Radiomics and Predictive analytics

Can it replace conventional feature detection analysis



Osman Ratib MD, PhD

Division of Medical Imaging Riviera-Chablais Hospital, Vevey Department of Radiology and Medical Informatics University of Geneva





The RADIOMICS revolution...

Characterization and classification of identified lesions...

Radiomics: Images Are More than Pictures, They Are Data¹

Robert J. Gillies, PhD Paul F. Kinahan, PhD Hedvig Hricak, MD, PhD, Dr(hc)

In the past decade, the field of medical image analysis has grown exponentially, with an increased number of pattern recognition tools and an increase in data set sizes. These advances have facilitated the development of processes for high-throughput extraction of quantitative features that result in the conversion of images into mineable data and the subsequent analysis of these data for decision support; this practice is termed radiomics. This is in contrast to the traditional practice of treating medical images as pictures intended solely for visual interpretation. Radiomic data contain first-, second-, and higher-order statistics. These data are combined with other patient data and are mined with sophisticated bioinformatics tools to develop models that may potentially improve diagnostic, prognostic, and predictive accuracy. Because radiomics analyses are intended to be conducted with standard of care images, it is conceivable that conversion of digital images to mineable data will eventually become routine practice. This report describes the process of radiomics, its challenges, and its potential power to facilitate better clinical decision making, particularly in the care of patients with cancer.

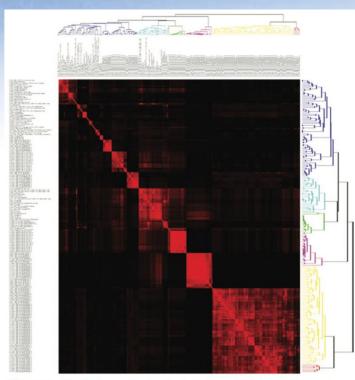
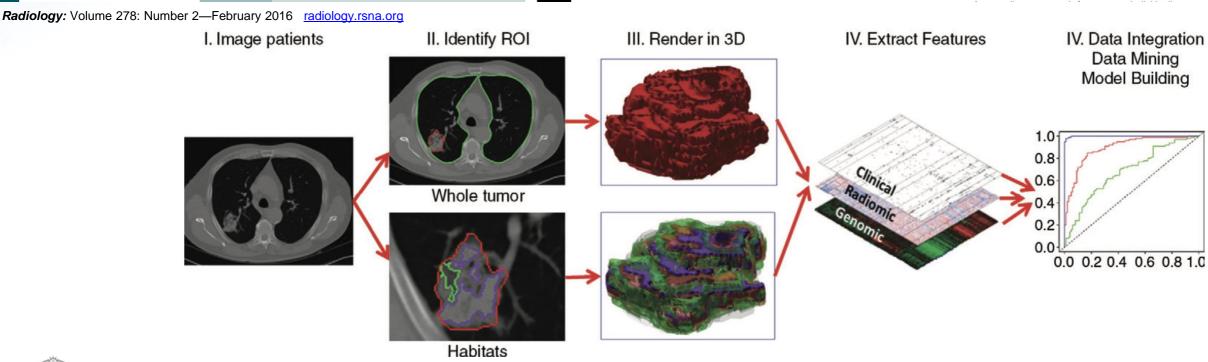


Figure 3: Covariance matrix of radiomic features. A total of 219 features were extracted from each non-small cell lung cancer tumor in 235 patients.









Radiomics in oncology

Multimodality analysis of Lung Cancer

EUROPEAN JOURNAL OF RADIOLOGY

NESSUE

Radiomics and its emerging role in lung cancer research, imaging biomarkers and clinical management: State of the art

Geewon Lee ^{a,b}, Ho Yun Lee (MD, PhD) ^{a,*}, Hyunjin Park ^c, Mark L. Schiebler ^d, Edwin J.R. van Beek ^e, Yoshiharu Ohno ^{f,g}, Joon Beom Seo ^h, Ann Leung ⁱ

Radiomics: Processing of Radiological Imaging Data 1. Image acquisition 2. ROI segmentation 3. Feature extraction Multispectral Imaging Discover Relationships **Genomics Pathology** Clinical Sex/Age Smoking Survival

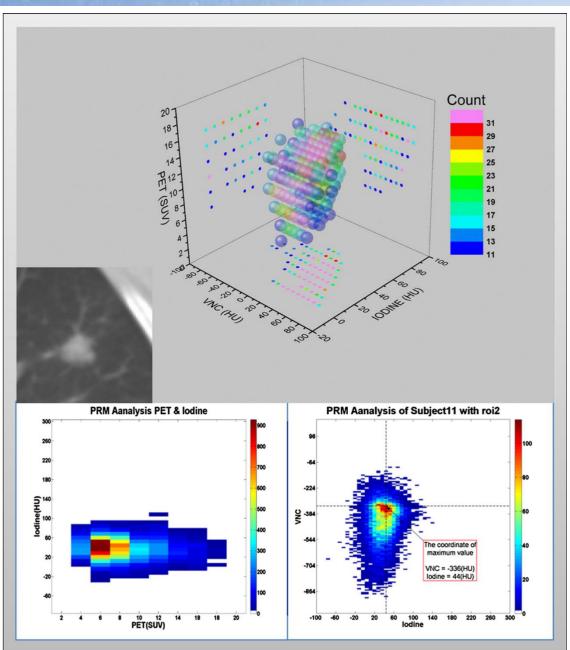


Fig. 4. Multispectral analysis.

Voxel-by-voxel correlations of virtual non-contrast images, iodine from contrast-enhanced CT scans, and maximum standard uptake value (SUVmax) from PET scans are shown for invasive adenocarcinoma





Structured Big DATA for data mining

Dr. Watson - from IBM



IBW MD

Robotics

IBM's Automated
Radiologist Can Read
Images and Medical
Records

Software that can read medical images and written health records could help radiologists work faster and more accurately.

ost smart software in use today specializes on one type of data, be that interpreting text or guessing at the content of

photos. Software in development at IBM has to do all those

at once. It's i

The software philosopher identify anat as CT scans, medical reco

Avicenna is i speed up the cardiology as up using and

Execution and the second secon

Mahmood, a IBM's Avicenna software highlighted possible embolisms on this CT scan in

California, and chief scientist on the project, says that her team and others in the company are already getting ready to start testing the software outside the lab on large volumes of real patient data. "We're getting into preparations for commercialization," she says.







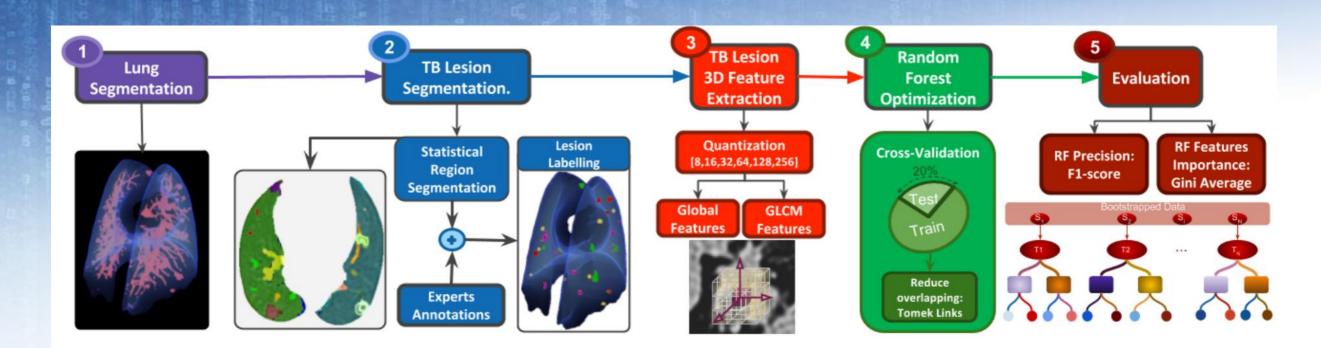


Figure 1. Fully-automatic radiomics workflow for the extraction of informative features on the lung parenchyma: 1) Lung segmentation and airway tree extraction; 2) Selection of relevant volumes employing the Statistical Region merging method matched with the expert annotations of lesions; 3) Extraction of texture features from each volume at 8, 16, 32, 64, 128 and 256 levels of quantization; 22 features are extracted from the Grey Level Co-Occurrence Matrix (GLCM) and 4 are global descriptor of the volume (Mean, Median, Maximum and Minimum); 4) Optimization of the Random Forest (RF) hyperparameters (number of trees, minimum number of samples for split and the maximum number of features to evaluate per node); The optimal Random Forest (RF) classifier is computed per quantization level and number of features employed. The optimization employ a grid search process with 100-fold cross validation where the training data (80% of the total) in each fold is filtered employing Tomek Links¹⁰ to handle class imbalance; 5) Two-fold evaluation: a) The weighted F₁-score is employed as a measure of the classification quality of the most frequent TB lesion types; b) The importance of each feature is evaluated using as merit figure the Gini importance.







Question 1:

Do computers really need images?

Why not analyze the raw data directly?





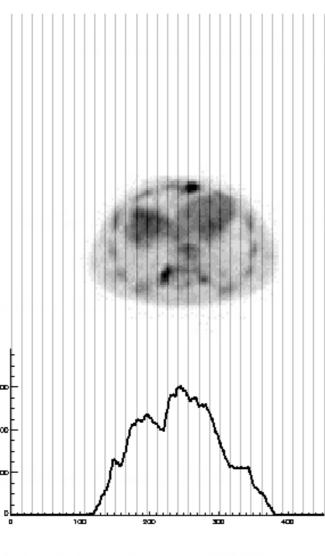


Radiomics feature extraction from raw data

Analyzing sinograms instead of reconstructed images

(True) Emission Volume

Sinogram (stored data)



Forward Projection

> angle 0

Theta (angl

Rho (offset)

intensity profile:

reated by Adam Leon Kesner, PhD, Medical Physicist :

- University of Colorado, Anschutz Medical Campus, Aurora, Colorado, USA

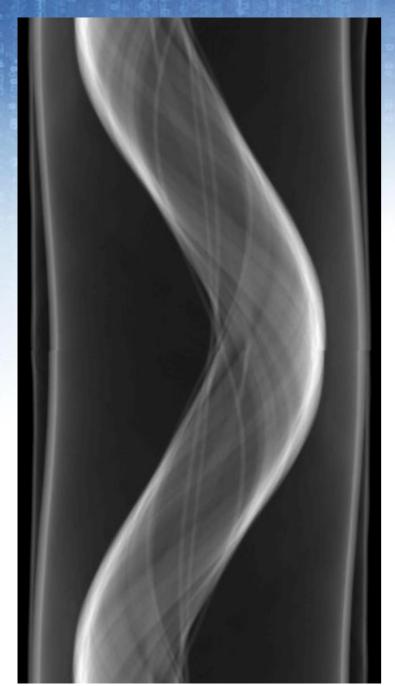




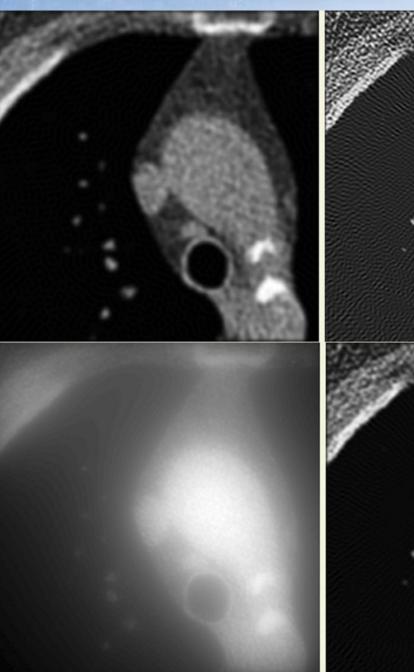


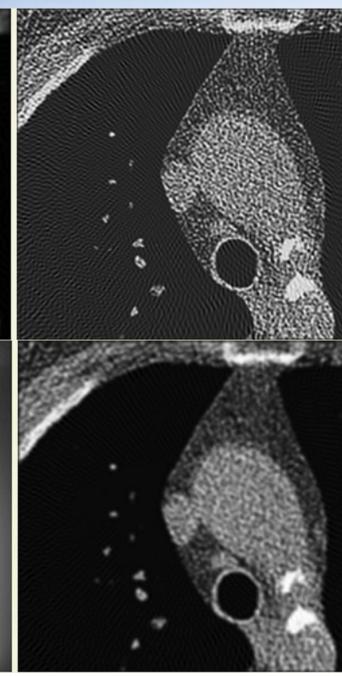
Radiomics feature extraction from raw data

Analyzing sinograms instead of reconstructed images















Question 2:

Will computers do better with raw data that with images?







Question 3:

How can we validate the accuracy of computer analysis of raw data?







Question 4:

What do we need to train a computer and evaluate its performance?













