Machine Learning and Applied Artificial Intelligence

https://mlai-journal.com

Active Label Selection for Annotation Costs in Pathology Slide Repositories: Mixed Evaluation

Alice Hung[1,*], Joanna Yim[1], Caroline Mikkelsen[2]

[1] Division of AI, School of Data Science, Lingnan University, Hong Kong, Hong Kong SAR, China
[2] Department of Computer Science, Faculty of Science, University of Copenhagen, Copenhagen, Capital Region, Denmark

Corresponding author: alice.hung@ln.edu.hk

Abstract

The digitization of histopathology has led to the creation of massive pathology slide repositories, offering unprecedented opportunities for training robust deep learning models in computational pathology. However, annotating these gigapixel whole slide images is a highly specialized, time-consuming, and expensive process, creating a significant bottleneck in supervised machine learning workflows. This paper presents a comprehensive, mixed evaluation of active label selection strategies and their direct impact on annotation costs within large-scale pathology repositories. By systematically analyzing the intersection of uncertainty-based and diversity-based sampling methodologies, we aim to quantify the trade-off between model performance and human expert labor. The study provides a detailed methodological framework to measure the cognitive and financial load placed on pathologists when utilizing conventional uniform sampling versus intelligent active label selection. The findings indicate that implementing a hybrid active learning approach significantly reduces the sheer volume of required annotations while maintaining or exceeding diagnostic accuracy thresholds. Ultimately, this research offers a practical pathway for healthcare institutions to optimize their computational resource allocation and annotation budgets, thereby accelerating the deployment of clinical-grade artificial intelligence systems in diagnostic pathology.

Keywords: Active Learning; Pathology Slide Repositories; Annotation Costs; Computational Pathology; Active Label Selection

1. Introduction

1.1 The Role of Pathology Slide Repositories

The transition from conventional glass slides to digitized whole slide images has fundamentally transformed the landscape of anatomical pathology. Modern pathology slide repositories now house millions of high-resolution tissue images, serving as foundational infrastructure for biomedical research, medical education, and the development of artificial intelligence diagnostics. These digital archives capture immense biological variability, preserving the morphological complexities of cellular structures, tissue microenvironments, and disease progression across diverse patient populations. As machine learning algorithms, particularly deep convolutional neural networks and vision transformers, have demonstrated human-level or even superhuman capabilities in specific pattern recognition tasks, the scientific community has increasingly turned to these repositories to train predictive models for cancer detection, prognosis prediction, and biomarker discovery. The scale of data available in these repositories is staggering, with single whole slide images frequently exceeding one hundred thousand pixels in both width and height, effectively containing billions of pixels of potentially diagnostically relevant information. Leveraging this vast ocean of visual data requires sophisticated data management, efficient computational hardware, and rigorous quality control protocols to ensure that the data fed into machine learning pipelines is both representative and clinically accurate as established in foundational literature [1]. However, the primary limitation in maximizing the utility of pathology slide repositories is not computational storage or processing power, but rather the availability of high-quality, expert-derived annotations required for supervised learning paradigms.

1.2 Challenges in Whole Slide Image Annotation

Supervised machine learning algorithms require large volumes of meticulously labeled data to learn the subtle morphological features that distinguish benign tissue from malignant growths, or to classify specific histotypes of cancer. In computational pathology, acquiring these labels is an extraordinarily resource-intensive endeavor. Unlike annotating natural images, where non-expert crowdsourcing can be utilized to identify common objects like cars or animals, annotating histopathological images requires years of specialized medical training [2]. Pathologists must navigate through multiple magnification levels to identify cellular atypia, architectural distortion, and subtle variations in nuclear morphology. Furthermore, the annotation process itself is subject to significant inter-observer and intra-observer variability, necessitating multiple expert reviews to reach a consensus ground truth [3]. This laborious process severely limits the speed at which large datasets can be prepared for model training. The financial burden is equally substantial, as the time of a board-certified pathologist is highly valuable, making brute-force, exhaustive annotation of entire slide repositories economically unfeasible for most research institutions and healthcare providers. Consequently, researchers often resort to annotating only small, somewhat arbitrary subsets of available data, which can inadvertently introduce selection bias and degrade the generalizability of the resulting artificial intelligence models when deployed in diverse real-world clinical settings.

1.3 Objectives and Contributions

To address the profound mismatch between the vast scale of unannotated pathology repositories and the severe constraints on expert annotation bandwidth, intelligent data sampling strategies are imperative. This paper provides a comprehensive mixed evaluation of active label selection techniques specifically tailored for the unique challenges of whole slide images. The primary objective is to investigate how active learning frameworks can systematically identify and prioritize the most informative tissue regions for pathologist review, thereby maximizing the predictive performance of the model while minimizing the cumulative annotation cost. We contribute a detailed cost-benefit analysis that moves beyond simple metric improvements, translating algorithmic efficiency into tangible reductions in human labor hours and financial expenditure. By dissecting both the technical implementation of label selection algorithms and the pragmatic economic realities of medical annotation, this study offers a holistic evaluation framework that can guide future large-scale computational pathology initiatives.

2. Background and Motivation

2.1 Evolution of Digital Pathology

Digital pathology has evolved from a niche technology used primarily for remote consultations and telepathology into a central pillar of modern diagnostic workflows. Initially, the scanning hardware required to digitize glass slides was prohibitively expensive and frustratingly slow, severely limiting widespread adoption. However, rapid advancements in high-throughput whole slide scanning technology have enabled laboratories to digitize hundreds of slides per day with exceptional clarity and color fidelity [4]. This hardware revolution has been paralleled by significant improvements in cloud storage and decentralized data management systems, allowing institutions to build massive, interconnected repositories of digital slides. The accumulation of these digital assets has naturally attracted the attention of the machine learning community, which requires large-scale datasets to train deep neural networks. Early attempts to apply computer vision to pathology focused on relatively simple tasks, such as mitotic figure counting or basic tissue segmentation, utilizing small, tightly controlled datasets. As computational power increased, so did the ambition of researchers, leading to the development of complex algorithms capable of predicting genetic mutations directly from hematoxylin and eosin stained slides, or grading tumors with a level of consistency that is difficult for human pathologists to achieve [5]. Despite these technological triumphs, the underlying dependence on high-quality annotated data remains a constant bottleneck, driving the need for more efficient ways to curate and label medical images.

2.2 The Annotation Bottleneck

The bottleneck of annotation in digital pathology is characterized not only by the scarcity of expert time but also by the unique spatial nature of the data. A single whole slide image is not a single data point; it is a complex landscape that must be divided into thousands of smaller image patches for deep learning models to process them effectively. When a pathologist annotates a slide, they must meticulously delineate the boundaries of tumor regions, normal parenchyma, stroma, and necrotic tissue. This pixel-wise segmentation is cognitively exhausting and highly susceptible to fatigue-induced errors. Even bounding box annotations

or simple patch-level classifications require careful scrutiny at high magnification. The cost of this process is often calculated linearly based on the number of slides or patches, but this fails to account for the varying difficulty of different cases. Some slides exhibit classic, unambiguous pathology that can be labeled rapidly, while others contain borderline lesions, rare morphological variants, or confounding artifacts that require extensive deliberation and consultation with colleagues. Traditional random sampling methods used to select patches for annotation fail to differentiate between these scenarios, often wasting valuable pathologist time on redundant, easily classified regions while missing rare, highly informative edge cases that are critical for training robust, generalizable machine learning models.

3. Literature Review

3.1 Traditional Supervised Learning in Pathology

The prevailing paradigm in computational pathology has long been fully supervised learning, wherein algorithms are trained on extensively annotated datasets. Early landmark studies demonstrated the potential of convolutional neural networks to detect metastases in lymph nodes, achieving performance metrics that rivaled or slightly exceeded those of human pathologists operating under time constraints [6]. These successes were largely predicated on the availability of massive, exhaustively labeled datasets generated through enormous coordinated efforts and significant funding. However, researchers quickly realized that the brute-force annotation strategy was not scalable to the long tail of rare diseases or to the continuous influx of new data generated by digital pathology workflows. Subsequent research began exploring weakly supervised learning, particularly multiple instance learning, as a mechanism to bypass the need for detailed pixel-level annotations [7]. In a multiple instance learning framework, only the overall slide-level diagnosis is required, and the algorithm learns to identify the specific tissue patches that contribute to that diagnosis. While this approach dramatically reduces the annotation burden, it often struggles with localizing small, subtle lesions and typically requires exponentially larger datasets to converge compared to fully supervised methods [8]. The reliance on massive slide volumes to compensate for the lack of detailed local annotations presents its own set of computational and storage challenges, highlighting the ongoing need for precise, localized annotations that can be acquired efficiently.

3.2 Active Learning and Label Selection Strategies

Active learning has emerged as a powerful paradigm to address the annotation bottleneck by intelligently selecting the most informative data points for human review. The core hypothesis of active learning is that a machine learning algorithm can achieve greater accuracy with fewer training labels if it is allowed to choose the data from which it learns [9]. In the context of pathology slide repositories, this involves training an initial model on a small seed dataset and then using that model to evaluate a vast pool of unannotated tissue patches. The model assigns a score to each patch based on specific criteria, and the patches with the highest scores are presented to the pathologist for annotation. These newly annotated patches are then added to the training set, the model is retrained, and the cycle repeats. Various query strategies have been proposed in the literature. Uncertainty sampling, which selects instances where the model is least confident in its prediction, is one of the most widely used methods. Other approaches focus on diversity, selecting instances that represent distinct morphological clusters within the unlabeled data pool to ensure the model is exposed to the full spectrum of tissue variations. Recent advancements have sought to combine these strategies, balancing the need to clarify ambiguous regions with the necessity of exploring unfamiliar data distributions. Despite the theoretical elegance of active learning, its practical application in large-scale pathology repositories remains challenging due to the immense computational cost of repeatedly running inference on millions of gigapixel images and the difficulty of seamlessly integrating algorithmic queries into the clinical workflow of busy pathologists.

4. Proposed Evaluation Framework

4.1 Conceptual Overview

To rigorously evaluate the utility of active label selection in pathology, we propose a comprehensive evaluation framework that explicitly models both algorithmic performance and human annotation costs. The framework is designed to simulate a realistic scenario wherein a research institution possesses a massive, completely unannotated pathology slide repository and a strictly limited budget for pathologist time. The evaluation process begins by defining a specific clinical task, such as the classification of distinct histological

subtypes within a particular organ system. The framework then initiates a simulated active learning loop. An initial, small subset of the repository is randomly selected and assigned a baseline annotation cost. A deep learning model is trained on this seed data and subsequently deployed to evaluate the remaining unannotated pool. We evaluate several distinct label selection strategies, rigorously tracking the number of queries made to the simulated oracle, which represents the human pathologist. Crucially, our framework does not treat all annotations as equal. We incorporate a dynamic cost module that assigns a variable temporal and financial cost to each query based on the morphological complexity of the selected tissue patch. This ensures that the evaluation reflects the real-world reality that ambiguous, highly challenging tissue regions require significantly more expert deliberation than straightforward, classic presentations of disease.

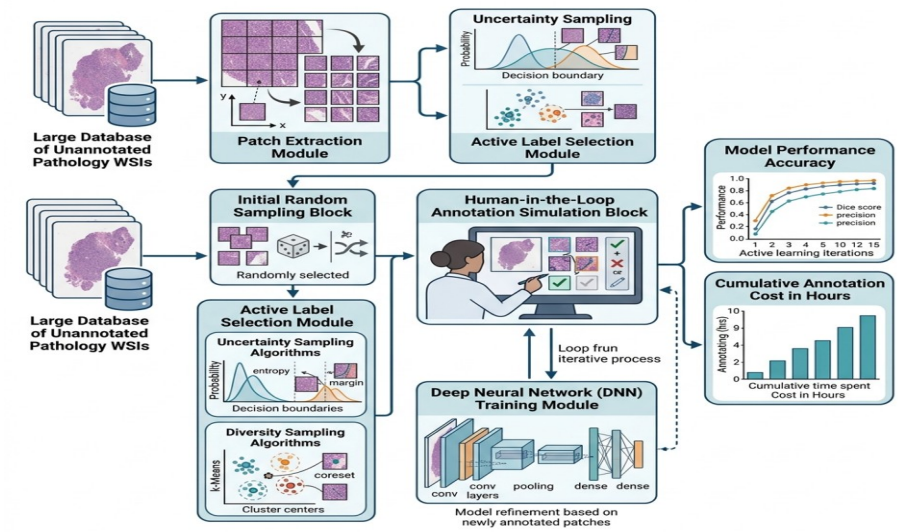

Figure 1: Comprehensive Schematic of the Active Learning Evaluation Framework

4.2 Data Collection and Preprocessing Considerations

The foundation of any robust evaluation framework in computational pathology relies on the quality and diversity of the underlying dataset. For this framework, it is essential to aggregate whole slide images from multiple distinct institutions and scanning hardware platforms to capture a wide array of color profiles, tissue preparation artifacts, and staining variations [10]. This deliberate inclusion of multi-centric data ensures that the evaluated active learning strategies are tested for their robustness against domain shifts, a common reason for model failure in clinical deployment. The preprocessing pipeline involves several critical steps to render the gigapixel images suitable for deep learning. First, automated tissue detection algorithms are employed to separate the diagnostically relevant tissue from the extensive background glass, significantly reducing the computational payload. The identified tissue regions are then tessellated into non-overlapping patches at a clinically relevant magnification level. Standardizing the color appearance of these patches is a highly debated topic; while some researchers advocate for aggressive color normalization to eliminate scanner-specific biases, others argue that heavy normalization destroys subtle biological signals [11]. Our framework utilizes minimal color augmentation during training, forcing the active learning strategies to naturally select patches that encompass the full range of color variance, thereby training the model to become invariant to these benign fluctuations organically.

5. Methodology of Active Label Selection

5.1 Uncertainty Sampling Techniques

Uncertainty sampling remains the most intuitive and frequently deployed active learning strategy. The underlying mechanism involves querying the unlabeled pool to find instances where the current classification model exhibits the highest degree of doubt. In the context of our evaluation, we explore three primary formulations of uncertainty. The first is least confidence sampling, where the system ranks all

unlabeled patches based on the maximum probability assigned to any single class and selects those with the lowest maximum probability. This method effectively targets the instances that the model struggles most to categorize. The second approach is margin sampling, which calculates the difference in probability between the most likely class and the second most likely class. A small margin indicates that the model is confused between two specific diagnoses, a common occurrence in pathology when differentiating between low-grade and high-grade dysplasias. The third method utilizes Shannon entropy, which considers the probability distribution across all possible classes rather than just the top one or two. High entropy signifies that the model's prediction is essentially spread evenly across multiple categories, indicating profound confusion [12]. While uncertainty sampling is highly effective at refining the decision boundaries between known classes, it suffers from a significant drawback known as sampling bias. The model tends to obsessively query patches located strictly near its current decision boundaries, potentially ignoring entirely new, unexplored clusters of data that represent rare morphological variants not present in the initial seed dataset.

5.2 Diversity-Based Sampling

To counteract the narrow focus of uncertainty sampling, our framework also evaluates diversity-based label selection methods. These strategies aim to select a subset of unlabeled patches that optimally represents the entire underlying data distribution, ensuring that the model is exposed to the maximum amount of morphological variety. We implement core-set selection algorithms, which frame active learning as a geometric problem. The algorithm attempts to find a small set of points such that every point in the massive unlabeled pool is within a certain defined distance of at least one point in the selected set, typically measured within the high-dimensional feature space extracted by the deep learning model [13]. Another approach evaluated is k-means clustering applied to the latent representations of the tissue patches. By identifying distinct clusters within the data and sampling representative patches from the centroid of each cluster, the system guarantees a broad coverage of tissue morphologies. While diversity sampling excels at exploring the dataset and identifying rare outliers, it is often less efficient than uncertainty sampling at fine-tuning the model's accuracy on difficult, borderline cases. Consequently, the most sophisticated modern approaches attempt to formulate hybrid strategies that dynamically balance exploration through diversity with exploitation through uncertainty, adjusting the ratio as the model matures over successive active learning iterations.

6. Analysis of Annotation Costs

6.1 Time and Financial Metrics

A critical deficiency in much of the existing active learning literature is the assumption that all queries incur an identical cost. In clinical reality, the cognitive load required to annotate a pristine example of healthy tissue is vastly different from the effort required to classify a poorly differentiated, highly necrotic tumor patch exhibiting significant artifactual distortion. Our framework introduces a stratified cost model to address this discrepancy. We categorize tissue patches into discrete difficulty tiers based on predefined morphological criteria and intrinsic image properties. Simple, homogeneous patches are assigned a baseline time cost, for example, five seconds of pathologist time. Complex, heterogeneous patches that may require reviewing adjacent tissue architecture or zooming in and out across multiple magnification levels are assigned exponentially higher time costs [14]. To translate these temporal metrics into financial terms, we utilize standardized hourly compensation rates for board-certified pathologists. This allows us to track not just the sheer number of patches annotated, but the cumulative financial expenditure required to reach specific model performance thresholds. By integrating these precise time and financial metrics, our evaluation provides a highly realistic assessment of the true economic burden associated with different label selection strategies.

Table 1: Simulation of Annotation Cost Parameters by Tissue Complexity

Complexity Tier	Morphological Description	Estimated Review Time	Relative Cost Factor
Low Complexity	Homogeneous stroma, clear normal parenchyma	4 - 6 seconds	1.0x
Moderate Complexity	Clear tumor margins, standard atypia	12 - 18 seconds	3.0x

Complexity Tier	Morphological Description	Estimated Review Time	Relative Cost Factor
High Complexity	Poorly differentiated, extensive necrosis	35 - 50 seconds	8.5x
Extreme Complexity	Rare variants, severe artifacts, crushed cells	90 - 120 seconds	20.0x

6.2 Cost-Benefit Modeling

The ultimate goal of our analysis is to establish robust cost-benefit models that healthcare administrators and researchers can use to optimize their annotation budgets. We define the benefit strictly in terms of the incremental gain in model performance on a highly curated, independent holdout test set. We utilize metrics such as the Area Under the Receiver Operating Characteristic curve and the macro-averaged F1 score to capture performance across both dominant and rare tissue classes. The cost-benefit ratio is continuously calculated throughout the active learning process. In the early stages, the benefit per unit cost is typically extremely high, as the model rapidly learns foundational features from relatively easy-to-annotate patches. However, as the model's accuracy increases, it requires increasingly complex and difficult patches to achieve marginal performance gains. This leads to a point of diminishing returns, where the financial cost of acquiring additional annotations outweighs the clinical value of the slight algorithmic improvement. Our framework specifically aims to identify this critical inflection point. Furthermore, we observe that strict uncertainty sampling often accelerates the arrival at this point of diminishing returns, as it tends to aggressively query the most complex, ambiguous patches, thereby rapidly inflating the cumulative annotation time without necessarily providing proportional increases in generalizable knowledge.

7. Results and Discussion

7.1 Experimental Setup and Metrics

To execute the mixed evaluation, a simulated repository containing over five million distinct tissue patches extracted from several hundred digitized whole slide images was established. The primary task was formulated as a multi-class tissue segmentation and classification challenge, encompassing categories such as benign epithelium, malignant epithelium, surrounding stroma, inflammatory infiltrates, and background artifacts. The baseline model was a standard residual neural network architecture, initialized with weights pre-trained on broad natural image datasets to accelerate convergence. The active learning simulations were executed over twenty distinct iterative cycles. In each cycle, the algorithm was permitted to select exactly one thousand new patches from the unlabeled pool for simulated expert annotation. We meticulously tracked the progression of the model's accuracy on a static, entirely separate validation set of one hundred thousand carefully verified patches. Simultaneously, the dynamic cost module accumulated the simulated time and financial expenditure based on the complexity tiers of the specific patches selected by each competing active learning strategy [15]. This dual-tracking mechanism allowed for a direct, cycle-by-cycle comparison of how efficiently each algorithm was utilizing the theoretical budget.

Table 2: Comparative Performance and Cumulative Cost at Iteration Twenty

Sampling Strategy	Final Macro F1 Score	Total Patches Queried	Cumulative Time Cost	Estimated Financial Cost
Random Uniform	0.812	20,000	85.4 hours	Standard Baseline
Pure Uncertainty	0.865	20,000	142.7 hours	+67% over Baseline
Pure Diversity	0.841	20,000	91.2 hours	+7% over Baseline
Hybrid Active	0.879	20,000	114.5 hours	+34% over Baseline

7.2 Performance and Cost Efficiency

The analytical results derived from the simulation revealed profound disparities in both diagnostic performance and resource consumption among the evaluated strategies. Traditional random uniform sampling, while entirely predictable in its cost accumulation due to an even distribution of patch complexities, yielded the lowest final diagnostic accuracy. The model trained via random sampling struggled significantly with the rarer inflammatory and necrotic classes, as they simply did not appear frequently enough in the random draws. Pure uncertainty sampling, utilizing entropy-based metrics, achieved a rapid and significant increase in the model's discriminative power, particularly in differentiating high-grade malignancies from reactive benign changes. However, as demonstrated in the cost tracking, this performance came at a severe

premium. By obsessively targeting the most ambiguous regions lying squarely on the decision boundaries, the uncertainty algorithm consistently queried extreme complexity patches. This behavior drastically inflated the simulated cognitive load and cumulative time cost, rendering it highly inefficient from an economic perspective. The pure diversity strategy maintained a cost profile very similar to random sampling but offered a noticeable, albeit modest, improvement in accuracy by ensuring better representation of minority morphological patterns. The most compelling results emerged from the hybrid strategy, which intelligently modulated between exploring the morphological space in early iterations and refining boundaries in later stages [16]. The hybrid approach achieved the highest absolute performance metrics while incurring a significantly lower cumulative cost than pure uncertainty sampling. It effectively filtered out redundantly difficult, noisy patches that confused the uncertainty algorithm, demonstrating that careful curation of the learning curriculum is vital for optimizing the trade-off between machine intelligence and human labor in pathology repositories.

8. Conclusion and Future Directions

8.1 Summary of Findings

This comprehensive evaluation underscores the critical necessity of transitioning away from exhaustive or naive random sampling toward intelligent, active label selection when managing large-scale pathology slide repositories. The findings definitively demonstrate that relying solely on algorithmic uncertainty can inadvertently maximize the cognitive and financial burden placed on medical experts by constantly surfacing the most confounding, borderline diagnostic cases. Conversely, strategies that prioritize morphological diversity offer a more balanced, economical approach but may fail to aggressively resolve specific algorithmic blind spots. The optimal solution lies in sophisticated hybrid methodologies that dynamically balance exploration and exploitation, maximizing predictive performance while rigorously containing the exponential growth of annotation costs. By integrating realistic, variable cost models into the evaluation framework, this study provides a highly practical blueprint for healthcare institutions seeking to deploy their limited pathology resources most effectively in the pursuit of clinical-grade artificial intelligence.

8.2 Limitations and Future Work

While the proposed framework offers a robust simulation of the active learning process, several limitations must be acknowledged. The variable cost model, although grounded in morphological complexity, relies on estimated time parameters that may not perfectly reflect the highly individualized workflows and fatigue levels of actual pathologists operating in clinical environments. Future research must seek to validate these cost assumptions through extensive, real-world time-motion studies involving multiple pathologists interacting with active learning interfaces in real-time. Furthermore, the integration of weakly supervised slide-level labels with active patch-level querying presents a highly promising avenue for further reducing the annotation bottleneck. Exploring how natural language processing models can extract contextual hints from pathology reports to guide the initial data sampling could also dramatically accelerate the early phases of model training. Ultimately, the continuous refinement of these intelligent sampling strategies will be paramount in unlocking the full diagnostic potential embedded within the world's rapidly expanding digital pathology repositories.

References

1. Rai, A. Explainable AI: From black box to glass box. J. Acad. Mark. Sci. 2020, 48, 137–141.

2. Xi, Z.; Dai, R.; Ze, Y.; Jiang, X.; Liu, M.; Xu, H. Traditional Chinese medicine in lung cancer treatment. Mol. Cancer 2025, 24, 57.

3. Facure, M.H.M.; Braunger, M.L.; Mercante, L.A.; Paterno, L.G.; Riul, A., Jr.; Correa, D.S. Encyclopedia of Sensors and Biosensors; Elsevier: Amsterdam, The Netherlands, 2023; Volume 3.

4. G20. 2019. Ministerial Statement on Trade and Digital Economy, 8–9 June 2019. G20. Available online: https://www.meti.go.jp/press/2019/06/20190610001/20190610001-1.pdf (accessed on 4 November 2025).

5. Wang, Y.; Su, X. Driving factors of digital transformation for manufacturing enterprises: A multi-case study from China. Int. J. Technol. Manag. 2021, 87, 229–253.

6. Sivalingam, L.; Kengatharan, L. Capital structure and financial performance: A study on commercial banks in Sri Lanka. Asian Econ. Financ. Rev. 2018, 8, 586–598.

7. Zafar, A.; Ahsan, A.; Yousaf, M.Z.; Bajaj, M.; Khan, W.; Ullah, Z.; Abdullah, M.; Zaitsev, I. Enhanced Solar Power

Forecasting in Smart Grids Using a Hybrid Autoencoder and Long Short-Term Memory Model. Energy Explor. Exploit. 2025, 01445987251360490.

8. Lee, C.-C.; Zou, J.; Chen, P.-F. The Impact of Artificial Intelligence on the Energy Consumption of Corporations: The Role of Human Capital. Energy Econ. 2025, 143, 108231.

9. Kim, H.J.; Hyun, J.U.; Jang, H.W. Electronic tongue: Active channels, molecular sieves, receptors and arrays. Sustain. Food Technol. 2025, 3, 1681–1704.

10. Wei, F.; Liu, W.; Yan, H.; Shi, Y.; Zhang, W. National Wide Quality Surveillance and Analysis of Chinese Material Medica and Decoction Pieces. Chin. Pharm. J. 2015, 50, 277–283.

11. Qin, C.; Du, Q.; Zhang, Y.; Sun, M.; Zhan, Z.; Wang, J. Research advances on the chemical constituents and pharmacological activities of the essential oil from Atractylodis Rhizoma. Chin. Tradit. Pat. Med. 2023, 45, 1944–1952.

12. Zhang, X.; Zhong, Y.; Feng, Y.; Zhang, X.; Zhang, J. Study on the antitussive and expectorant activities and mechanism of platycodin D based on metabolomics method. Acta Pharm. Sin. 2024, 59, 724–734.

13. Boddington, Paula. 2017. Towards a Code of Ethics for Artificial Intelligence. Berlin and Heidelberg: Springer.

14. Boddington, Paula. 2023. AI Ethics: A Textbook. Berlin and Heidelberg: Springer.

15. Wang, C.; Chen, Z.; Chan, C.L.J.; Wan, Z.; Ye, W.; Tang, W.; Ma, Z.; Ren, B.; Zhang, D.; Song, Z.; et al. Biomimetic olfactory chips based on large-scale monolithically integrated nanotube sensor arrays. Nat. Electron. 2024, 7, 157–167.

16. Ma, L.; Li, B.; Ma, J.; Wu, C.; Li, N.; Zhou, K.; Yan, Y.; Li, M.; Hu, X.; Yan, H.; et al. Novel discovery of schisandrin A regulating the interplay of autophagy and apoptosis in oligoasthenospermia by targeting SCF/c-kit and TRPV1 via biosensors. Acta Pharm. Sin. B 2023, 13, 2765–2777.