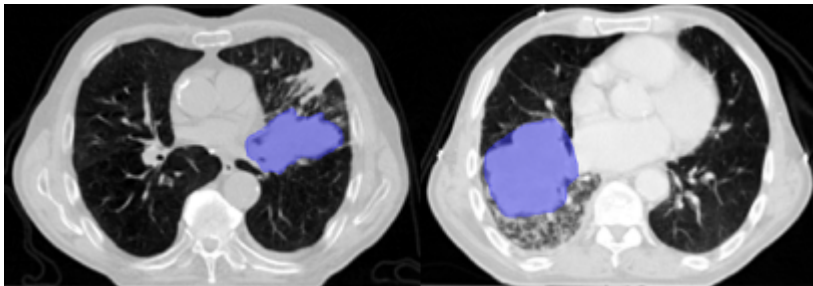0 = Normal tissue or non-primary tumor

1 = Primary lung tumor

Illustration of GT annotation from the LOTUS paper

We will also provide anonymized **metadata** in csv files: train_metadata.csv = [sex, age, cancer stage, histology], train_dicom_metadata.csv = image information such as: [scanner manufacturer, kVp, mA, ..]. Meta data is not complete, and has missing values, which is typical of real clinical data registries.

To help you focus on some structures of interest in the provided images (here the lungs), we provide for the training set the binary segmentation **masks of the lungs** (0 = outside the lungs; 1 = lungs). Note that the lung masks are not perfect, as some have connections between the left and right lungs. This type of issue is typical of real outcomes from automated segmentation tools. You can use these additional segmentations to guide the training of your tumor segmentation model.

You can use metadata and lung masks for the training but not for the inference. Such additional training data can help you for example to stratify your train/validation sets, to learn some attention mechanism, or to infer accessory tasks in addition to the segmentation.

At test time:

Your solution will be evaluated on N=100 test CT scans against their ground-truth primary tumor segmentations. Neither metadata nor lung segmentation masks will be available for the test scans.

Training and submission rules

You can only use the provided images and tumor masks as input training data for this challenge. The evaluation, submission and rankings of participants to reach the finalist stage will be done via the Kaggle competition platform. You will be invited to the competition shortly after the kick-off meeting and will then be able to submit your predictions to Kaggle as a .csv file. Note that a package handling the conversion from the NIfTI files with the predicted tumor masks to the submission file is provided (link in the Kaggle overview).

Machine Learning models accepted

You can use any publicly shared baseline model. If using a deep-learning architecture, it is mandatory that all model weights are initialized and trained from scratch by the participants. No transfer of pre-trained weights will be tolerated.

When writing your code, pay attention to control essential parameters for "accurate" replication of your training set up (required in the finalist stage - with replication level up to some reasonable margin), such as: random seeds and order of CT scans used for training. A list of references describing well-known models will be found at the end of this page but you can find many more articles in the literature. We recommend Pubmed and Google Scholar as search tools. By using any IP-PAris VPN, you should have access to most of the relevant scientific journals.

Evaluation Criteria

Evaluation of the challenge will be based on the ranking in the Kaggle leaderboard, which evaluates the **segmentation** performance of your model on the held-out test dataset, based on the three metrics presented below.

The 5 teams with the best rankings on the leaderboard on the FREEZE DAY (see below) of the challenge will proceed to a final stage to rank the final 3 winning teams.

The final ranking of the 3 winning teams will involve running the submitted codes (reproducibility) and include metrics related to network size and inference time (**frugality**)

Timeline

- **FREEZE DAY**: all teams have submitted their final test results on Kaggle. Leaderboard is now frozen on Kaggle and made public. Top 5 performing teams are selected for the final round. The Top-5 teams will be the teams with the best segmentation performance metrics (see below) on the test set.
- **CODE DAY**: the 5 finalist teams submit a working implementation of their model (both training and inference code).
  This should include a requirements.txt file to replicate your working python environment (necessary libraries) and all self-contained code, including a main_train.py script to execute to re-train your model and a main_test.py to make predictions on the train and test data. Please also provide a brief summary of your approach (as a markdown or text format of your choice). This step will exclude teams found ineligible - for not respecting the rules of this challenge- or showing too much discrepency between inference quality on test results with the open test cohort, between the Leadearbord and our re-trained performance on the same test set. In the event of one or more top-5 teams being excluded, the top 6+ team(s) will be included to define the final roster of 5 finalist teams.
- **AWARD DAY**: for the 5 finalist teams, the organizing team will have fully retrained the submitted models and re-perform train and test-time inferences by executing the submitted scripts. No further instructions from the participating teams will be admissible. At this stage, the winner will be based on segmentation performance metrics (see below) as well as computational complexity of the model: network size and inference time per scan.

Evaluation Metrics

The following metrics will be used to evaluate the quality of the segmentation masks:

[1] the **mean 3D DICE** score across the test subjects:

[2] The mean error on **RECIST** maximal diameter [code provided]

[3] The mean error on **tumoral volume** [code provided]

Leaderboard

The test data is split between a public leaderboard (which can be probed during the competition) and a private leaderboard, which is only revealed at the end of the competition. For this challenge, 20% of the test data is shared and used in the public leaderboard and 80% in the held-out leaderboard.

Timeline

- **Nov 14**, 2023: Registration opens
- **Nov 20** 2023: Registration closes
- **Nov 21** 2023: Introductory meeting for registered teams. At Telecom Paris.
- **Dec 1st**, 2023: Challenge starts (with guest accounts created by Telecom Paris)
- —------------------
- **January 29th**, 2023: **FREEZE DAY** + the 5 finalists are announced
- **February 1st**, 2023: **CODE DAY** - finalists submit their code
- **February 29th**, 2023: Final ranking and prize award ceremony at Telecom Paris.

Computing Resources

Some access will be provided to the GPU cluster of Télécom Paris.
 Shortly after joining the challenge, a session will be organized to guide the teams with access through using the resources of this cluster.
You will be encouraged to use the resources of the cluster sparsely.

Methods and software tools

OS for submitted code: Linux
Code: Python >= 3.8
DL libraries: Pytorch or Tensor flow

Suggested packages: numpy, scikit-learn (machine learning), pyradiomics, nibabel, MONAI, SimpleITK.

You can use any code - coded by you or not - as permitted by their license. Keep in mind that for your team to be eligible to win, the organizing team should be able to independently re-run both the training and inference stages of your model given a file of library requirements, and code including main_train.py and main_test.py scripts to execute.

All CT scans and segmentation masks in the dataset are saved as NIfTI files (.nii or .nii.gz extensions). You can read, write and analyze such images using one of several python libraries, such as nibabel, TorchIO, MONAI, or Pigeon.

To visualize CT scans and segmentations, you may directly display 2D slices in a jupyter notebook using matplotlib after loading the image, or you may use dedicated free visualization software tools such as ParaView (after loading the AnalyzeNIfTIReaderWriter plugin), 3D Slicer, ITK Snap and others.

References

[1] https://www.who.int/news-room/fact-sheets/detail/lung-cancer
[2] https://www.nature.com/articles/s41467-022-30841-3
[3] https://arxiv.org/pdf/2201.00458.pdf [LOTUS]