Towards Reliable Lung Cancer Prediction: A Hybrid Framework for Noise Reduction and Uncertainty Control

Sourojit Pal[1], Sandip Roy[2], Pratip Rana[3], Avishek Banerjee[4], Koushik Majumder[1], Sachin Shetty[2]

[1]Department of Computer Science & Engineering, MAKAUT, West Bengal 741249, India
[2]Center for Secure & Intelligent Critical Systems, Old Dominion University, Virginia, USA
[3]Department of Computer Science, Old Dominion University, Virginia, USA
[4]Department of IT, Asansol Engineering College, West Bengal 713305, India
psouro1990@gmail.com, {sroy, prana, sshetty}@odu.edu, avishekbanerji@gmail.com, koushik-zone@yahoo.com

Abstract

Uncertainty remains a critical challenge in healthcare AI, since predictive errors can directly compromise patient safety and undermine trust. Structured clinical datasets in healthcare are frequently characterized by heterogeneous acquisition protocols, incomplete records, and inconsistent or noisy encodings. This inflates aleatoric uncertainty and weakens calibration. These challenges are exemplified in lung cancer risk modeling, where small cohorts, variable collection practices, and limited feature quality make the problem especially acute. Significant advances in uncertainty quantification (UQ) have been achieved in imaging and signal processing through Bayesian inference, evidential learning, and robust architectural designs. In contrast, tabular clinical datasets remain a critical yet underexplored domain. Addressing this gap requires methods that are lightweight, certifiable, and effective on noisy datasets without relying on large models or data. Considering this challenges, we propose a frequency-aware hybrid representation that combines Principal Component Analysis (PCA) with the Discrete Cosine Transform (DCT). Using mutual information (MI)–based feature ordering, the framework suppresses high-frequency artifacts while preserving discriminative structure. As the framework was applied to a publicly available lung cancer dataset, it demonstrated an accuracy improvement from 98.1% to 99.7%, reduced Negative Log-Likelihood (NLL) by $\approx$ 82% from 5.25% to 0.94%, lowered aleatoric uncertainty from 10.50% to 3.35% ($\approx 68\%$ reduction), and preserved AUROC at 99%. We evaluated the framework across three publicly available lung-cancer datasets where it demonstrated a reduction in aleatoric uncertainty by 7% on an average, confirming generalizability. The Wilcoxon signed-rank test confirms that the results are statistically significant. This work shows that part of the 'irreducible' variability is actually compressible noise, thereby facilitating more reliable and uncertainty-aware AI for healthcare.

Introduction

Lung cancer prediction in clinical settings frequently relies on *tabular* data (questionnaires, vitals, labs, comorbidities) (Liao et al. 2023). Such data are impacted by noise due to heterogeneous collection protocols, missing values, categorical recoding inconsistencies, small cohort sizes, and site shifts. These imperfections elevate *aleatoric uncertainty*, a type of uncertainty due to inherent randomness the in data, and weaken calibration and downstream decision reliability (Kendall and Gal 2017). Clinical tabular methodologies remained underexplored in uncertainty-aware preprocessing despite their ubiquity in screening and risk stratification.

While aleatoric uncertainty is often described as irreducible, its boundary can blur when models are misspecified or latent variables drive apparent randomness (Gawlikowski et al. 2023). Advances in measurement, representation, or modeling can make portions of seemingly "irreducible" variability tractable (Heid et al. 2023; Li 2025). This motivates frequency-aware preprocessing for noisy clinical tables.

Motivation: Empirically, appropriate representations can suppress measurement jitter and stabilize estimates. In healthcare, datasets are prone to incomplete entries, inconsistent encoding, and measurement variability. All these factors elevate aleatoric uncertainty and undermine the reliability of predictive models. We seek a simple, tabular-first denoising step that improves reliability without sacrificing global structure. Furthermore, clinical screening often operates in resource-constrained environments, where lightweight preprocessing is preferable to complex deep architectures. Existing approaches that focus on imaging or signal domains do not directly address tabular noise, leaving a gap in uncertainty-aware preprocessing for structured records. Bridging this gap is essential for improving trust in predictive systems and supporting safer deployment in routine clinical workflows.

Contributions: To address the above challenges, we translate the motivation into the following contributions.

1. We propose a lightweight, tabular-first PCA–DCT hybrid framework for lung-cancer prediction, designed to denoise clinical features while preserving key statistical structure.
2. We provide a unified uncertainty evaluation across multiple lung-cancer datasets, reporting AUROC, NLL, calibrated metrics, risk–coverage, and paired Wilcoxon tests on per-sample aleatoric change.
3. We highlight that part of the variability often deemed "irreducible" is compressible measurement noise, underscoring the value of frequency-aware preprocessing in clinical tabular pipelines.

4. We demonstrate consistent performance gains, with the hybrid reducing mean aleatoric uncertainty by approximately 7%, lowering NLL by 3.1% approximately, and preserving AUROC relative to PCA-only and DCT-only baselines.

The rest of this article is structured as follows: related-work provides a brief literature survey on various aleatoric uncertainty quantification approaches. In proposedsoln, we describe the "PCA-DCT hybrid" framework. In experiments, we present extensive results to validate the merits of the proposed framework. Finally, we conclude the paper in conclusion.

Related Works

Uncertainty in machine learning is commonly partitioned into *aleatoric* (data-inherent), *epistemic* (model/knowledge), and *distributional* (shift) components (Kendall and Gal 2017; Malinin and Gales 2018), (Pal et al. 2023). While the textbook view treats aleatoric effects as irreducible, recent analyses emphasize that the boundary is representation- and measurement-dependent: misspecification, latent confounders, and aggregation choices can make seemingly "irreducible" variability partially tractable (Gawlikowski et al. 2023; Heid et al. 2023; Li 2025). Evidence from metrology and model building further shows that information structure and measurement discipline critically shape observed variability and replicability (Wang et al. 2025).

Methodologically, aleatoric uncertainty has been tackled by heteroscedastic likelihoods, Bayesian surrogates (Kendall and Gal 2017), evidential distributions and other robust techniques that can separate data from model uncertainty (Sensoy, Kaplan, and Kandemir 2018). Data-centric denoising strategies such as Noise2Noise (Lehtinen et al. 2018) and test-time augmentation (Wang et al. 2019) reduce variance without requiring clean labels. From the evaluation evaluation standpoint, calibration metrics like ECE and performance under dataset shift remain crucial. (Ovadia et al. 2019).

Prior studies emphasize that clinical noise, missingness, and heterogeneous acquisition conditions significantly inflate aleatoric effects, thereby limiting confidence in model outputs and clinical decision-making (Alizadehsani et al. 2024; Tsaneva-Atanasova, Pederzanil, and Laviola 2025). Recent works explores multi-modality to reduce uncertainty (Hoarau et al. 2025) and tackles noise through both label-level and signal-level approaches. In the signal domain, PCA variants have been effective denoisers but can struggle under spatially correlated noise (Henriques et al. 2023). Frequency-domain methods are also gaining traction: Fourier-basis augmentation targets robustness gaps left by standard visual augmentations (Vaish, Wang, and Strisciuglio 2024); in X-ray imaging, Poisson PCA combined with nonlocal means reduces photon-limited noise (Kipele and Greyson 2023).

Synthesis : Despite this progress, most uncertainty and denoising advances focus on *images* and continuous signals. In contrast, *tabular* clinical methodologies that are critical for screening and risk stratification, receive comparatively less attention. Label-noise methods and calibration metrics help, but they do not directly address column-wise high-frequency artifacts. PCA captures global covariance yet tends to retain these localized irregularities; DCT suppresses high-frequency noise but can remove discriminative structure if applied naively to mixed clinical features. This gap motivates a frequency-aware, *tabular-first* representation that aligns informative variation, filters jitter with low-frequency DCT coefficients, and preserves global structure.

Proposed PCA-DCT Hybrid Framework

This section discusses about the background and the framework of the research undertaken.

The Approach

Our approach aims for reliable lung cancer prediction from tabular clinical data using an easy-to-use plug-and-play hybrid embedding. The framework unifies *mutual-information (MI)–guided feature ordering*, *low-frequency discrete cosine transform (DCT)* filtering, and *principal component analysis (PCA)* into a single hybrid embedding (Figure 1).

Working principle

We pre-process the data by standardizing numeric fields, one-hot encoding categorical variables, and imputing missing values. After preprocessing, we compute MI between each feature and the label. The features are then sorted by MI to concentrate informative variation at lower "frequencies" along the feature axis. On the MI-ordered table we perform two computations per sample: (i) PCA scores that capture *global covariance* (Jolliffe and Cadima 2016), and (ii) a 1D-DCT across the feature axis (Ahmed, Natarajan, and Rao 1974). For DCT, we retain only the first k low-frequency coefficients ($k \in \{2, 4, 6, 8, 12\}$) and suppress column-wise high-frequency artifacts like sparse one-hot toggles, discretization edges, sparse binary indicators, threshold cutoffs, and others. Concatenating PCA scores with the retained DCT coefficients yields a compact, denoised vector that preserves discrimination while attenuating noise. An ensemble classifier is then trained on this hybrid representation and results are computed.

Performance and efficacy

Handling Information Leakage : Information leakage occurs when test data unintentionally influences model training, producing overly optimistic results. PCA-based hybrids are particularly vulnerable to information leakage. If mutual information ordering, PCA variance components, or DCT coefficients are estimated using test data, the results may appear more reliable than they truly are. Such leakage undermines the validity of uncertainty estimates, creating a false sense of improved calibration that can lead to serious failures in real-world deployment.

To prevent information leakage, MI ordering, PCA, and DCT parameters are estimated strictly on the training portion within each split, and the validation is performed on held-out data.

Hyperparameter definitions : We keep hyperparameters intentionally modest (refer table 1). This consists of narrow PCA explained-variance target and a small discrete set for k in the DCT block. The base model is set to "gradient boosting" and train to test ratio of $75:25$.

Performance calibration : Reliability is quantified using AUROC for discrimination; negative log-likelihood (NLL) and expected calibration error (ECE) for calibration; and risk–coverage for selective prediction under a reject option. We further report predicted/expected entropy and mutual information to decompose predictive uncertainty. We also compute a simple shift score via Mahalanobis distance in the PCA subspace,

$$d_M(x) \;=\; \sqrt{(z-\mu)^\top \Sigma^{-1} (z-\mu)}\,,$$

where z denotes PCA coordinates and (μ, Σ) are the training-set mean and covariance. Statistical evidence is provided using paired Wilcoxon signed-rank tests on per-sample changes in aleatoric uncertainty (Post–Pre), with effect sizes. The effect sizes shows the magnitude of reduction in uncertainty. For practical robustness, we include *noise-injection* studies that add controlled feature or label noise to emulate realistic clinical perturbations (miscoding, missingness patterns), informed by prior label-noise and augmentation literature.

Uncertainty calibration : The proposed framework doesn't just reduce noise, it also measures how much uncertainty changes after PCA–DCT hybridization. The measures are as discussed below:

1. Calibration metrics : PCA-DCT framework computes Negative Log-Likelihood (NLL) and Expected Calibration Error (ECE). These measures measure whether predicted probabilities match true outcomes, that is, how "honest" the uncertainty estimates are.
2. Risk–coverage curves : By allowing a reject option, the framework plots risk versus prediction coverage. This shows how uncertainty can guide safer selective predictions (low risk when the model is confident).
3. Entropy and mutual information : Predictive entropy and MI are reported per sample, breaking down the uncertainty into aleatoric (data-driven randomness) and epistemic components.
4. Wilcoxon signed-rank tests : The framework directly test per-sample changes in aleatoric uncertainty (Post–Pre) between PCA-DCT hybrid and baseline, thereby quantifying whether reductions are statistically significant (with effect sizes).
5. Distributional analysis : The framework compares CDFs of aleatoric uncertainty and scatterplots of uncertainty vs. confidence to show how the hybrid reshapes the distribution of uncertainty across patients.

Benchmarking the framework : PCA-DCT hybrid framework was initially trained and tested on publicly available lung cancer dataset called virtual10. This is available in website: hugging face. Apart from this, two more publicly available datasets on lung cancer datasets were used (refer to table 4). Number of cases in the datasets ranged between $(309 - 4500)$ with 16 features each. When applied to these datasets, the proposed procedure *consistently reduced aleatoric uncertainty* by approximately 7%, *lowered NLL* by **3.1%**, and *preserves AUROC* consistently, relative to standardized PCA and DCT baselines.

These results suggest that, once features are MI-aligned and high-frequency artifacts are filtered, a clinically meaningful portion of variability behaves as "*compressible measurement noise*". The approach is lightweight, requires few design choices, and integrates naturally as a **frequency-aware, tabular-first** preprocessing block amongst lung-cancer risk modeling methodologies—appropriate to motivate broader, multi-center validation. The framework is computationally inexpensive, easy to configure, data efficient and broadly applicable, making it practical for clinical research settings with limited data and resources. Refer algorithm 1.

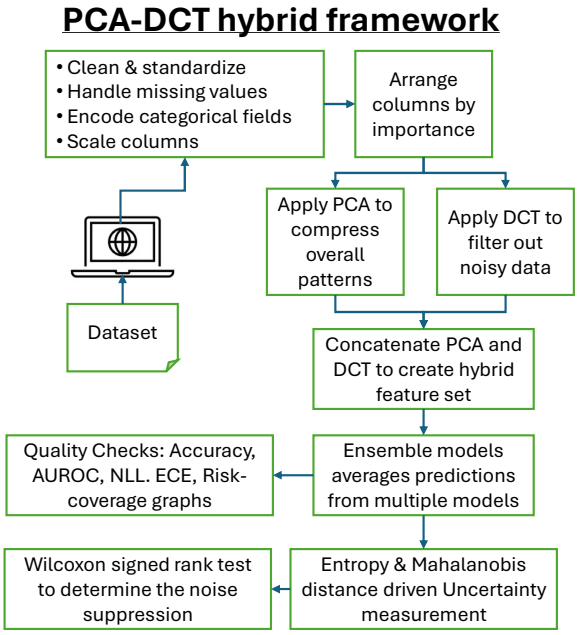

Figure 1: Frequency-aware preprocessing for tabular lung-cancer prediction: MI ordering + low-frequency DCT, PCA embeddings, and unified uncertainty evaluation.

Performance Evaluation

In this section we undertake the experiments to visuallize the impact of PCA-DCT hybrid framework on the noise in clinical tabulated dataset.

Evaluation of PCA-DCT hybrid framework

In this section, we review the functioning of PCA-DCT hybrid framework over dataset 1 and evaluate its performance in terms of accuracy, aleatoric uncertainty, reliability and trust.

Pre-Post discrimination and calibration: This experiment evaluates the impact of PCA-DCT hybridization on

Symbol	Description	Value
k	Default k for DCT in ablation	4
PCA_VAR	Default explained variance for PCA	99%
BASE_MODEL	Model used for train/test	gb
ORDER_MODE	Mode for method execution	mi
CALIBRATE	Learn isotonic mapping on calibration split	FALSE
PCA_SWEEP	Operating range for PCA	0.90–0.99
K_SWEEP	Operating set for DCT	{2, 4, 6, 8, 12}
TEST_SIZE	Size of the test population	25%

Table 1: Brief description of the global parameters used in the PCA-DCT framework.

Dataset ID	Description	# cases	# features
Dataset 1	base_hf_virtual10	4500	16
Dataset 2	test_hf_survey (nateraw)	310	16
Dataset 3	test_kaggle_survey	310	16

Table 2: Dataset summary used in the PCA–DCT framework.

Algorithm 1: PCA–DCT Hybridization

Input set : Table X with labels Y, PCA variance τ, DCT keep-set K

Output set: Classifier f with performance metrics

▷ ***preprocessing***

▷ ***Numeric scaling***

$X \leftarrow \text{Standardize}(X_{\text{num}})$

▷ ***Categorical encoding***

$X \leftarrow \text{OneHot}(X_{\text{cat}})$

▷ ***Train-only stats***

$X \leftarrow \text{Impute}(X;\, \theta_{\text{train}})$

▷ ***MI order***

$X \leftarrow \text{SortByMI}(X, Y)$

▷ ***Base representation***

$z \leftarrow \text{PCA}_{\tau}(X_{\text{train}})$

▷***Hybridization & evaluation***

for *each* $k \in K$ **do**

  $c = \text{DCT}(x)$

  $\hat{c}_k = c_{1:k}$

  $h = [\, z \parallel \hat{c}_k \,]$

  $f_k \leftarrow \text{Learn}(h_{\text{train}}, Y_{\text{train}})$

  $\{Acc, AUROC, NLL, ECE\} \leftarrow f_k(h_{\text{test}})$

return *Accuracy, AUROC, NLL, ECE*

discrimination and calibration using Dataset 1. The baseline (Pre: standardized features) is compared against the PCA-DCT hybrid (Post: PCA-DCT transformed features). Performance is measured using Accuracy, AUROC, Negative Log-Likelihood (NLL), calibrated NLL, and aleatoric uncertainty.

The results show that accuracy improves from 0.9812 to 0.9973, while AUROC remains nearly unchanged at 0.9988 and 0.9985, respectively (Table 3). A substantial improvement in calibration is observed, with NLL decreasing from 0.0525 to 0.00939 (∼82% reduction) (Fig.2). Aleatoric uncertainty is also reduced, dropping from 0.1050 to 0.0335 (∼68% reduction). In addition, the NLL distribution demonstrates a leftward shift, supporting the improvement in calibration and reduced uncertainty. The experiment demonstrates that frequency-aware preprocessing through PCA-DCT hybridization substantially improves calibration (large reduction in NLL) while preserving discrimination. The marked decrease in aleatoric uncertainty highlights effective compression of noise-driven variability, resulting in more reliable and stable predictions.

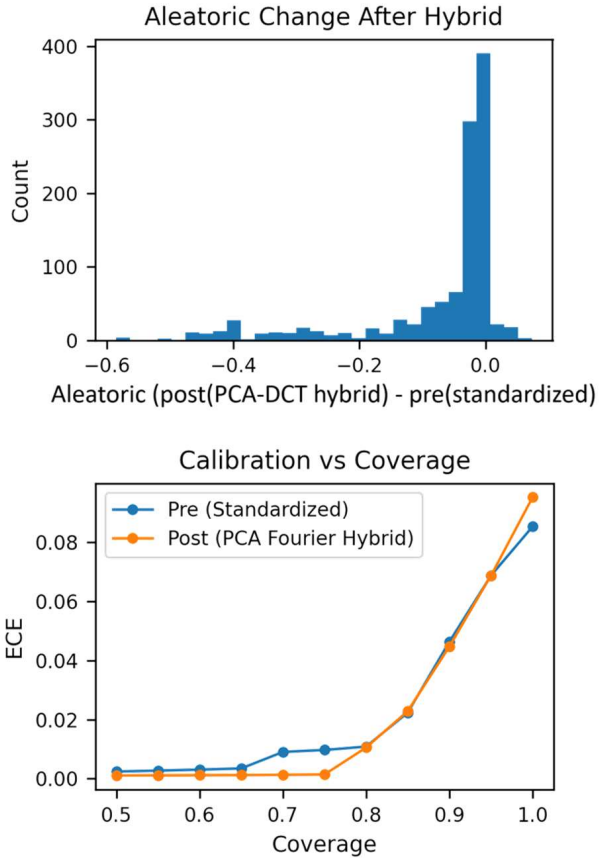

Figure 2: Aleatoric change represents NLL and Calibration-coverage represents ECE. The PCA–DCT Hybrid reduces aleatoric uncertainty compared to the standardized baseline and demonstrates improved calibration, achieving lower NLL and ECE across coverage levels.

Calibration quality & selective prediction: This experiment evaluates the behavior of expected calibration error (ECE) versus coverage, along with the selective risk curve when a reject option is introduced.

The coverage-calibration curve (Fig. 2) for the Hybrid lies consistently below the baseline across most coverage levels, indicating lower calibration error at comparable throughput.

Measure	Dataset: Dataset 1	
	Pre	Post
Accuracy	0.9812	**0.9973**
AUROC	0.9988	**0.9985**
NLL	0.0525	**0.0094**
NLL (calibrated)	0.0525	**0.0094**
Aleatoric Uncertainty	0.1050	**0.0335**

Table 3: Pre/Post metrics for Dataset 1.

In addition, the reliability diagram and risk-coverage curve (Fig. 3) move closer to the ideal, with the hybrid framework achieving lower risk for matched coverage.

The proposed framework produces more honest probabilities and enables safer selective prediction, both of which are critical for clinical decision support.

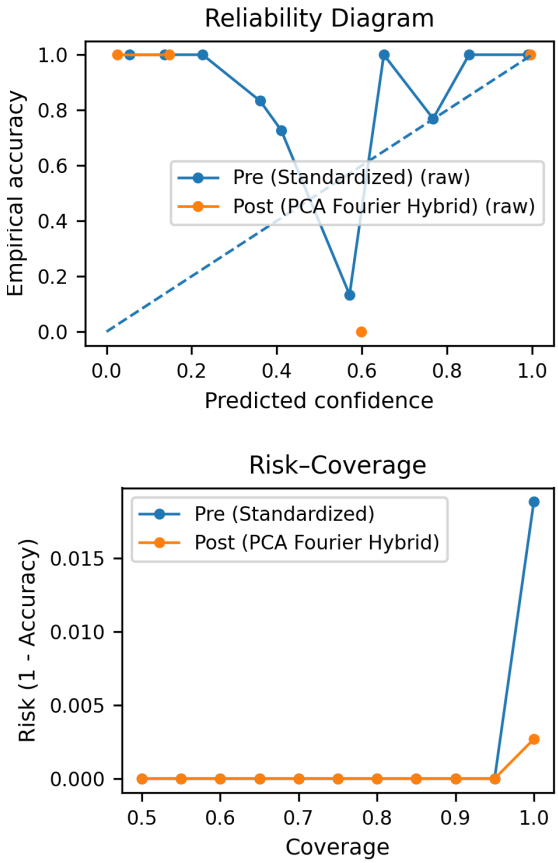

Figure 3: Reliability diagram and risk–coverage curve. The PCA–DCT Hybrid is closer to ideal calibration and achieves lower risk at the same coverage, indicating safer selective prediction.

Uncertainty behavior and ablations: In this evaluation we probe the impact of removing components (ablation sensitivity) and examine how aleatoric uncertainty relates to model confidence. We compare the Baseline (Pre) with the PCA-DCT Hybrid (Post) using: (a) mean ablation error as components are removed, (c) the cumulative distribution of aleatoric uncertainty (CDF), and (e) scatter of aleatoric versus predictive confidence.

Mean ablation error is lowest for the Hybrid across removal levels, indicating greater robustness to component drop (Fig. 4). The aleatoric CDF for the Hybrid lies left/above the baseline, showing lower uncertainty mass. Aleatoric versus confidence exhibits a stronger inverse relation for the Hybrid, with higher confidence paired with lower aleatoric, reflecting tighter alignment between uncertainty and data difficulty.

Improvements are broad rather than localized, demonstrating that frequency-aware representation reduces sensitivity to ablations and better calibrates confidence to difficulty, thereby mitigating over-confident errors.

Robustness to injected noise: We injected controlled feature and label noise into the dataset and tracked the changes in accuracy and aleatoric uncertainty. The Baseline (Pre) is compared against the PCA–DCT Hybrid (Post) under progressively increasing noise conditions.

Across increasing noise levels, the hybrid framework consistently maintained higher accuracy and lower aleatoric uncertainty compared to the baseline. This demonstrates that the framework representation is less sensitive to noise-driven perturbations. The outcome can be seen in Fig. 5 and Fig. 6.

The observed improvements are robust to clinically realistic perturbations such as miscoding and missingness, supporting the generalizability of the Hybrid approach beyond a single clean split.

Statistical note. Paired Wilcoxon signed-rank tests on per-sample aleatoric deltas (Post–Pre): $p \approx 3.1 \times 10^{-127}$, Cliff's $\delta \approx -0.616$; $\sim 81\%$ of show a negative median shift.

Distributional profile of uncertainty : This evaluation characterizes how the hybrid framework reshapes the distribution of per-sample aleatoric uncertainty and its coupling with predictive confidence.

The CDF of aleatoric uncertainty in (Fig. 7) lies above and to the left of the baseline across most quantiles, indicating a global shift toward lower uncertainty. In aleatoric vs confidence plot (Fig. 7), the aleatoric–confidence scatter shows a tighter inverse relationship, where high-confidence predictions concentrate at lower aleatoric values. The outliers are observed to shrink in magnitude.

Beyond mean or median summaries, the distribution-level evidence demonstrates a broad reduction in uncertainty and improved alignment between confidence and data difficulty. These properties are particularly important for setting triage thresholds and for transparent communication of risk in clinical settings.

Cross-dataset benchmarking

In this section we summarize the comparison of Baseline (Pre) versus Hybrid (Post) across three public lung-cancer tabular datasets: Dataset 1, Dataset 2, and Dataset 3. Performance is evaluated using Accuracy, AUROC, Negative Log-Likelihood (NLL), and aleatoric uncertainty.

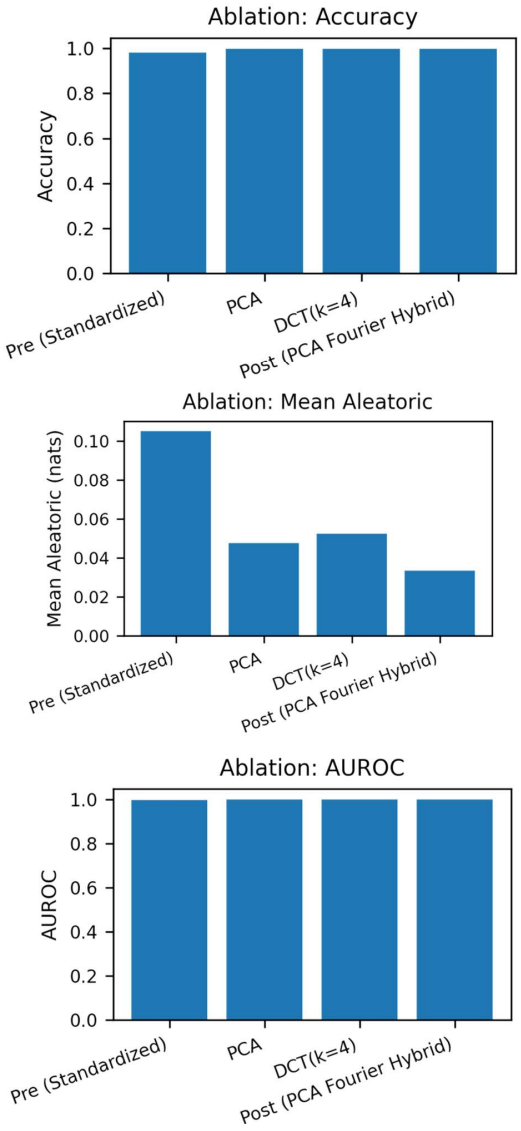

Figure 4: Ablation comparing Standard, PCA-only, DCT-only, and PCA–DCT Hybrid. The hybrid yields the lowest error/uncertainty across settings, indicating complementary benefits of PCA (global structure) and DCT (noise suppression).

For Dataset 1, accuracy improves from 0.9812 to 0.9973, while AUROC remains stable at 0.9988 and 0.9985. NLL decreases from 0.0525 to 0.00939 (approximately 82% reduction), and aleatoric uncertainty drops from 0.1050 to 0.0335 (approximately 68% reduction). For Dataset 2, accuracy rises from 0.8846 to 0.9103, AUROC improves from 0.9574 to 0.9618, NLL decreases from 0.2037 to 0.1786, and aleatoric uncertainty reduces from 0.1123 to 0.0406. For Dataset 3, the same pattern is observed as in Dataset 2, with accuracy increasing from 0.8846 to 0.9103, AUROC improving from 0.9574 to 0.9618, NLL decreasing from

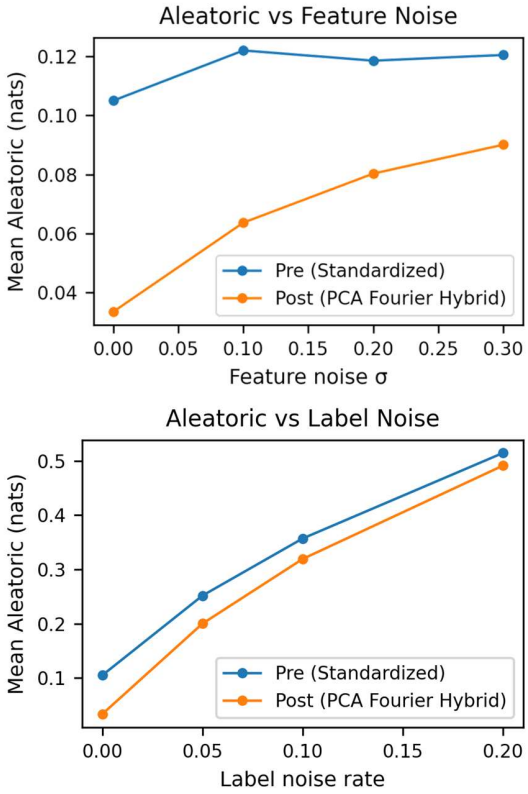

Figure 5: Robustness under injected noise. As feature/label noise increases, the hybrid maintains lower aleatoric uncertainty than the baseline, demonstrating resilience to realistic clinical perturbations.

0.2037 to 0.1786, and aleatoric uncertainty reducing from 0.1123 to 0.0406. This shows that the hybrid framework consistently lowers NLL and aleatoric uncertainty while preserving or slightly improving AUROC across heterogeneous tabular sources. Together with evaluation from PCADCTeval, the evidences forms a strong, preliminary case for frequency-aware preprocessing in lung-cancer tabular methodologies. The results are observable in table 4.

Measure	Dataset	Dataset 1	Dataset 2	Dataset 3
Accuracy	Pre	0.9812	0.8846	0.8846
	Post	**0.9973**	**0.9103**	**0.9103**
AUROC	Pre	0.9988	0.9574	0.9574
	Post	**0.9985**	**0.9618**	**0.9618**
NLL	Pre	0.0525	0.2037	0.2037
	Post	**0.0094**	**0.1786**	**0.1786**
NLL (calibrated)	Pre	0.0525	0.2037	0.2037
	Post	**0.0094**	**0.1786**	**0.1786**
Aleatoric uncertainty	Pre	0.1050	0.1123	0.1123
	Post	**0.0335**	**0.0406**	**0.0406**

Table 4: Pre/Post performance across datasets.

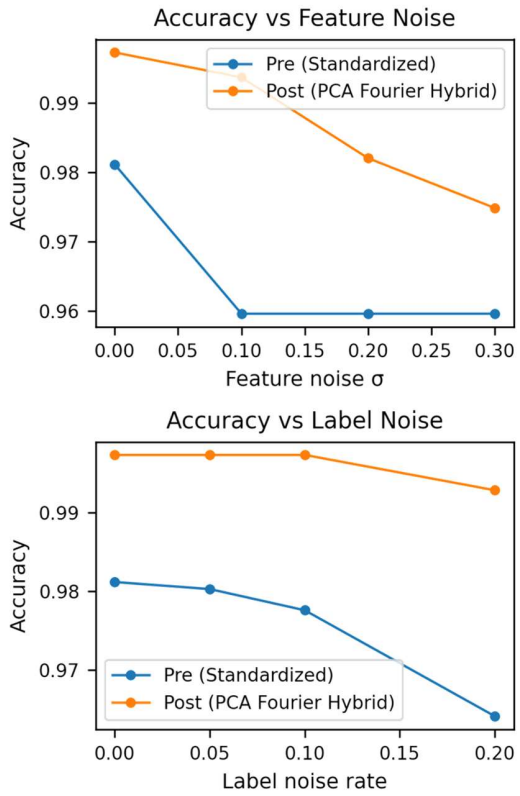

Figure 6: Accuracy degradation under injected noise. The PCA–Fourier hybrid framework consistently outperforms the standardized baseline across both feature noise (top) and label noise (bottom). As noise levels increase, the hybrid maintains higher accuracy, demonstrating robustness and resilience to data perturbations.

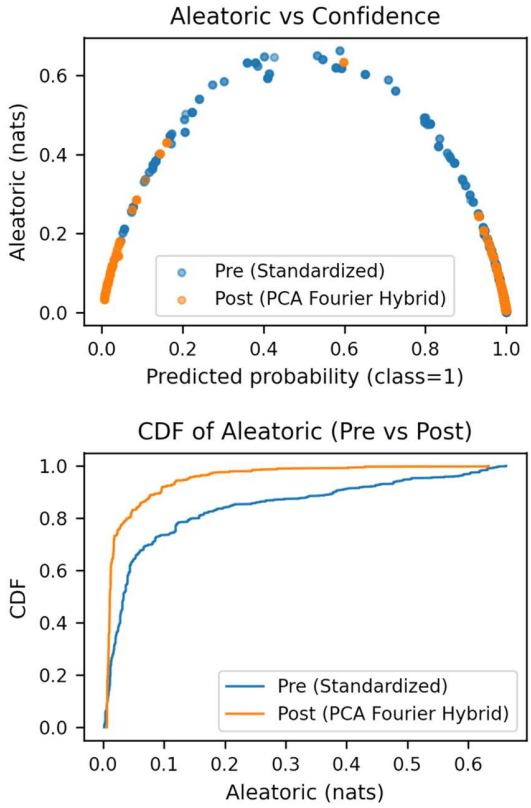

Figure 7: Aleatoric uncertainty plotted against predictive confidence (top) and its cumulative distribution (bottom). The hybrid PCA–DCT representation compresses uncertainty around the decision boundaries and shifts the distribution toward lower values, reflecting more stable predictions.

Conclusion

In this paper, we introduce a frequency-aware PCA–DCT hybrid for lung-cancer tabular prediction. The framework orders features by mutual information, retains low-frequency DCT coefficients, and preserves global structure via PCA, yielding a compact representation. Across three public datasets, **the hybrid reduced mean aleatoric uncertainty by approximately 7%, lowered NLL by 3.1%, and preserved AUROC. These effects persisted under injected noise and were supported by paired Wilcoxon tests.** The method is simple, data-efficient, and integrates readily into existing workflows. Future work targets broader validation, automated tuning, improved selective prediction, and integration with shift and causality-aware clinical models.

References

Ahmed, N.; Natarajan, T.; and Rao, K. R. 1974. Discrete Cosine Transform. *IEEE Transactions on Computers*, 23(1): 90–93.

Alizadehsani, R.; Roshanzamir, M.; Hussain, S.; Khosravi, A.; Koohestani, A.; Zangooei, M.; Abdar, M.; Beykikhoshk, A.; Shoeibi, A.; Zare, A.; Panahiazar, M.; Nahavandi, S.; Srinivasan, D.; Atiya, A.; and Acharya, U. 2024. Handling of uncertainty in medical data using machine learning and probability theory techniques: a review of 30 years (1991-2020). *Annals of Operations Research*, 339(3): 1077–1118.

Gawlikowski, J.; Tassi, C. R. N.; Ali, M.; Lee, J.; Humt, M.; Feng, J.; Kruspe, A.; Triebel, R.; Jung, P.; Roscher, R.; et al. 2023. A survey of uncertainty in deep neural networks. *Artificial Intelligence Review*, 56(1): 1513–1518.

Heid, E.; McGill, C. J.; Vermeire, F. H.; and Green, W. H. 2023. Characterizing uncertainty in machine learning for chemistry. 63(13): 4012–4029.

Henriques, R. N.; Ianus, A.; Novello, L.; Jovicich, J.; Jespersen, S. N.; and Shemesh, N. 2023. Efficient PCA denoising of spatially correlated redundant MRI data. *Imaging Neuroscience*, 1: 1–26.

Hoarau, A.; Quost, B.; Destercke, S.; and Waegeman, W. 2025. Reducing Aleatoric and Epistemic Uncertainty through Multi-modal Data Acquisition. *arXiv preprint arXiv:2501.18268*.

Jolliffe, I. T.; and Cadima, J. 2016. Principal component analysis: a review and recent developments. *Philosophical*

transactions of the royal society A: Mathematical, Physical and Engineering Sciences, 374(2065): 20150202.

Kendall, A.; and Gal, Y. 2017. What uncertainties do we need in bayesian deep learning for computer vision? *Advances in neural information processing systems*, 30.

Kipele, D.; and Greyson, K. A. 2023. Poisson noise reduction with nonlocal-pca hybrid model in medical x-ray images. *Journal of Image and Graphics*, 11(2): 178–184.

Lehtinen, J.; Munkberg, J.; Hasselgren, J.; Laine, S.; Karras, T.; Aittala, M.; and Aila, T. 2018. Noise2Noise: Learning Image Restoration without Clean Data. In *Proceedings of the 35th International Conference on Machine Learning*, volume 80 of *Proceedings of Machine Learning Research*, 2965–2974.

Li, X. 2025. On Context-Content Uncertainty Principle.

Liao, W.; Coupland, C. A.; Burchardt, J.; Baldwin, D. R.; Gleeson, F. V.; Baldwin, D.; and Gleeson, F. V. 2023. Predicting the future risk of lung cancer: development, and internal and external validation of the CanPredict (lung) model in 19.67 million people and evaluation of model performance against seven other risk prediction models. *The Lancet Respiratory Medicine*, 11(8): 685–697.

Malinin, A.; and Gales, M. 2018. Predictive uncertainty estimation via prior networks. *Advances in neural information processing systems*, 31.

Ovadia, Y.; Fertig, E.; Ren, J.; Nado, Z.; Sculley, D.; Nowozin, S.; Dillon, J.; Lakshminarayanan, B.; and Snoek, J. 2019. Can you trust your model's uncertainty? evaluating predictive uncertainty under dataset shift. *Advances in neural information processing systems*, 32.

Pal, S.; Roy, S.; Banerjee, A.; Majumdar, K.; Gupta, U.; Rana, S.; and Shetty, S. 2023. A Unified Approach for Identification and Analysis of the Sources of Uncertainty in Machine Learning Techniques. In *International Conference on Computing and Communication Networks*, 443–453. Springer.

Sensoy, M.; Kaplan, L. M.; and Kandemir, M. 2018. Evidential Deep Learning to Quantify Classification Uncertainty. In *Advances in Neural Information Processing Systems*, volume 31, 3183–3193.

Tsaneva-Atanasova, K.; Pederzanil, G.; and Laviola, M. 2025. Decoding uncertainty for clinical decision-making. *Philosophical Transactions A*, 383(2292): 20240207.

Vaish, P.; Wang, S.; and Strisciuglio, N. 2024. Fourier-basis functions to bridge augmentation gap: Rethinking frequency augmentation in image classification. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 17763–17772.

Wang, G.; Li, W.; Aertsen, M.; Deprest, J.; Ourselin, S.; and Vercauteren, T. 2019. Aleatoric Uncertainty Estimation with Test-Time Augmentation for Medical Image Segmentation with Convolutional Neural Networks. *Neurocomputing*, 338: 34–45.

Wang, T.; Wang, Y.; Zhou, J.; Peng, B.; Song, X.; Zhang, C.; Yan, L. K.; et al. 2025. From aleatoric to epistemic: Exploring uncertainty quantification techniques in artificial intelligence.