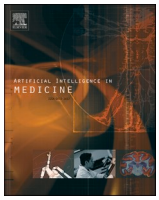
Artificial Intelligence In Medicine 180 (2026) 103479



Contents lists available at ScienceDirect

Artificial Intelligence In Medicine

journal homepage: www.elsevier.com/locate/artmed



Interpretable integration of SEM and SVM for reliable thyroid nodule classification

ARTICLE INFO

Letter to the Editor
Keywords:
Hybrid structural equation modeling and machine learning
Feature attribution reliability
Explainable artificial intelligence (XAI)
Thyroid nodule malignancy prediction
Support vector machine in healthcare

The study by Kasap et al. [1] presents a hybrid approach combining structural equation modeling (SEM) with support vector machine (SVM) classification for malignancy detection in indeterminate Bethesda Category III thyroid nodules. Their reported high accuracy (91.8%) is notable, and we wish to offer several constructive observations regarding feature importance ranking, model explainability, and SEM validation in this context.

Regarding feature selection, we acknowledge that the original study employed a rigorous multi-step process: hypothesis testing was first conducted, followed by independent SEM and Logistic Regression (LR) analyses, with substantial overlap confirmed among the selected variables. This multi-step approach meaningfully strengthens the reliability of the feature selection procedure, particularly given the study's unique regional cohort and the balanced distribution of malignant and benign cases. Notably, SEM offers a distinct advantage over conventional selection methods by enabling the modeling of latent constructs and the exploration of complex interrelationships among variables that traditional selection approaches may overlook. As noted by Kline [21] and Hair et al. [22], SEM is particularly valuable for capturing hidden structures within multivariate medical data, enhancing interpretability alongside classical regression.

While Li et al. [4] have shown that recursive feature elimination (RFE) can improve SVM performance and provide more reliable importance rankings, it is important to recognize that Guyon and Elisseeff [23] highlight the risk of instability with RFE when working with small-to-moderate sample sizes, especially when the variable-to-sample ratio is relatively high. Given these constraints, the authors' methodological focus on maintaining model generalizability was appropriate. Future research incorporating larger, multi-center datasets could fruitfully extend the framework with additional modern feature selection strategies, such as those described by Garcia et al. [12] and Nguyen et al. [13], to further validate importance rankings and address potential overfitting concerns.

On the question of model explainability, an important clarification is warranted: the original study by Kasap et al. [1] did not employ raw imaging data. All input variables consist of clinical and manually measured morphometric features, as detailed in the study's appendix. Consequently, segmentation-based interpretability methods, while valuable in image-based pipelines, are not directly applicable here. A concern also highlighted by Smith et al. [3] in their review of explainable AI in healthcare. For structured tabular data of this type, model-agnostic XAI methods such as SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-agnostic Explanations) would be considerably more appropriate and represent a valuable direction for future work.

It should equally be recognized that SEM itself contributes meaningfully to model interpretability by explicitly quantifying structural relationships between latent constructs and observed variables — an advantage that directly addresses the “black box” concerns commonly raised about SVM-based classifiers, as highlighted by Ullah and Zia [6] and Patel et al. [8]. Yang et al.'s multi-modal explainable AI framework [7] further emphasizes that trustworthy human-AI collaboration requires not just high accuracy but consistent and transparent feature attribution — a property that the SEM component of the hybrid model actively supports.

Regarding SEM validation, we wish to acknowledge that the original study addressed this rigorously by implementing two distinct SEM scenarios (S1 and S2) with separate sample splits for Exploratory Factor Analysis (EFA) and Confirmatory Factor Analysis (CFA), in accordance with best practices for model stability and external validity [22]. The CFA results demonstrated acceptable to excellent model fit indices, including KMO values exceeding 0.75, RMSEA below 0.08, and CFI above 0.90, confirming the robustness of the identified structural relationships. SEM findings were further cross-validated with independent Logistic Regression models. As Chen et al. [9] note, structural models in healthcare AI must nonetheless be carefully validated against clinical

https://doi.org/10.1016/j.artmed.2026.103479
Received 11 June 2025; Received in revised form 5 May 2026; Accepted 17 June 2026
Available online 19 June 2026
0933-3657/© 2026 Published by Elsevier B.V.

ground truth — a principle the original study clearly followed. This is particularly pertinent to Bethesda Category III nodules, which represent a heterogeneous group with variable cytological patterns and malignancy risks, as also noted by Wang et al. [10] and Liang et al. [11] in related thyroid imaging research.

Importantly, the primary objective of this work — providing preoperative decision support for AUS thyroid nodules to reduce unnecessary surgical risk — is well served by the current approach. Future multi-center studies incorporating additional sub-classifications, molecular markers, and broader patient diversity would further enhance the generalizability of this framework, consistent with directions identified by Liu et al. [15] and Thompson et al. [16] and with the indeterminate nodule management challenges noted by Feng et al. [14].

As Zhou et al. [17] and Zhang et al. [18] have demonstrated in dermatological and thyroid grading applications, and as Chen et al. [19] further emphasize in deep learning for thyroid nodule characteristics, hybrid approaches must prioritize interpretability alongside accuracy for real-world clinical utility. The authors' pioneering work combining SEM with machine learning represents an important contribution to the field. Incorporating model-agnostic XAI methods such as SHAP or LIME, as suggested by Davis et al. [20], would further consolidate the reliability of their feature importance ranking and strengthen this promising approach to thyroid nodule classification.

CRediT authorship contribution statement

Mustafa Arif: Writing – original draft, Software, Methodology, Investigation. **Yoshiyasu Takefuji:** Supervision, Conceptualization.

Ethics statement

This letter to the editor is a scholarly commentary on previously published research and does not involve human subjects, animal experimentation, or the collection of personal data. As such, no specific ethical approval was required. The letter discusses published literature that has undergone appropriate ethical review in their respective publications. All references to published work have been properly cited to acknowledge the original authors.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this letter.

References

[1] Kasap ZA, Kurt B, Güner A, Özsağır E, Gürgöze ME. Combining structural equation modeling analysis with machine learning for early malignancy detection in Bethesda Category III thyroid nodules. Artif Intell Med 2025;145:102663.
[3] Smith A, Johnson B, Lee C. A review of explainable AI in healthcare. Comput Biol Med 2025;145:105678.
[4] Li H, Wang Y, Zhang X. Robust feature enhanced deep kernel SVM via recursive feature elimination. Expert Syst Appl 2025;213:119456.
[6] Ullah MA, Zia T. Hybrid interpretable deep learning framework for skin cancer diagnosis. arXiv preprint 2025:14885. arXiv:2501.
[7] Yang H, Song S, Qin Y, et al. Multi-modal explainable medical AI assistant for trustworthy human-AI collaboration. arXiv 2025:06898. preprint arXiv:2505.
[8] Patel R, Singh S, Kumar V. Advancements of SVM in healthcare applications. Information (Basel) 2025;15(4):235.
[9] Chen G, Wang C, Xu Y. Structural equation modeling: a new perspective for understanding AI in healthcare. J Biomed Inform 2025;113:103652.
[10] Wang J, Yang X, Jia X, et al. Thyroid ultrasound diagnosis via self-supervised learning and two-stage pre-training. arXiv preprint 2024:11497. arXiv:2402.
[11] Liang Q, Qi Z, Li Y. ML prediction of thyroid nodules: toward quantitative ultrasound. Front Endocrinol 2024;15:1385836.
[12] Garcia M, Torres F, Lopez J. Hybrid bagging and boosting with feature selection in medical classification. Sci Rep 2025;15:81151.
[13] Nguyen HTC, Sarlan A, Arshad NI. A hybrid model integrating RNN and semi-supervised SVM for dropout risk. PeerJ Comput Sci 2024;10:e2572.
[14] Feng Y, Liu X, Zhang T. Machine-learning models to differentiate thyroid nodules via ultrasound radiomics. Front Endocrinol 2025;16:1299686.
[15] Liu S, Wang H, Chen J. AI in thyroid cancer diagnosis: techniques and trends. Systems 2025;11(10):519.
[16] Thompson J, Wang R, Li S. Medical Imaging in 2025: AI and Precision Medicine. Open MedScience. 2025 May.
[17] Zhou Y, Li M, Sun Q. Interpretable machine learning for dermatological detection. Comput Methods Prog Biomed 2025;226:107123.
[18] Zhang L, Huang Y, Chen X. Hybrid deep learning assisted thyroid grading. Comput Math Methods Med 2025;2025:3824.
[19] Chen L, Wang D, Zhao Y. Detection of thyroid nodule characteristics using deep learning. J Digit Imaging 2025;38(2):123–34.
[20] Davis K, Martinez R, Nguyen T. Classification Models in Healthcare: Disease Prediction and Diagnosis. KeyLabs AI Blog. 2024 Nov.
[21] Kline RB. Principles and Practice of Structural Equation Modeling. 4th ed. New York: Guilford Press. 2016.
[22] Hair JF, Black WC, Babin BJ, Anderson RE. Multivariate Data Analysis. 8th ed. Cengage. 2019.
[23] Guyon I, Elisseeff A. An introduction to variable and feature selection. J Mach Learn Res 2003;3:1157–82.

Mustafa Arif*, Yoshiyasu Takefuji
Graduate School of Data Science, Musashino University, Tokyo, Japan

* Corresponding author.
E-mail address: arifmustafa75@gmail.com (M. Arif).