

## Abstract

### **Radiomics of pulmonary nodules as a predictor of lung cancer in SOS: a monocenter prospective interventional study for lung cancer early diagnosis with digital chest tomosynthesis**

**Adriano De Maggi**

Scuola di Specializzazione in Fisica Medica, Università degli Studi di Torino  
AO Santa Croce e Carle, Cuneo

**Purpose:** Lung cancer is the leading cause of cancer-related death around the world. In 2008 there were nearly 1.6 million new cases worldwide, accounting for the 12.7% of all new cancer diagnoses[1]. Despite decreasing trends in smoking and resulting decrease in lung cancer mortality, the population at risk for lung cancer continues to be large[2] and the overall survival of lung cancer patient is still low. In the past years, several programs have been developed to screen for lung cancer using low-dose chest computed radiography (CT). Even if Computed tomography (CT) is considerably more accurate in detecting lung nodules than chest radiography CT has significantly higher radiation dose and higher cost. In Santa Croce and Carle Hospital, in this context, enrolled around 2000 subjects in a monocenter prospective no-profit interventional study for lung cancer early diagnosis with Digital chest tomosynthesis (DTS). Digital tomosynthesis is a limited angle tomography that allows reconstruction of multiple image planes from a set of projection data acquired over a relatively small angle of X-ray tube movement[3]. The interpretation of radiological 2D and 3D chest images results difficult due to the presence of the ribs, vessels and the heart. During the last ten years several automated methods for computerized detection of nodules have been developed. The main obstacles in nodules detection and characterization are the non-lung structures and the intra and inter reader repeatability [4]. Over the last two decades, various methods have been developed to improve diagnostic accuracy by reducing the number of incorrect diagnoses. Computer Aided Detection (CAD) methods have become a central tool in thoracic images, in particular for the detection and characterization of lung nodules[5]. In such setting the CAD system have a high sensitivity at the cost of a low specificity, which is provided by the analysis of the radiologist. The growth of technology in the last ten years has boosted the introduction of artificial intelligence in medical field such as Machine learning technique and radiomics. The process of analyzing images through machine learning often starts with algorithms that deal with the extraction of features that are considered relevant in making the prediction or diagnosis of interest. The machine learning algorithm then identifies the best combination of these image features for classifying the image or computing some metric for a given image region[6]. The

aim of this work is to study the radiomics indices of pulmonary nodules as a predictor of lung cancer in SOS. Developed different machine learning to classify lung nodule on a training set, test it on an independent test set and compare the diagnostic accuracy of ML models respect to visual analysis and radiology semantic.

**Methods and materials:** Between December 2010 and August 31, 2018, 1843 subjects were enrolled in the monocenter SOS “Studio Osservazionale” clinical trial (NCT number NCT03645018). Participants considered eligible were smokers or former smokers aged 45 to 75 years, with a smoking history of at least 20 pack-years; for former smokers, the maximum time since quitting smoking was 10 years. [7]. Even if the overall number of patients used in this investigation is relatively high and complex, 1843 patients followed-up through 7 years, the number of events, i.e. lung cancer, is relatively low, about 1%. It is known that small sample sizes increase both the type-I (incorrectly detecting a difference) and type-II (not detecting an actual difference) error rates. Images coming from SOS trial were analyzed. Since DTS exams are a sequence of bidimensional images as a first step DTS images were transformed in 3D matrix with an in-house developed script in MATLAB. 3D matrix were wrote as DICOM CT images in order to analyse them with segmentation software. No conversions were made to rescale pixel values (PV) to CT HU units. Then segmentation was carried out in PET Encore (MIM software Inc., Cleveland, OH, USA) workstation. A radiologist and a medical physicist manually contoured all the nodules. The lung nodules contour was first applied to the central reconstructed tomosynthesis slice in which the lesion was more clearly visible and then manually propagated to the next slices taking to account the presence of ribs overlapping lung structures and "blurring" artifacts. The contour were then adapted to the nodule. Texture features were extracted from the DICOM-RT structures and DTS images using the open-source and validated PORTS texture radiomics toolkit (Nyflot MJ, 2015, Hatt M, 2016), developed at University of Washington. Several features were extracted: 1<sup>st</sup> order (histogram) and 2nd order based on different matrices. To calculate the images features shall be isotropic. We hence resampled using a spline function at a fixed voxel dimension of 3mmx3mmx3mm from 0.2mmx0.2mmx3mm of the original pixel's size.

In this work two different populations with no statistical difference in the clinical variables were used to train and test the machine learning algorithm. The training dataset consisted of 85 patients and 120 nodules while the test dataset consisted in 132 patients and 209 nodules. All the nodules were characterized by size, location, type and LUNGRADS classificaton. Several machine learning techniques were applied to the train dataset: Stochastic Gradient Boosting, Support Vector

Machines and Random Forest. These different algorithms were trained on the train dataset. Predictive models were tested with different values of characteristic parameter.

**Results:** Radiomics calculation failed in 38 of 329 nodules. Besides this not all radiomics indexes were correctly calculated. Only 50/329 (15%) nodules had Neighboring Gray Tone Difference Matrix (NGTDM) radiomics features calculated, while about half, 195/329 (60%), of the Gray Tone Spatial Dependence Matrix (GTSDM) were computed. Gray Level Size Zone (GLSZM) radiomics features were computed in (265/329) 80% of the cases. The area under the receiver operating characteristic curve (AUC) was used to quantify the association between lung cancer and radiomic features. The AUC between lung cancer and radiomic features, in the entire dataset, ranged from 0.496 in histogram\_skewness to 0.845 in histogram\_entropy. Most of the GTSDM features were significantly related to lung cancer, in particular gtsdm\_sum\_entropy had an AUC of 0.841 [0.781;0.901] .To underline the good behavior of the gtsdm\_sum\_entropy in Figure 1 is shown the distribution of the variable in lung cancer and non lung cancer patients while in Figure 2 is shown the distribution of the GTSDM\_sum\_entropy variable around the nodule.

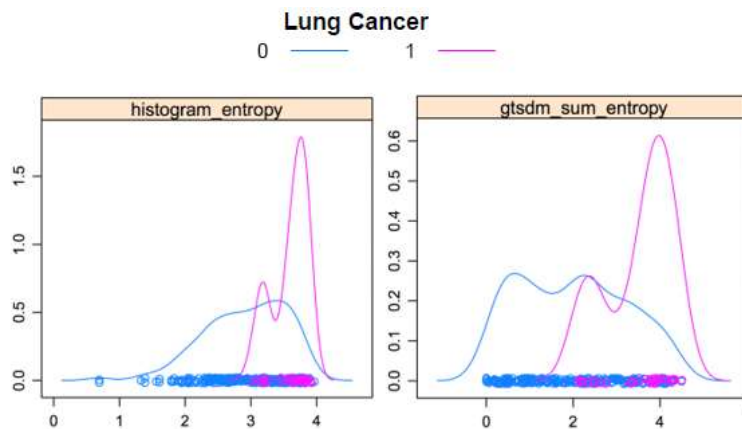


Figure 1 Histograms of gtsdm\_sum\_entropy features

The optimized predictive models were then applied to the test dataset to classify nodules between lung cancers and non lung cancers. Confusion Matrix between predicted and actual values were evaluated for each predictive model along with accuracy metrics. SOS dataset was used to compare the different predictive models on a dataset of 209 nodules of which 20 lung cancer, the binary visual was able to identify 19 positive nodules, the LUNGRADS classification correctly identify 13 positive nodules, GBM and Random Forest respectively detected 8 and 6 lung cancer.

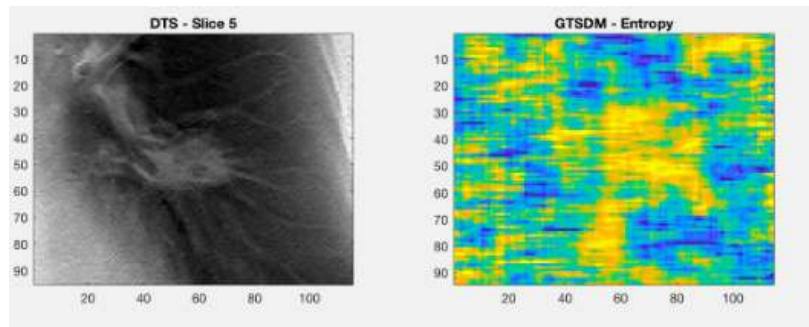


Figure 2 Distribution of the GTSDM entropy around the nodule

**Conclusion:** In this investigation, predictive methods that can classify nodule as lung cancer or non lung cancer were evaluated. All the segmented nodules were imported on PORTS and radiomics indexes were calculated. Radiomics calculation of segmented nodules failed in 38 of 329 nodules. More than the 80% of the failed nodules had diameter below 8 mm. This behavior is probably due to the limited amounts of voxels in the structure that does not allow the calculation of radiomics indexes spread over limited volume

We found several features with that seem to have a role in the prediction of lung cancer, such as histogram\_entropy (AUC=0.845) and gtsdm\_sum\_entropy (AUC=0.841) . We use three different machine learning techniques to define a predictor model, results show that there were no significant difference between gbm and random forest in terms of summary statistical parameters such as accuracy and Cohen's Kappa. Support vector machines at the moment results to be not a good prediction model. All predictive models were good in detecting lung cancer but with a high rate of false positive. Indeed, negative predictive value (NPV) was quite low, less than 0.4. Since the number of lung cancers are low respect to the non lung cancer case the algorithms suffers in founding a high NPV. The predictive model was applied to an independent population of DTS demonstrating a good sensitivity with a relatively low specificity. The model could hence be of helps to radiologist in discriminating positive from negative lung cancers, as an assistance to their evaluation. To further evaluate the model we have in program the validation of the predictive model on an external independent cohort.

## References

- [1] M. Malvezzi, P. Bertuccio, F. Levi, C. La Vecchia, and E. Negri, "European cancer mortality predictions for the year 2012," *Ann. Oncol. Off. J. Eur. Soc. Med. Oncol. ESMO*, vol. 23, no. February, pp. 1044–1052, 2012.
- [2] D. R. Aberle *et al.*, "Reduced lung-cancer mortality with low-dose computed tomographic screening," *N. Engl. J. Med.*, vol. 365, no. 5, pp. 395–409, 2011.
- [3] J. T. Dobbins and H. P. Mcadams, "Chest tomosynthesis: technical principles and clinical update," *Eur. J. Radiol.*, vol. 72, no. 2, pp. 244–251, 2009.

- [4] K. N. Jeon *et al.*, “Computer-aided nodule detection and volumetry to reduce variability between radiologists in the interpretation of lung nodules at low-dose screening computed tomography,” *Invest. Radiol.*, vol. 47, no. 8, pp. 457–61, Aug. 2012.
- [5] L. Giger, R. Hoffmann, A. Schmidt, E. Metz, and M. Beblen, “An ‘Intelligent’ Workstation for Computer-aided Diagnosis,” pp. 647–656.
- [6] B. J. Erickson, P. Korfiatis, Z. Akkus, and T. L. Kline, “Machine learning for medical imaging,” *Radiographics*, vol. 37, no. 2, pp. 505–515, 2017.
- [7] A. Terzi, L. Bertolaccini, A. Viti, and L. Comello, “Lung Cancer Detection with Digital Chest Tomosynthesis,” *J. Thorac. Oncol.* vol. 8, no. 6, pp. 685–692, 2013.