Master's Thesis Engineering Technology 2022-2023

Radiomics for NSCLC ^{18}F-FDG PET CT imaging: Correlations and prognostic model: a statistical study

Daan Baert
Master IW Nucleaire Technologie

Kobe Verlinden
Master IW Nucleaire Technologie

OBJECTIVE

Lung cancer has the second highest mortality rate among cancer phenotypes, as shown in Figure 1, and is linked with an increasing incidence of 2.4 million by 2035, raising the need for more efficient diagnosis and accurate prognosis.

Figure 1: 5-year relative survival [1]

Radiomics can be a useful tool to support both purposes. This master's thesis aims to uncover the possible distinguishing powers of radiomics data between healthy and lung cancer tissue.

MATERIALS AND METHODS

- ◊ Acquisition: A cohort of 49 patients was diagnosed with NSCLC by ^{18}F-FDG PET imaging for more accurate staging and underwent a lobectomy.
- ◊ Segmentation: The Volume of interest (VOI) of the lung lesion was first segmented in the ACCURATE tool. Then, the previously segmented VOIs are exactly translated in the ACCURATE tool to the opposite lung including only healthy lung tissue. Radiomics features of both VOIs were extracted, as visualized in Figure 2.
- ◊ Data analysis: Paired t-test, data visualisation (dendogram), Principal Component Analyses (PCA), and discriminative models in Matlab are used for the statistical analysis.

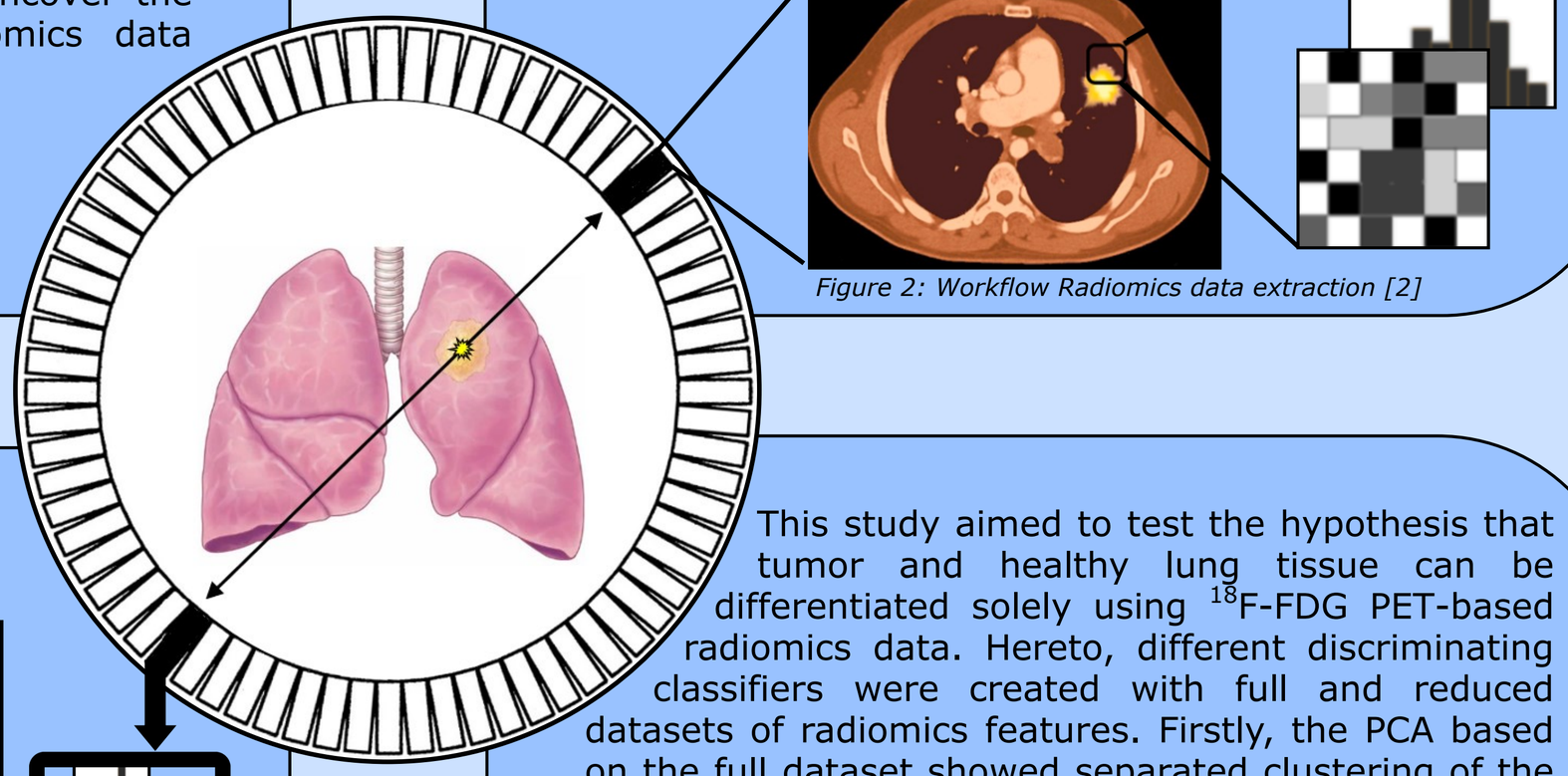

Figure 2: Workflow Radiomics data extraction [2]

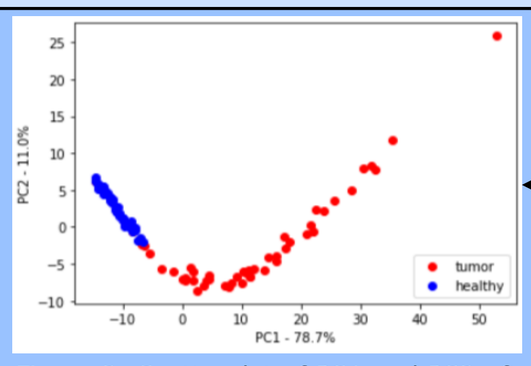

Figure 3: Scatterplot of PC1 and PC2 after noise reduction indicating clustering of healthy and tumor tissue.

A paired t-test unveiled 69 irrelevant features in differentiating both tissue types.

The PCA indicates separate clustering of both tissue types, with the noise-reduced dataset of 269 features showing the best results, as shown in Figure 3 . A set of 30 features still performed adequately.

The Fine and Weighted K-Nearest Neighbors, and Bagged Trees and Subspace Discriminant ensemble learning models show an accuracy of 98.3%. The Fine KNN outperformed the others with a predicting accuracy of 89.5%.

RESULTS

This study aimed to test the hypothesis that tumor and healthy lung tissue can be differentiated solely using ^{18}F-FDG PET-based radiomics data. Hereto, different discriminating classifiers were created with full and reduced datasets of radiomics features. Firstly, the PCA based on the full dataset showed separated clustering of the tumor and the healthy tissue. This differentiation was even more clear after noise-suppression, reducing the dataset to the 269 features with the highest loading scores. A set of 30 features still performed adequately. The second major finding was that four different machine learning models attained an accuracy of 98.3% in predicting the tissue type: Fine KNN, Weighted KNN, Bagged trees, and Subspace Discriminant models. The Fine Tree model achieved an accuracy 96.7% using only one radiomics feature: glcmFeatures2Davg-inverse difference.

This study adds promising results to the rapidly expanding field of radiomics. The identified models may assist in the further development of (semi) automated tools and provide a basis for further research in ^{18}F-FDG PET- based lung cancer diagnosis.

CONCLUSION

Supervisors / Co-supervisors / Advisors: Prof. Dr. Liesbet Mesotten, Dr. Elien Derveaux, Prof. Dr. Dirk Valkenborg, Prof. Dr. Brigitte Reniers en Prof. Dr. Peter Adriaensens

[1] "Cancer of the lung and Bronchus - Cancer stat facts," SEER. [Online]. Available: https://seer.cancer.gov/statfacts/html/lungb.html. [Accessed: 7-May-2023].

[2] R. Manafi-Farid, N. Karamzade-Ziarati, R. Vali, F. M. Mottaghy, and M. Beheshti, "2-[18F]FDG PET/CT radiomics in lung cancer: An overview of the technical aspect and its emerging role in management of the disease," Methods, vol. 188, pp. 84–97, 2021.

Ziekenhuis Oost-Limburg · future health BY ZOL GENK

De opleiding industrieel ingenieur is een gezamenlijke opleiding van UHasselt en KU Leuven

UHASSELT · KU LEUVEN