Machine Learning Based Malignancy Prediction in Thyroid Nodules Malignancy: Radiomics Analysis of Ultrasound Images

Sajad P. Shayesteh, Mostafa Nazari, Ali Salahshour, Atlas Haddadi Avval, Ghasem Hajianfar, Mohsen Araabi, Maziar Khateri, Hamid Abdollahi, Hossein Arabi, Isaac Shiri, Habib Zaidi, *Fellow, IEEE*

Abstract– **The aim of this work was to use sonographic image features as biomarker to assess the malignancy of thyroid nodules in patients recommended to FNA according to ACR TI-RADS guideline. Two hundred and ten patients with FNA test report were included in this study. Eighty Different quantitative radiomic features were extracted from sonographic images. Minimum Redundancy Maximum Relevance (MRMR) and logistic regression (LR) algorithms were used as feature selector and classifier, respectively. The evaluation of the models was performed using accuracy, sensitivity, specificity, and area under the receiver operating characteristic curve (AUC). AUC of MRMR feature selection algorithm and LR classifier was 0.87 (with accuracy of 0.74, sensitivity of 0.85 and specificity of 0.60). In the validation dataset, the AUC was 0.92 (with accuracy of 0.70, sensitivity of 0.81 and specificity of 0.58). The proposed model could be potentially used as alternative to FNA as non-invasive tools in clinical setting.**

Index Terms—**Radiomics, Ultrasound, Thyroid, Machine Learning**

I. Introduction

ULTRASOUND is widely used for benign and malignant thyroid nodules differentiation as it is noninvasive, and cost-effective examination for target evaluation. According to the report of American College of Radiology (ACR) Thyroid Imaging Reporting and Data System (TI-RADS) Committee, visual interpretation of this ultrasound based features are used for thyroid nodules scoring and sum of these factors used to determine probability for malignancy of thyroid nodules [1-3].

Radiomic features analysis is the transformation of medical images into quantitative and comprehensive features that enables high-dimentional data mining and could guide treatment decision and response prediction [4-7]. The ability of radiomic features and machine learning based analysis as accurate, reliable and non-invasive biomarker for cancer diagnosis, prognosis and response prediction have been investigated in previous studies [8-12]. Also, application of radiomics models in better characterizing thyroid nodules and in identifying nodules or tumors behavior is a nowadays important research topic [2, 11, 13, 14]. The aim of this study was to use the sonographic images features as biomarker to assess the malignancy of thyroid nodules in patients recommended to FNA according to ACR TI-RADS guideline.

II. Materials and Methods

We performed a retrospective study of 210 patients (20 males and 55 females; mean age, 45.5 ±13.1 years; range, 22 to 76 years) selected randomly as training and 40 patients (20 males and 55 females; mean age, 45.5 ±13.1 years; range, 22 to 76 years) selected randomly as external validation dataset. Patients with TI-Rad 3, 4 and 5 according to TI-RADS report, underwent ultrasound-guided FNA and pathological examination reports were included in this study.

Segmentation of thyroid nodules on the gray scale ultrasound images was done by a seven years' experience radiologist using the freely available LIFEx software (version v5.10, lifexsoft.org). Nodules were first reviewed and then were drawn on the US images. Various radiomic features from different feature sets including intensity, shape and texture-based features were extracted from gray scale ultrasound images. Extracted features included shape features (n=17)[14], intensity histogram features (n=9), intensity direct (n=19), neighbor intensity difference (n=5), co-occurrence matrix features (n=19), and gray level run-length matrix features (n=11) [15-18].

Prior model building, we applied MRMR (Max-Relevance and Min-Redundancy) Feature selection method. We used supervised logistic regression (LR) algorithm to building Radiomics model. For grid based hyperparameter tuning performed bootstrap method [19]. We used AUC, accuracy, sensitivity and specificity for models evaluation.

Manuscript was submitted Dec 20, 2020. This work was supported by the Swiss National Science foundation under grant SNFN 320030_176052.

S P.Shayesteh and M. Araabi are with Department of Physiology, Pharmacology and Medical Physics, Alborz University of Medical Sciences, Karaj, Iran (e-mail: shayeste_sajad@yahoo.com, Arabic.moh@gmail.com)

M. Nazari is with Department of Biomedical Engineering and Medical Physics, Shahid Beheshti University of Medical Sciences, Tehran, Iran. (e-mail: msnaa1392@gmail.com)

A. Salahshour is with department of Radiology, Alborz University of Medical Sciences, Karaj, Iran (e-mail: Alisalahshoor11@yahoo.com)

A. Haddadi Avval is with school of Medicine, Mashhad University of Medical Sciences, Mashhad, Iran (e-mail: atlas.ha78@gmail.com)

G. Hajianfar is with Rajaie Cardiovascular Medical and Research Center, Iran University of Medical Science, Tehran, Iran (e-mail: hajianfar.gh@gmail.com)

M. Khateri is with finetech in Medicine Research Center, Iran University of Medical Sciences (IUMS), Tehran, Iran (e-mail: mazym4004@gmail.com)

H. Abdollahi is with department of Radiologic Sciences and Medical Physics, Kerman University of Medical sciences, Kerman, Iran (e-mail: hamid_rbp@yahoo.com)

I. Shiri, H. Arabi and H. Zaidi are with division of Nuclear Medicine and Molecular Imaging, Geneva University Hospital, CH-1211, Geneva 4, Switzerland. (e-mail: Isaac.shirilord@unige.ch, Hossein.Arabi@unige.ch habib.zaidi@hcuge.ch).

H. Zaidi is with Geneva University Neurocenter, Geneva University, Geneva, Switzerland.

III. RESULT

Results of our machine learning based analysis using MRMR feature selection algorithm and LR classification model was shown in Fig. 1 for training and validation dataset. As was shown, in training data set, AUC of MRMR feature selection algorithm and LR classifier was 0.87 (with accuracy of 0.74, sensitivity of 0.85 and specificity of 0.60). And, in validation data set, AUC was 0.92 (with accuracy of 0.70, sensitivity of 0.81 and specificity of 0.58).

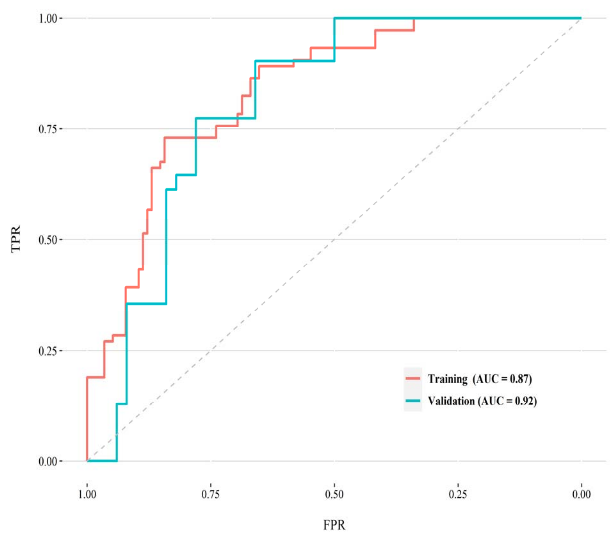

Fig 1. The results of MRMR feature selection and logistic regression model in training and validation set by using 1000 bootstrapping sample.

IV. DISCUSSION AND CONCLUSION

In this study results of training dataset was validated with external validation dataset which confirming the reliability and reproducibility of thyroid nodules malignancy prediction model [11, 15]. At first, we used MRMR feature selection algorithm to reduce overfitting and throw out redundant quantitative features. In MRMR, the features are ranked according to the minimal redundancy and maximal relevance to target output [16]. In this study we used LR classifier to investigate the ability of machine learning based algorithms for thyroid nodules malignancy prediction. Proposed model could be potentially used as alternative to FNA as non-invasive tools in clinical routines. The main limitation of current study was the size of training sets and applying only one feature selector and one classifier for madling. Testing different combination of feature selection and classification algorithms could results in high accurate model for diagnostic and prognostic model. Also, the reproducibility and repeatability of radiomics features were neglected in model building and should be considered in future studies [17, 18].

REFERENCES

[1] F. N. Tessler, W. D. Middleton, and E. G. Grant, "Thyroid imaging reporting and data system (TI-RADS): a user's guide," *Radiology,* vol. 287, no. 1, pp. 29-36, 2018.

[2] J. Liang *et al.*, "Predicting malignancy in thyroid nodules: radiomics score versus 2017 American College of Radiology Thyroid Imaging, Reporting and Data System," *Thyroid,* vol. 28, no. 8, pp. 1024-1033, 2018.

[3] T. Liu *et al.*, "Comparison of the application of B-mode and strain elastography ultrasound in the estimation of lymph node metastasis of papillary thyroid carcinoma based on a radiomics approach," *International journal of computer assisted radiology and surgery,* vol. 13, no. 10, pp. 1617-1627, 2018.

[4] H. Abdollahi, I. Shiri, and M. Heydari, "Medical Imaging Technologists in Radiomics Era: An Alice in Wonderland Problem," (in eng), *Iran J Public Health,* vol. 48, no. 1, pp. 184-186, Jan 2019.

[5] G. Hajianfar *et al.*, "Noninvasive O6 Methylguanine-DNA Methyltransferase Status Prediction in Glioblastoma Multiforme Cancer Using Magnetic Resonance Imaging Radiomics Features: Univariate and Multivariate Radiogenomics Analysis," *World Neurosurgery,* vol. 132, pp. e140-e161, 2019/12/01/ 2019.

[6] S. Mostafaei *et al.*, "CT imaging markers to improve radiation toxicity prediction in prostate cancer radiotherapy by stacking regression algorithm," (in eng), *Radiol Med,* vol. 125, no. 1, pp. 87-97, Jan 2020.

[7] M. Nazari *et al.*, "Noninvasive Fuhrman grading of clear cell renal cell carcinoma using computed tomography radiomic features and machine learning," (in eng), *Radiol Med,* vol. 125, no. 8, pp. 754-762, Aug 2020.

[8] I. Tsougos, A. Vamvakas, C. Kappas, I. Fezoulidis, and K. Vassiou, "Application of radiomics and decision support systems for breast MR differential diagnosis," *Computational and mathematical methods in medicine,* vol. 2018, 2018.

[9] Y. S. Choi *et al.*, "Machine learning and radiomic phenotyping of lower grade gliomas: improving survival prediction," *European Radiology,* pp. 1-9, 2020.

[10] H. Wang *et al.*, "Machine learning-based multiparametric MRI radiomics for predicting the aggressiveness of papillary thyroid carcinoma," *European Journal of Radiology,* vol. 122, p. 108755, 2020.

[11] H. Zhou *et al.*, "Differential Diagnosis of Benign and Malignant Thyroid Nodules Using Deep Learning Radiomics of Thyroid Ultrasound Images," *European Journal of Radiology,* p. 108992, 2020.

[12] M. Nazari, I. Shiri, and H. Zaidi, "Radiomics-based machine learning model to predict risk of death within 5-years in clear cell renal cell carcinoma patients," (in eng), *Comput Biol Med,* vol. 129, p. 104135, Nov 23 2020.

[13] M. Sollini, L. Cozzi, A. Chiti, and M. Kirienko, "Texture analysis and machine learning to characterize suspected thyroid nodules and differentiated thyroid cancer: Where do we stand?," *European journal of radiology,* vol. 99, pp. 1-8, 2018.

[14] T. Liu *et al.*, "Prediction of lymph node metastasis in patients with papillary thyroid carcinoma: a radiomics method based on preoperative ultrasound images," *Technology in cancer research & treatment,* vol. 18, p. 1533033819831713, 2019.

[15] Y. Sim, S. E. Lee, E.-K. Kim, and S. Kim, "A Radiomics Approach for the Classification of Fibroepithelial Lesions on Breast Ultrasonography," *Ultrasound in Medicine & Biology,* 2020.

[16] M. Radovic, M. Ghalwash, N. Filipovic, and Z. Obradovic, "Minimum redundancy maximum relevance feature selection approach for temporal gene expression data," *BMC bioinformatics,* vol. 18, no. 1, p. 9, 2017.

[17] M. Edalat-Javid *et al.*, "Cardiac SPECT radiomic features repeatability and reproducibility: A multi-scanner phantom study," (in eng), *J Nucl Cardiol,* Apr 24 2020.

[18] I. Shiri *et al.*, "Repeatability of radiomic features in magnetic resonance imaging of glioblastoma: Test-retest and image registration analyses," (in eng), *Med Phys,* vol. 47, no. 9, pp. 4265-4280, Sep 2020.