

Tutorial 3: Cancer Imaging Phenomics Toolkit (CaPTk): A Quantitative Imaging Analysis Tool for Feature Extraction and Predictive Modeling of Clinical Outcomes

Organizers

Spyridon Bakas, University of Pennsylvania, USA Sarthak Pati, University of Pennsylvania, USA

Overview

We seek to give a hands-on tutorial with clear and focused material on a newly developed, light-weight, dynamically growing software platform for analysis of radiographic cancer images, currently focusing on image analysis and machine learning tools for brain, breast, and lung cancer.

The Cancer Imaging Phenomics Toolkit (CaPTk) [www.med.upenn.edu/cbica/captk], is a freely distributed open-source toolkit incorporating algorithms essential for quantitative biomedical image analysis. CaPTk aims to facilitate the swift translation of advanced computational algorithms into routine clinical quantification, analysis, decision making, and reporting workflow, without requiring substantial computational background. Its long-term goal is providing advanced imaging analytics in cancer prediction, precision diagnosis and prognosis, as well as in better understanding the biological mechanisms of cancer development, by utilizing multi-parametric imaging signatures that relate to clinical outcome, response to treatment, as well as cancer molecular characteristics.

During the 1st part of the tutorial, the scientific coordinator of CaPTk (Spyridon Bakas) will present algorithms incorporated in CaPTk. Specific case studies and sample data will be provided to guide the participants towards using these algorithms through both the Graphical User Interface (GUI) and the command-line, for both individual and batch processing. We will attempt to offer a comprehensive analysis pipeline, starting from immediately after downloading multi-file DICOM files from a medical PACS and covering i) NIfTI conversion, while keeping the DICOM header information, ii) pre-processing algorithms, including co-registration of multimodal volumes for assessing voxels in tandem, skull-stripping, careful correction for magnetic field inhomogeneities, denoising, and intensity normalization, iii) segmentation, including brain tumor segmentation, and breast tissue segmentation and density estimation, iv) quantitative feature extraction and parameterization, including morphology, histogrambased, and texture, v) specialized tools for glioblastoma, including a radiogenomic biomarker of EGFR pathway involvement and prediction of clinical outcome; as well as v) targeted neurosurgical planning. To showcase the potential clinical value of specialized predictive modelling, an example case study will be used that can show how integration of CaPTk applications can potentially directly influence patient treatment.

During the 2nd part of the tutorial, the lead developer of CaPTk (Sarthak Pati) will present the various ways that computational imaging scientists (such as the ISBI audience) can incorporate their existing algorithms (available as either C++ classes, or standalone applications written in Python) into the CaPTk's GUI, allowing for direct use of the easy-to-use and lightweight interface, enabling use of their advanced algorithm by non-computational experts, e.g. clinicians, thereby leading to improved visibility/citations of their tools from others. In addition, possible ways to contribute to the feature extraction pipeline will be discussed.

Topics

1. Introduction to Quantitative Biomedical Image Analysis, Radiomics, and Radiogenomics

- 2. Use of CaPTk and incorporated algorithms through GUI and batch-processing (Spyridon Bakas)
- 2.1. Pre-processing algorithms, e.g., bias correction, WhiteStripe intensity normalization
- 2.2. Segmentation
 - 2.2.1. Multimodal Brain Tumor Segmentation (GLISTRboost, ITK-SNAP, DeepMedic)
 - 2.2.2. Breast Tissue Segmentation and Cancer Risk Assessement (LIBRA)
- 2.3. Quantitative Imaging Phenomic Feature Extraction
- 2.4. Specialized Algorithms for Glioblastoma
 - 2.4.1. Radiogenomic biomarker of EGFR pathway activation
 - 2.4.2. Prediction of potential region of recurrence
 - 2.4.3. Targeted neurosurgical planning
- 2.5. Showcasing potential use of the software in a clinical pipeline
- 2.6. Q & A

Coffee Break

3. Extendability of CaPTk (Sarthak Pati)

(How computational scientists can integrate their algorithms in CaPTk, exploiting its ease-to-use GUI)

- 3.1. Incorporating Standalone algorithms/apps written in any language, e.g. MatLab, Python.
- 3.2. Source code integration of algorithms written in C++
- 3.3. Contribution to Quantitative Imaging Phenomic Feature Extraction
- 3.4. Q & A

For the most updated version of the schedule please visit <u>www.med.upenn.edu/cbica/captk-2018-isbi-</u> tutorial.html

Audience

CaPTk targets both computational imaging scientists (i.e., the ISBI audience) and non-computational experts (e.g. clinicians, neuro-scientists). Computational scientists of all levels (e.g., graduate students, post-docs) can i) use existing algorithms implemented in CaPTk in high throughput imaging data (batch processing), as well as ii) incorporate their algorithms into CaPTk's GUI that is based on Qt and relies on ITK, VTK, and OpenCV and therefore take advantage of the easy-to-use interface and user database to gain visibility/citations of their algorithms. Furthermore, CaPTk targets the non-computational experts (e.g., clinicians) by enabling immediate use of complex algorithms for clinically relevant studies through a user-friendly, platform-independent interface.

Prerequisites:

Intermediate knowledge of programming. Use of CMake and ITK will be beneficial. Attendees are expected to have installed CaPTk and downloaded the sample data beforehand, facilitating the hands-on tutorial. The exact version intended to be used during the tutorial, as well as further updates will be posted on www.med.upenn.edu/cbica/captk-2018-isbi-tutorial.html The organizers will also provide USB drives with the source code and sample data, comprising anonymized i) multimodal MRI volumes of patients diagnosed with glioblastoma and ii) digital mammography images of breast cancer patients. Hands-on computer work will be involved, but participants will also be able to follow along in their own time using the presentation material that will be provided, e.g., slides, sample data, source code.