2018 IEEE VIP-CUP

2018 IEEE Video and Image Processing Cup
"Lung Cancer Radiomics"

Supported by:

IEEE Signal Processing Society (SPS)
Concordia Institute for Information System Engineering (CIISE), Concordia University
The Cancer Imaging Archive (TCIA)
Intelligent Signal & Information Processing (I-SIP) Lab

Contact: *Arash Mohammadi*
Contact Email: *arash.mohammadi@concordia.ca*
HomePage: *https://users.encs.concordia.ca/ i-sip/2018VIP-Cup*

2018 IEEE Video and Image Processing (VIP) Cup: Lung Cancer Radiomics

ORGANIZING COMMITTEE
Prof. Arash Mohammadi, Concordia University, Montreal, Canada, arash.mohammadi@concordia.ca
Prof. Amir Asif, Concordia University, Montreal, Canada, amir.asif@concordia.ca
Prof. Anastasia Oikonomou, Sunnybrook Health Sciences, Toronto, Canada, anastasia.oikonomou@sunnybrook.ca
Dr. Keyvan Farahani, Cancer Imaging Program of the National Cancer Institute (NCI), farahank@mail.nih.gov
Justin Kirby, Frederick National Laboratory for Cancer Research, kirbyju@mail.nih.gov

INTRODUCTION
The volume, variety, and velocity of medical imaging data generated for medical diagnosis are exploding. Interpretation of such large amount of diagnostic images, however, highly depends on the experience of the radiologist can be extremely time-consuming. Referred to as Radiomics, the ability to process such large amounts of data promises to decipher the un-decoded information within medical images; Develop predictive and prognosis models to design personalized diagnosis; Allow comprehensive study of tumor phenotype, and; Assess tissue heterogeneity for diagnosis of different type of cancers. More specifically, Radiomics refers to the process of extracting and analyzing several features (e.g., attenuation, shape, size, and location) from medical images with the ultimate goal of obtaining predictive or prognostic models. Segmentation and prediction are considered as critical steps among different processing tasks within the Radiomics pipeline, and are the focus of this competition. The 2018 VIP-CUP challenge is on segmentation and prediction of Lung Cancer Tumor region via screening Computed Tomography (CT) scans. Images from several patients along with the annotations will be provided for training and validation purposes. The evaluation will be performed based on test sets provided closer to the submission deadline.

PARTICIPATION GUIDELINES
Teams satisfying the eligibility criteria outlined below, are invited to participate in the VIP-CUP. Detailed competition instructions together with the data sources will be provided at https://users.encs.concordia.ca/ i-sip/2018VIP-Cup.

Eligibility Criteria: Each team must be composed of: (i) One faculty member (the Supervisor); (ii) At most one graduate student (the Tutor), and; (iii) At least three but no more than ten undergraduates. At least three of the undergraduate team members must be either IEEE Signal Processing Society (SPS) members or SPS student members. Postdocs and research associates are not considered as faculty members. A graduate student is a student having earned at least a 4-year University degree at the time of submission. An undergraduate student is a student without a 4-year degree. Questions about the 2018 VIP-CUP should be directed to Dr. Arash Mohammadi.

- **Eligibility of 3rd and 4th-year Undergraduate Students**: All undergraduate students are eligible to participate in the VIP-Cup. They do not have to be in the first two years of their program.
- **Eligibility of 1st and 2nd-year Masters Students whose Bachelor's degree was 3 years**: First and second-year Master's students are eligible to act a graduate student member of a team, but not as an undergraduate member (regardless of the duration of their bachelor's degree). A team may have at most one graduate student member. Furthermore, graduate students are expected to act as tutors for the team, not as primary participants.

  Since the VIP-Cup is an undergraduate oriented competition, Master's students are not eligible to participate as regular team members, only as graduate student mentors/tutors.
- **Eligibility of Students in a 5-year program that will culminate in a Master's degree**: Students who do not hold a bachelors and are in their 4th year of a 5-year program that will culminate in a master's degree are eligible to participate in the VIP-Cup as regular undergraduate team members.

ENROLLING IN THE PIAZZA INTERACTIVE PAGE
Please join the VIPCUP2018 page created on the Piazza to interact with organizing committee members; ask questions, and; access supporting documents. Below are the steps to join the Piazza page:

- Go to https://piazza.com
- Click the "Sing Up" tab on the top left corner of the page, and select "Students Get Stared"
- In the "Search Schools:" type in "IEEE SPS" and select the pop-up option.
- Select **Summer 2018** term.
- In Class 1 tab:

  - Write “VIPCUP2018: IEEE Video and Image Processing (VIP) Cup 2018”
  - Password is “sps007”.
  - Select “Join as: Student”.
  - Click on Join the Class.
- Proceed with signing up to get access to the VIP-CUP Piazza page

TENTATIVE SCHEDULE

Competition webpage & Preliminary Info:	May 1st, 2018	Finalist Teams Announced:	August 30th, 2018
Initial Train Dataset available:	May 25th, 2018	VIP Cup at ICIP in Athens:	October 7th, 2018
Team registration on IEEE VIP-Cup:	June 10, 2018	Results Submission:	August 5th, 2018

1 Radiomics Segmentation Challenge

The volume, variety, and velocity of medical imaging data is exploding, making it impractical for clinicians to properly utilize the available information resources in an efficient fashion. At the same time, interpretation of such large amount of medical imaging data by humans is significantly error prone reducing the possibility of extracting informative data. The ability to process such large amounts of data promises to decipher the un-decoded information within medical images; Develop predictive and prognosis models to design personalized diagnosis; Allow comprehensive study of tumor phenotype, and; Assess tissue heterogeneity for diagnosis of different type of cancers. Recently, there has been a great surge of interest on Radiomics, which refers to the process of extracting and analyzing several semi-quantitative (e.g., attenuation, shape, size, and location) and quantitative features (e.g., wavelet decomposition, histogram, and gray-level intensity) from medical images with the ultimate goal of obtaining predictive or prognostic models. Radiomics workflow, typically, consists of the following four main processing tasks:

(i) Image acquisition/modality;

(ii) Image segmentation;

(iii) Feature extraction and qualification, and;

(iv) Statistical analysis and model building.

The Radiomics features can be extracted from different imaging modalities including Magnetic Resonance Imaging (MRI); Positron Emission Tomography (PET), and; Computed Tomography (CT). Among these imaging modalities, the CT is the most used because of its imaging sensitivity, high resolution and isotropic acquisition in locating the lung lesions.

The 2018 IEEE VIP-CUP consists of 2 Tasks:

Task 1: ***Segmentation of Tumor region in Lung CT Images***, is the initial task of the 2018 VIP-Cup challenge. *Segmentation (i.e., providing contour or bounding box)* and *Localization (i.e., a class label is supposed to be assigned to each pixel)* are considered as the critical step within the Radiomics workflow, as typically Radiomics features are extracted from segmented sections. Although, manual delineation of the gross tumour is the conventional (standard) clinical approach, it is time consuming and extensively sensitive to inter-observer variability. Automatic and semi-automatic segmentation methods together with their ability to minimize manual error and increase consistency of delineating regions are of paramount importance. In recent years, different tumor segmentation methods have been proposed, which can be categorized generally as follows:

- Tresholding Methods;
- Histogram-based Methodologies;
- Active contours;
- Morphological Methods;
- Deformable Models;
- Clustering approaches;
- Graph-cut Mechanisms;
- Markov Random fields, and;
- Neural Networks and Deep Networks.

Task 2: ***Prediction Analysis***, details will be finalized and announced shortly.

2 Dataset

The challenge dataset contains images from non-small cell lung cancer (NSCLC) subjects selected from NSCLC-Radiomics dataset [1, 2] provided by The Cancer Imaging Archive (TCIA) [3]. The first training data consist of pre-treatment CT scans from 100 subjects, together with manual delineation by a radiation oncologist of the 3D volume of the gross tumor volume. Fig 1 illustrates an example of CT images of lung cancer patients with tumor contours. Test data will be realised, on a later date, including unseen CT images without inclusion of the annotations. Please also consider the following further comments:

- Consider 2D segmentation problem, however, the information from neighbouring slices can be used too.
- There are 2 set of initial training datasets:

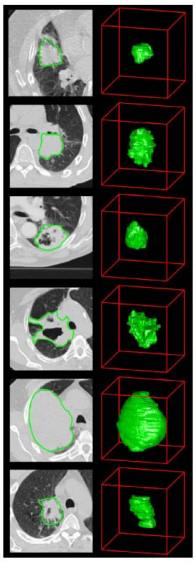

Figure 1: Example computed tomography (CT) images of lung cancer patients. CT images with tumour contours [2].

(i) 100 subjects with all slices, and;

(ii) A selected set of subjects with tumor only slices.

- Use folder names including "StudyID" as an identifier for each subject, e.g., "01-01-2014-**StudyID-42151**".
- Annotations are provided through an extra DICOM file for each patient (Each patient has two folders. One including the original images and one including the annotations). By reading the annotation DICOM file you have access to an attribute called "contour data" that consists of contour points for tumors.
- To view the RTSTRUCT segmentation on top of the images (visualize RTSTRUCT files over the CT scan) you can use the following two options:
  - *Dicompyler*: An open-source platform that can visualize DICOM images and their segmentations; http://www.dicompyler.com/.
  - *Slicer*: Another open-source platform that can visualize DICOM images and their segmentations. This software can also provide reconstruction of other angles (for further understandings); https://www.slicer.org/.

3 Evaluation Criteria

The segmented contours provided by the competitors will be compared against the manual contours for all test images using evaluation metrics similar to the following ones.

C1- Dice Coefficient: This is a measure of relative overlap, where 1 represents perfect agreement and 0 represents no overlap.

$$D = \frac{2|X| \cup |Y|}{|X| + |Y|}, \tag{1}$$

where X and Y are the ground truth and test regions.

C2- Mean surface distance: The directed average Hausdorff measure is the average distance of a point in X to its closest point in Y, i.e.,

$$\vec{d}_{H,\text{avg}}(X,Y) = \frac{1}{|X|} \sum_{x \in X} \min_{y \in Y} d(x,y). \tag{2}$$

The (undirected) average Hausdorff measure is the average of the two directed average Hausdorff measures given by

$$d_{H,\text{avg}}(X,Y) = \frac{\vec{d}_{H,\text{avg}}(X,Y) + \vec{d}_{H,\text{avg}}(Y,X)}{2} \tag{3}$$

Hausdorff distance (95% Hausdorff distance): The directed percent Hausdorff measure, for a percentile r, is the rth percentile distance over all distances from points in X to their closest point in Y. For example, the directed 95% Hausdorff distance is the

point in X with distance to its closest point in Y is greater or equal to exactly 95% of the other points in X. In mathematical terms, denoting the rth percentile as K_r, this is given by

$$\vec{d}_{H,r}(X,Y) = K_r\Big(\min_{y\in Y} d(x,y)\Big)\forall x \in X \tag{4}$$

The (undirected) percent Hausdorff measure is defined again with the mean:

$$d_{H,r}(X,Y) = \frac{\vec{d}_{H,r}(X,Y) + \vec{d}_{H,r}(Y,X)}{2} \tag{5}$$

References

[1] Aerts, Hugo J. W. L., Rios Velazquez, Emmanuel, Leijenaar, Ralph T. H., Parmar, Chintan, Grossmann, Patrick, Carvalho, Sara, Lambin, Philippe, "Data From NSCLC-Radiomics," *The Cancer Imaging Archive*, 2015.

[2] Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Cavalho, S., Lambin, P., "Decoding Tumour Phenotype by Noninvasive Imaging using a Quantitative Radiomics Approach," *Nature Communications*, Nature Publishing Group, 2014.

[3] "The Cancer Imaging Archive (TCIA)," *http://www.cancerimagingarchive.net/*.