CoSolve Sprint Challenge

Challenge Title: **Advanced signal detection in MRI scans to enable detection and quantification of cancers**

Abstract:

AstraZeneca is seeking an automated tool for the detection of tumor-relevant regions from images generated via a proprietary imaging approach. The participant will be provided with image scans and boundaries of tumorigenic *areas within the images. AstraZeneca seeks data analytics solutions that use innovative approaches to help detect and quantify the signal changes in these tumor regions over time.*

This challenge requires written documentation and submission of the source code and/or an executable for testing the algorithm.

Structure/Image/Logo:

Summary:

ABOUT THE SEEKER

AstraZeneca is a global, research-based, biopharmaceutical company with a focus on five key therapeutic areas: 1) cardiovascular, renal & metabolism 2) oncology, 3) respiratory and immunology, 4) vaccines and immune therapies, 5) rare diseases. As an innovation-driven research organization, AstraZeneca recognizes that great ideas come from many sources. Open innovation is an avenue by which ideas can be shared and AstraZeneca recently launched CoSolve to facilitate the advancement of pharmaceutical research.

Detailed Description and Solution Requirements:

BACKGROUND

AstraZeneca is exploring the use of advanced AI approaches for the accurate detection of tumorigenic areas using images constructed using proprietary technologies. These images are supplied with this challenge and correspond to 3 dimensional MRI images. The images are provided in the Neuroimaging Informatics Technology Initiative (NIfTI) open file format. The data includes MRI scans from 20 mice for 2 or 3 different time points: the baseline scan (time point 1), 15 minutes after the baseline scan (time point 2), and a day after the baseline scan (time point 3). In total, there are 50 MRI scans. Each MRI scan also has a corresponding annotation mask in a separate NIfTI file.

Posting Period: 90 days. Evaluation Period: 60 days

THE CHALLENGE

The images and a readme file are available to the registered participant for download (see data location link, below). The task is to devise an algorithm that is able to maximally separate tumorigenic areas from the rest of the areas in the images for those images in timepoints 2 and 3. Specifically, the algorithm should be able to calculate relevant features in the tumor, such that the features can be used to demonstrate that areas within tumor regions can be used to differentiate tumor from other regions in those images. It should be noted that timepoint 1 corresponds to baseline imaging without any MRI imaging agent and can be thus used to enhance the signals needed to differentiate tumor regions seen at timepoints 2 and 3. If the submitter needs to, multiple images from a given timepoint (for timepoints 2-3) can be used to enhance relevant signals, along with any sets of images from timepoint 1. The submitter must demonstrate that the aforementioned features are able to statistically separate regions within a tumorigenic boundary and non-tumorigenic regions using at least one timepoint from timepoints 2 and 3. This demonstration should minimally consist of showing the calculated feature scores within the tumorigenic regions are statistically different from similarly bootstrapped samples from nontumorigenic areas of the test images in the image. Additional metrics and analysis are at the discretion of the participant and should focus on convincingly demonstrating that the algorithm has strong discriminatory power, especially focusing on potential interpretability and comparison across the different time points. For instance, one can imagine that an optimally performing algorithm might reveal that signal intensity is dependent on the timepoint. To be clear, there is no guarantee that such dependencies exist in the current images.

1. A detailed description of the proposed Solution and how it addresses each **Technical Requirement** presented above. This should also include a **thorough description of the algorithm used and the code shared** in the Solution accompanied by a well-articulated rationale for the method employed. A description of the method in a text form of up to 3 pages with up to 10 appendix pages (the evaluation will be done only based on the main pages, but the Reviewers will use the appendix pages in case there is a need for clarification) must be provided to help the Reviewers to assess the approach.

2. **Output** from the proposed algorithm will need to be applied to the images in the test folder (10 images from 4 mice) that will not be used for training the model.

3. **The participants must submit** the source code and/or an executable with sufficient documentation to enable the Seeker to compile, execute the algorithm, and validate the method using additional test data sets. If there are challenges in submitting the code with the submission,repositories such as github can be utilized and relevant links must be provided in the submission report.
Furthermore, the Solution must rely on open source or similar licensing models that would allow the Seeker to use the tool without incurring additional licensing costs. The Challenge award is contingent upon evaluation and validation of the submitted Solutions by the Seeker. During the evaluation period, the Seeker will validate submissions using additional images similar to the example images provided in the challenge.

To receive an award, **the Solvers will not have to transfer their exclusive IP rights to the Seeker**. Instead, Solvers will grant to the Seeker a *non-exclusive license* to practice their solutions.

Submissions to this Challenge must be received **by 11:59 PM (US Eastern Time) on Nov 30th, 2024**.
Late submissions will not be considered. The proposal should not include any personal identifying information (name, username, company, address, phone, email, personal website, resume, etc.) or any information the Solvers may consider as their Intellectual Property they do not want to share.

Appendix

Use of public data: allowed

Help from medical imaging experts or others: Allowed. It is the responsibility of the submitters to properly acknowledge such contributions.

Evaluation criteria: The submission will be evaluated for both innovation and for significance of the proposed approach in identifying tumorigenic regions in images that are similar to those provided with this challenge. The seeker will give a maximum of two awards of up to USD $12,500 per award. While the specifics of the final award amount rests solely on the seeker, the seeker expects to give at least one award of USD $12,500, and upto a maximum of USD 25,000 for a single award. The Challenge award is contingent upon evaluation and validation of the submitted Solutions by the Seeker. During the evaluation period, the Seeker will validate submissions using additional images similar to the example images provided in the Challenge.

Data location: https://forms.office.com/e/Q357rfhR6m (Please answer few questions to gain access to the dataset)