2026 JULY 19–22 VANCOUVER, BC
JOINT AAPM | COMP MEETING

From Radiomics to TI-RADS: A Stability-Aware Clinically-Interpretable Dictionary for Explainable Thyroid Nodule Classification on Ultrasound

Mohammad Salmanpour [1,2,3], Shahram Taeb, PhD[4], Mohammad Ayazi [5], Siavash Hosseinpour Saffarian, Somayeh Sadat Mehrnia [7], Mehdi Maghsudi [3], Ali Fathi Jouzdani [3], Amirhossein Shahsavani Asl[8], Arman Rahmim[1,2] and Ilker Hacihaliloglu [2,9]

(1) Department of Basic and Translational Research, BC Cancer Research Institute; (2) Department of Radiology, University of British Columbia; (3) Technological Virtual Collaboration Company (TECVICO CORP.), Vancouver, BC, Canada; (4) Department of Radiology, School of Paramedical Sciences, Guilan University of Medical Sciences, Rasht, Iran; (5) School of Medicine, Kermanshah University of Medical Science, Kermanshah, Iran; (6) College of Interdisciplinary Science and Technology, Faculty of Intelligent Systems, University of Tehran, Tehran, Iran; (7) Department of Radiology Technology, Faculty of Allied Medical Sciences, Shahid Beheshti University of Medical Sciences, Tehran, Iran; (8) Faculty of Medicine, Kerman University of Medical Sciences, Kerman, Iran; (9) Department of Medicine, University of British Columbia, Vancouver, BC, Canada,

ABSTRACT

- This study developed an **explainable ultrasound (US)-based thyroid nodule classification (TNC) framework** by linking radiomic features (RFs) with American College of Radiology Thyroid Imaging Reporting and Data System (ACR TI-RADS) descriptors.
- A total of 5,542 thyroid nodules from three multicenter datasets were analyzed. RFs were extracted using PyRadiomics, and multiple machine learning (ML) models were tested. SHapley Additive exPlanations (SHAP) was used to explain the model results.
- Key radiomic features were associated with important ACR TI-RADS findings, including echogenicity, margins, shape, and echogenic foci. The best model achieved a strong receiver operating characteristic area under the curve (ROC-AUC) of **0.941 ± 0.005**.
- Overall, the framework improves the **transparency and clinical usefulness** of artificial intelligence (AI)-assisted thyroid cancer risk assessment.

INTRODUCTION

- Thyroid nodules are commonly detected by US, but **distinguishing benign from malignant** nodules remains challenging due to overlapping imaging features and operator-dependent assessment, making accurate risk stratification important for biopsy and treatment decisions [1].
- This study links RFs with ACR TI-RADS descriptors to create an explainable AI-based framework for reproducible thyroid nodule classification [2].

METHODS AND MATERIALS

A. Development of a Clinically Informed Feature Interpretation Dictionary. Nine experts, including four physicians, three physicists, one radiology expert, and one biologist, refined the feature mappings. A Radiological and Biological Dictionary was developed to link ACR TI-RADS ultrasound descriptors with IBSI-compliant RFs. These mappings used first-order, shape, and texture RFs related to TI-RADS findings.

B. Explainable Machine Learning Classification Task. Chart 1 shows the main steps performed for the explainable ML classification task.

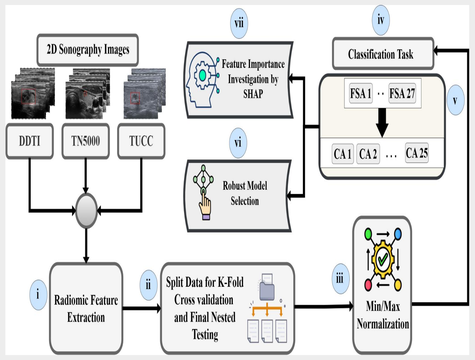

Chart 1. Overview of the proposed radiomics-based ML pipeline for thyroid US image classification.

RESULTS

Table 1 summarizes the proposed radiological and biological dictionary, linking radiomic features to TI-RADS semantic descriptors.
Chart 2 presents the performance of the top five models selected using the proposed stability-aware model selection framework, while Chart 3 shows SHAP analysis demonstrating consistent feature importance patterns across the best-performing models.

Table 1 provides an overview of the summarized RF dictionary, systematically mapping TI-RADS semantic descriptors to their corresponding RFs.

TI-RADS Category	Semantic Descriptor	Risk Level	Mapped Radiomic Feature(s)	Quantitative Signature (Interpretation)
Composition	Cystic / Spongiform	Low	Shape: Circularity, Compactness Shape: Elongation	High Value: Indicates round/regular shape. Low Value: Indicates isodiametric shape.
	Mixed / Solid	Increased	First Order: Entropy	Low Value: Indicates heterogeneity. High Value: Indicates complexity
Echogenicity	Hyperechoic / Isoechoic	Low / Mild	First Order: Mean GLDM: Large Dependence High Gray-Level Emphasis (LDHGLE)	High Value: Indicates high average brightness. High Value: Indicates coarse, bright texture.
	Hypoechoic / Very Hypoechoic	Mod. / High	First Order: Mean GLRLM: Low Gray-Level Run Emphasis (LGRE) NGTDM: Coarseness	Low Value: Indicates overall darkness. High Value: Indicates dominance of dark texture runs. Low Value: Indicates fine, complex texture.
Shape	Wider-than-Tall	Low	Shape: Elongation (Aspect Ratio)	Low Value (≤ 1): Transverse axis ≥ AP dimension.
	Taller-than-Wide	High	Shape: Elongation (Aspect Ratio)	High Value (> 1): AP axis > Transverse axis
Margin	Smooth	Low	Gradient: Gradient Magnitude. Gradient: Edge Sharpness	High Value: Indicates steep intensity change at border. High Value: Indicates a distinct interface.
	Irregular / Ill-defined	High	Gradient: Gradient Magnitude. GLCM: Contrast, Dissimilarity	Low Value: Indicates gradual/blurred transition. High Value: Indicates boundary heterogeneity.
Echogenic Foci	Punctate Echogenic Foci	High	First Order: Kurtosis. Filter: Laplacian of Gaussian (LoG)	High Value (> 3.0): Indicates "peaked" histogram (outliers). High Response: Detects small bright blobs.
	Macrocalcifications	Variable	GLDM: Large Dependence High Gray-Level Emphasis (LDHGLE)	High Value: Indicates large, bright contiguous zones.
	Large Comet-Tail Artifacts	Benign	GLDM: Small Dependence Low Gray-Level Emphasis (SDLGLE)	High Value (> 0.7): Captures small dark dependencies near bright seeds.

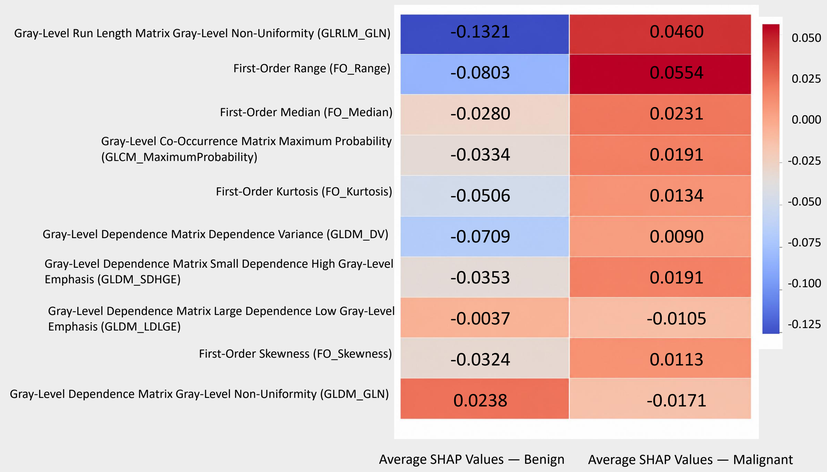

Chart 3. SHAP analysis showed consistent feature importance across top models, with texture heterogeneity as the strongest marker of malignancy.

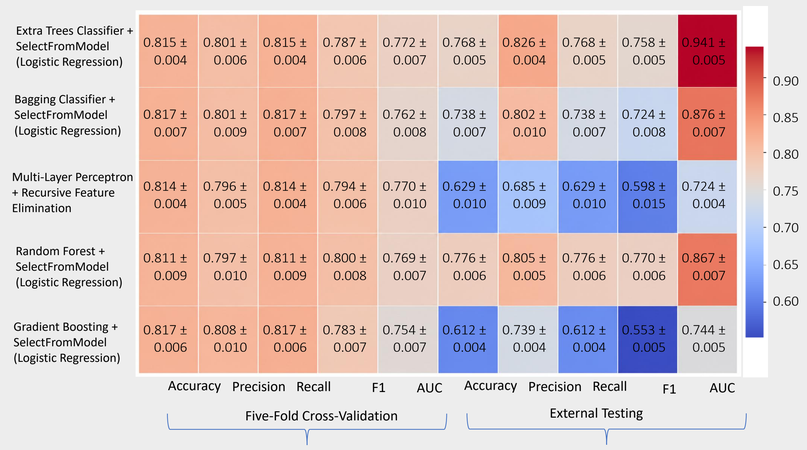

Chart 2. Heatmap illustrating the performance of the top five ranked models across validation and test metrics. Each cell reports the mean value ± standard deviation for Accuracy, Precision, Recall, F1-Score, and AUC under both validation and test conditions. Rows correspond to classifier–feature-selection combinations, and columns represent the evaluated metrics. Red shades indicate higher values, while blue shades indicate lower values.

CONCLUSIONS

This study introduces a **clinically interpretable radiomics dictionary and stability-aware model selection framework** that overcomes black-box limitations, enabling transparent and trustworthy thyroid nodule-risk stratification from ultrasound.

REFERENCES

1. G. Grani, M. Sponziello, S. Filetti and C. Durante, "Thyroid nodules: diagnosis and management," Nature Reviews Endocrinology, vol. 20, no. 12, pp. 715-728, 2024.
2. M. R. Salmanpour, S. Amiri, S. Gharib, A. Shariftabrizi, Y. Xu, W. Weeks, A. Rahmim and I. Hacihaliloglu, "Biological and Radiological Dictionary of Radiomics Features: Addressing Understandable AI Issues in Personalized Prostate Cancer; Dictionary version PM1. 0," Journal of imaging informatics in medicine, pp. 1-22, 2025.

CONTACT

Mohammad Salmanpour, Ph.D.
Department of Basic and Translational Research, BC Cancer Research Institute, Vancouver, BC, Canada
msalman@bccrc.ca

© Poster Template by Genigraphics® 1.800.790.4001 www.genigraphics.com