



Letter to the Editor

Methodological Concerns in Radiomics: Addressing Bias in LASSO and SHAP for Thyroid Tumor Analysis



To the Editor

Ying Fu et al. conducted a study that focused on intra- and peritumoral radiomics derived from ultrasound images to aid in the preoperative differentiation of follicular thyroid adenoma, follicular carcinoma, and follicular tumors of uncertain malignant potential [1]. Their analysis employed the least absolute shrinkage and selection operator (LASSO) regression to identify the most relevant features, utilizing five-fold cross-validation to ensure the robustness of the training set. Additionally, to enhance the interpretability of the model's decisions, the researchers analyzed the impact of the identified features on the model output using the SHapley Additive exPlanations (SHAP).

LASSO regression raises methodological concerns that could affect the validity of Fu et al.'s findings. As a linear and parametric method, LASSO eliminates significant nonlinear features, introducing critical biases [2–8]. It selects only one variable from highly correlated groups, potentially missing key predictors. Its tendency to shrink coefficients to zero may oversimplify models by excluding relevant variables. Additionally, LASSO's sensitivity to the regularization parameter means that improper tuning can heavily influence performance, resulting in configuration-specific biases. Over 300 peer-reviewed articles have documented substantial biases in feature importance rankings derived from machine learning models, including LASSO.

SHAP also inherits and potentially amplifies biases in feature importances from the models it explains, distorting interpretations and conclusions [9–13]. The function structure “explain = SHAP(model)” reveals SHAP's reliance on the model. Since SHAP relies on model outputs to determine feature importance, it inevitably inherits the model's biases. This can lead to misleading interpretations and undermine the reliability of the findings. Machine learning models often prioritize prediction accuracy, which can result in overfitting, meaning that high prediction accuracy does not guarantee reliable feature importance rankings.

The absence of ground truth values for feature importance leads to varied methodologies among models, necessitating careful interpretation of results. To reveal genuine relationships between target variables and features, researchers should prioritize three key areas: understanding data distribution for suitable modeling techniques, exploring statistical relationships, especially nonlinear ones, and validating findings through hypothesis testing and p-values. A multi-method approach is encouraged, integrating complementary techniques to capture complex feature interactions effectively.

For inherently nonlinear, nonparametric datasets, employing Spearman's rank correlation, feature agglomeration (FA), and high-variance gene selection (HVGS) is recommended. Spearman's correlation effectively identifies both linear and monotonic relationships while minimizing the influence of outliers [14,15]. FA reduces dimensionality and noise by clustering similar features, enhancing interpretability [16,17].

Meanwhile, HVGS highlights highly variable features that offer biological insights [18]. Together, these techniques provide a robust framework for understanding variable interactions and improving the reliability of feature importance assessments.

In conclusion, incorporating the proposed robust analytical methods is crucial for significantly improving the reliability of Fu et al.'s findings. By actively addressing the biases inherent in LASSO and SHAP, and prioritizing genuine associations, they can not only enhance the integrity of their work but also unveil more profound and unbiased insights into the underlying relationships among features in thyroid tumor radiomics.

Conflict of interest

The authors declare no competing interests.

CRediT authorship contribution statement

Naoki Iwata: Writing – original draft, Investigation. **Souichi Oka:** Writing – review & editing, Conceptualization. **Yoshiyasu Takefuji:** Supervision, Project administration.

Acknowledgments

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Data availability statement

No new data were generated or analyzed in support of this research.

References

- [1] Fu Y, Mei F, Shi L, Ma Y, Liang H, Huang L, et al. Intra- and peritumoral radiomics based on ultrasound images for preoperative differentiation of follicular thyroid adenoma, carcinoma, and follicular tumor with uncertain malignant potential. *Ultrasound Med Biol* 2025. doi: [10.1016/j.ultrasmedbio.2025.04.005](https://doi.org/10.1016/j.ultrasmedbio.2025.04.005).
- [2] Wüthrich K, Zhu Y. Omitted variable bias of lasso-based inference methods: a finite sample analysis. *Rev Econ Stat* 2023;105(4):982–97. doi: [10.1162/rest_a.01128](https://doi.org/10.1162/rest_a.01128).
- [3] Freo M, Luati A. Lasso-based variable selection methods in text regression: the case of short texts. *ASTA Adv Stat Anal* 2024;108:69–99. doi: [10.1007/s10182-023-00472-0](https://doi.org/10.1007/s10182-023-00472-0).
- [4] Basu T, Einbeck J, Troffaes MCM. Uncertainty quantification in lasso-type regularization problems. In: Vasile M, editor. *Optimization under uncertainty with applications to aerospace engineering*. Cham: Springer; 2021. p. 11–20. doi: [10.1007/978-3-030-60166-9_3](https://doi.org/10.1007/978-3-030-60166-9_3).
- [5] Fridgerisson EA, Williams R, Rijnbeek P, Suchard MA, Reys JM. Comparing penalization methods for linear models on large observational health data. *J Am Med Inform Assoc* 2024;31(7):1514–21. doi: [10.1093/jamia/ocae109](https://doi.org/10.1093/jamia/ocae109).
- [6] Hernández-Lemus E, Ochoa S. Methods for multi-omic data integration in cancer research. *Front Genet* 2024;15:1425456. Published 2024 Sep 19. doi: [10.3389/fgene.2024.1425456](https://doi.org/10.3389/fgene.2024.1425456).
- [7] Li X, Jacobucci R. Regularized structural equation modeling with stability selection. *Psychol Methods* 2022;27(4):497–518. doi: [10.1037/met0000389](https://doi.org/10.1037/met0000389).
- [8] Jain R, Xu W. HDSI: High dimensional selection with interactions algorithm on feature selection and testing. *PLoS One* 2021;16(2):e0246159. Published 2021 Feb 16. doi: [10.1371/journal.pone.0246159](https://doi.org/10.1371/journal.pone.0246159).

<https://doi.org/10.1016/j.ultrasmedbio.2025.08.016>

Received 13 July 2025; Accepted 12 August 2025

- [9] Huang X, Marques-Silva J. On the failings of Shapley values for explainability. *Int J Approx Reason* 2024;171:109112.
- [10] Bilodeau B, Jaques N, Koh PW, Kim B. Impossibility theorems for feature attribution. *Proc Natl Acad Sci U S A* 2024;121(2):e2304406120.
- [11] Molnar C, König G, Herbringer J, Freiesleben T, Dandl S, Scholbeck CA, Casalicchio G, Grosse-Wentrup M, Bischl B. General pitfalls of model-agnostic interpretation methods for machine learning models editors.. In: Holzinger A, Goebel R, Fong R, Moon T, Müller KR, Samek W, editors. *xxAI – Beyond explainable AI. Lecture notes in computer science*, vol 13200. Cham: Springer; 2022. p. 39–68. doi: [10.1007/978-3-031-04083-2_4](https://doi.org/10.1007/978-3-031-04083-2_4).
- [12] Wilming R, Budding C, Müller K-R, Haufe S. Scrutinizing XAI using linear ground-truth data with suppressor variables. *Mach Learn* 2022;111(5):1903–23.
- [13] Watson DS. Conceptual challenges for interpretable machine learning. *Synthese* 2022;200(2):65.
- [14] Yu H, Hutson AD. A robust Spearman correlation coefficient permutation test. *Commun Stat Theory Methods* 2024;53(6):2141–53. doi: [10.1080/03610926.2022.2121144](https://doi.org/10.1080/03610926.2022.2121144).
- [15] Eden SK, Li C, Shepherd BE. Nonparametric estimation of Spearman's rank correlation with bivariate survival data. *Biometrics* 2022;78(2):421–34. doi: [10.1111/biom.13453](https://doi.org/10.1111/biom.13453).
- [16] Zhang J, Wu X, Hoi SCH, Zhu J. Feature agglomeration networks for single stage face detection. *Neurocomputing* 2020;380:180–9. doi: [10.1016/j.neucom.2019.10.087](https://doi.org/10.1016/j.neucom.2019.10.087).
- [17] Sreeja MU, Kooroor BC. Enhanced block-based feature agglomeration clustering for video summarization editors. In: Srinivas M, Sucharitha G, Matta A, Chatterjee P, editors. *Machine learning algorithms and applications*. John Wiley & Sons, Inc.; 2021. doi: [10.1002/9781119769262.ch7](https://doi.org/10.1002/9781119769262.ch7).
- [18] Xie Y, Jing Z, Pan H, Xu X, Fang Q. Redefining the high variable genes by optimized LOESS regression with positive ratio. *BMC Bioinformatics* 2025;26:104. doi: [10.1186/s12859-025-06112-5](https://doi.org/10.1186/s12859-025-06112-5).

Naoki Iwata^{a,*}, Souichi Oka^b, Yoshiyasu Takefuji^a

^a Faculty of Data Science, Musashino University, Tokyo, Japan

^b Science Park Corporation, Kanagawa, Japan

*Corresponding author. Faculty of Data Science, Musashino University,
3-3-3 Ariake Koto-ku, Tokyo 135-8181, Japan
E-mail address: s2322008@stu.musashino-u.ac.jp (N. Iwata).