Cross-Modal Attention Fusion of Radiomic Lung Nodule Features and Transcriptomic Phase-Separation Signatures for Explainable Cancer Risk Stratification

Arjun D. Shah
Department of Electrical Engineering and Computer Science, University of Kansas, Lawrence, KS, USA.
arjuns@ku.edu

Pierre D. Kregory
Department of Electrical Engineering and Computer Science, University of Missouri, Columbia, MO, USA.
pierrework@missouri.edu

Kasper Butler
School of Computing, Clemson University, Clemson, SC, USA.
kasper.butler@clemson.edu

Massimo Taylor
Department of Computer Science, University of New Hampshire, Durham, NH, USA.
massimot@unh.edu

Abstract

Precision oncology demands risk stratification frameworks that synthesize heterogeneous biological evidence into actionable and interpretable predictions. This paper presents a system-level architecture for cross-modal attention fusion that integrates radiomic features extracted from lung nodule imaging with transcriptomic signatures shaped by biomolecular phase separation, aiming to deliver explainable cancer risk stratification. The work does not propose a single novel model; rather, it offers a sustained examination of the structural, governance, infrastructure, and policy dimensions associated with deploying such a cross-modal system in real clinical contexts. We argue that contemporary multimodal fusion methods, while powerful, often obscure the reasoning behind predictive outcomes, creating barriers to clinical trust, regulatory approval, and equitable deployment. By centering attention mechanisms that explicitly decouple modality-specific feature extraction from cross-modal interaction, the architecture promotes transparency through interpretable attention pathways. We discuss how radiomic features, long established for noninvasive tumor phenotyping, can be enriched by dual-attention segmentation networks, and how transcriptomic phase-separation signatures, which reveal condensate-driven gene regulation, provide a dynamic molecular complement. The paper critically examines trade-offs between predictive performance and explainability, explores infrastructure requirements for ingesting and aligning radiology and sequencing data at scale, and addresses fairness, robustness, and governance challenges that arise when such systems are embedded in clinical decision support. Throughout, the discussion emphasizes systemic resilience, policy alignment with evolving medical AI regulation, and sustainability of cross-institutional data collaborations. By treating

the fusion pipeline as a socio-technical infrastructure, we provide a forward-looking perspective that connects attention-driven representation learning with the broader imperatives of responsible translation in oncology.

Keywords

cross-modal attention, radiomics, transcriptomics, phase separation, lung nodule, risk stratification, explainable AI, multimodal fusion, governance, clinical deployment.

1. Introduction

Lung cancer remains the leading cause of cancer mortality globally, and early detection through low-dose computed tomography has demonstrated mortality reduction yet simultaneously produces a high volume of indeterminate pulmonary nodules whose management depends on nuanced risk assessment. Radiomics, the high-throughput extraction of quantitative image features, has moved beyond purely descriptive imaging phenotypes toward predictive modeling of tumor behavior and clinical outcomes [1], [2], [3]. In parallel, transcriptomic profiling has illuminated molecular subtypes and prognostic gene expression signatures, and recent discoveries in biomolecular phase separation have opened a new axis of regulation that connects the physical organization of transcriptional condensates to oncogenic programs. Integrating these two domains—radiomic imaging features and transcriptomic phase-separation signatures—through a cross-modal attention fusion framework holds the promise of more accurate and biologically grounded risk stratification. However, the system-level challenges of designing, validating, deploying, and governing such a multimodal pipeline demand far more than a purely technical optimization exercise; they require an interdisciplinary examination of architecture, explainability, fairness, infrastructure, and policy.

The present paper adopts a large-scale systems perspective to analyze how cross-modal attention fusion can bridge the macroscopic view provided by radiology with the microscopic regulatory landscape illuminated by phase-separated transcriptional complexes. We situate this investigation within a broader discourse on responsible AI in medicine, emphasizing that explainability is not merely a post-hoc add-on but must be woven into the architectural fabric of the fusion mechanism. Drawing on recent advances in attention decoupling, dual-attention segmentation networks, and transcriptomic condensate biology, we map the trade-offs associated with different design choices and evaluate their implications for clinical governance, algorithmic robustness, and health equity. The analysis proceeds without mathematical notation or algorithmic pseudocode, focusing instead on conceptual structural reasoning and system-level argumentation that bridges machine learning research, molecular biology, and health policy.

2. Radiomic Feature Extraction and Attention-Based Nodule Analysis

Radiomics fundamentally repositions medical images as mineable data fields, extracting shape, texture, intensity, and wavelet features that are often beyond human visual perception. Foundational work established that such features correlate with tumor histology, staging, and gene expression patterns, enabling noninvasive phenotyping [1], [2], [3]. In lung nodule assessment, the precise segmentation of the nodule boundary is a critical prerequisite for reliable feature extraction; inaccuracies at this stage propagate errors through all downstream modeling. Recent segmentation algorithms have leveraged path aggregation and dual attention mechanisms to enhance boundary delineation, as exemplified by PDU-Net architectures that combine multi-scale feature aggregation with channel and spatial attention

to refine nodule margins [4]. Such attention-augmented segmentation not only improves the Dice coefficient and volumetric accuracy but also produces richer intermediate feature maps that can be repurposed for subsequent diagnostic reasoning.

Attention mechanisms have become indispensable in modern medical image analysis. While early convolutional neural networks extracted features in a spatially uniform manner, attention-based models selectively amplify salient regions, mimicking a clinician's focal examination of suspicious areas. In the context of lung nodule classification and false positive reduction, multi-view convolutional architectures have historically relied on aggregating decisions from multiple two-dimensional planes [5]. More recently, three-dimensional deep learning models trained end-to-end on low-dose CT volumes have shown radiologist-level performance in cancer screening, implicitly learning attention-like weighting through their deep architectures [6]. The explicit introduction of self-attention, as popularized by the Transformer [7], pushed this further by enabling global context modeling, which is particularly beneficial for nodules that exhibit contextual relationships with surrounding vascular and pleural structures.

Despite these advances, the transition from a unimodal imaging pipeline to a multimodal fusion system that incorporates transcriptomic data amplifies the need for careful attention design. In a cross-modal setting, attention must not only weigh spatial features within the image but also decide how much each modality contributes to the final risk prediction for a given case. The survey of multimodal machine learning provides a taxonomy of fusion strategies, including early fusion, late fusion, and hybrid approaches, each with distinct implications for representational entanglement and interpretability [8]. Early fusion integrates features before decision-making, often blending modalities in a latent space that obscures their individual contributions, while late fusion combines independent unimodal predictions and offers modularity at the cost of missing intermodal interactions. Cross-modal attention offers a middle ground, allowing the model to dynamically reweight radiomic features based on transcriptomic context and vice versa, thereby modeling interactions without fully entangling representations. The architectural decision thus becomes a governance choice: how much transparency is preserved by maintaining separable modality-specific processing streams before the attention-based interaction layer.

3. Transcriptomic Phase-Separation Signatures and Their Prognostic Value

In parallel to imaging, transcriptomic profiling has long been used to define molecular subtypes of lung cancer and to derive prognostic gene expression signatures. However, the emerging field of biomolecular phase separation adds an important physical layer to transcriptional regulation that is not captured by conventional differential expression analysis. Phase separation describes the process by which proteins and nucleic acids demix from the surrounding milieu to form membraneless condensates, which can concentrate transcription factors, coactivators, and RNA polymerase II at super-enhancer regions, thereby driving oncogenic gene expression programs [9], [10]. The concept that transcriptional control is organized through dynamic condensates reshapes our understanding of how a relatively static genomic blueprint yields context-dependent transcriptomic output.

A particularly instructive example is the phase separation behavior of the YAP-MAML2 fusion protein, which differentially regulates the transcriptome depending on its condensation state. Chung et al. demonstrated that the physical phase transition of this oncogenic chimera modulates a gene expression program distinct from that driven by its non-condensed counterpart, revealing that risk-relevant transcriptional heterogeneity can arise from

biophysical states invisible to conventional sequencing readouts [11]. This finding underscores the need to move beyond gene lists and instead incorporate condensate-associated transcriptional signatures that capture the functional consequences of phase separation. The Cancer Genome Atlas characterization of lung adenocarcinoma provided a foundational molecular taxonomy, but its data predate systematic incorporation of phase-separation annotations [12]. Subsequent studies integrating both tumor and stromal gene expression signatures have improved survival stratification, yet they have not explicitly modeled the contribution of condensate-driven regulatory programs [13].

When such transcriptomic phase-separation signatures are positioned as a modality to be fused with radiomics, they offer a complementary biological perspective: while radiomics captures macroscopic tissue architecture and vascular patterns that reflect tumor-stroma interactions, the phase-separation transcriptomic signature captures the intracellular biophysical states that drive proliferation, invasion, and drug resistance. The challenge from a systems perspective is that these two data modalities operate at entirely different spatial and temporal scales. A CT image is a snapshot at millimeter resolution that integrates structural information across the entire nodule, whereas a transcriptomic signature is a bulk or spatially averaged molecular measurement from a biopsy or resection specimen. Bridging this scale gap through attention mechanisms requires the model to learn which radiomic textures correspond to specific condensate-regulated transcriptional programs—a mapping that is neither linear nor directly observable but must be inferred through supervised learning on paired cohorts.

4. Cross-Modal Attention Fusion Architecture

Designing an architecture to fuse radiomic features and transcriptomic phase-separation signatures demands a principled approach to cross-modal interaction that preserves interpretability while capturing complex dependencies. The notion of decoupling feature-driven attention from multimodal fusion attention, originally developed for person re-identification [16], provides a valuable conceptual template. In that framework, feature-driven attention focuses on extracting discriminative patterns within each modality independently, while multimodal fusion attention explicitly models the interactions between modality-specific representations only after they have been refined. Applying this decoupling principle to oncology risk stratification means that the network first computes radiomic attention maps that highlight morphologically suspicious regions within the nodule and transcriptomic attention distributions that weigh the relevance of different gene modules, especially those annotated as phase-separation-regulated. Only then does a cross-attention module allow the radiomic pathway to query the transcriptomic context and vice versa, generating modality-aware risk representations that remain traceable back to the original features.

The advantages of this decoupling are systemic rather than merely computational. For regulatory reviewers and clinical end users, it becomes possible to inspect whether a high-risk prediction is driven primarily by imaging features—such as spiculated margins or heterogeneous texture—by molecular signatures of YAP-MAML2-like condensate activity, or by a specific synergy between the two. Explainability methods such as SHAP [14] and LIME [15] can be applied to the fused representation, but their effectiveness depends on the underlying feature space being semantically structured. When feature-driven extraction is separated from cross-modal interaction, Shapley values can be computed not only for raw features but also for the attention-modulated contributions of each modality, providing explanations that map to clinically meaningful concepts. This contrasts with strongly

entangled early-fusion models, where the contribution of a gene expression measurement is inseparably mixed with a wavelets coefficient deep inside a black-box latent space.

Implementing such a decoupled attention architecture in a real-world clinical environment imposes substantial data engineering and modeling discipline. Paired imaging and transcriptomic data are scarser than unimodal datasets; large-scale biobanks and clinical trials that collect both CT images and RNA sequencing from the same patient are only beginning to mature. Consequently, the fusion model must be designed to leverage pretraining on large unimodal corpora—such as the National Lung Screening Trial CT collections for imaging and The Cancer Genome Atlas or Genotype-Tissue Expression datasets for transcriptomics—before fine-tuning on the more limited paired data. Transfer learning and domain adaptation become critical infrastructure components, as radiomic feature extractors trained on diagnostic classification may not directly transfer to the risk stratification task without careful recalibration. The decoupled attention architecture supports this modular pretraining because the unimodal feature encoders can be developed and validated independently before being plugged into the cross-attention layer, reducing the coupling of failure modes and allowing each component to be updated as new molecular annotation resources or imaging protocols become available.

5. Explainability Framework for Risk Stratification

Explainability in a cross-modal attention fusion system must be conceptualized at multiple levels: local explanation for an individual patient's risk score, global understanding of the model's learned associations, and structural transparency of the fusion mechanism itself. Local post-hoc explanations can be produced by computing attention weight visualizations over the nodule image and transcriptomic heatmaps, akin to the Grad-CAM approach that highlights discriminative image regions [20]. When combined with SHAP value analysis [14], the clinician can be shown that a particular nodule received a high-risk score largely because a specific region of ground-glass opacity attended strongly to a condensate-related gene module involved in epithelial-mesenchymal transition. Such narratively structured explanations are far more actionable than a single probability output and align with emerging regulatory expectations for clinical decision support software.

However, the provision of explainability must not be conflated with full mechanistic understanding. The biological reality is that phase-separated condensates and radiographic textures are linked through intermediating layers of cell density, extracellular matrix remodeling, and angiogenesis that are not directly modeled. The attention mechanism captures statistical associations that are context-dependent and may fail under distributional shifts, such as when a nodule appears in a patient with interstitial lung disease that alters background lung texture. Robustness to such shifts is therefore an essential companion to explainability. Adversarial testing, where carefully perturbed images or gene expression profiles are fed to the model, can reveal whether small clinically imperceptible changes cause drastic swings in attention patterns and risk scores [21]. A system that passes adversarial stress tests while maintaining consistent explanations across plausible variations is more likely to earn clinical trust.

From a governance perspective, explainability reports generated by the system must be integrated into clinical workflows in a manner that does not increase cognitive burden. Decision support tools that present rich attention maps and transcriptomic annotations might overwhelm a radiologist or oncologist who is making decisions under time pressure. Human-computer interaction research in medical AI indicates that effective explanation interfaces

should be customizable, allowing the user to drill down from a summary evidence statement into detailed modality-specific justifications on demand. The decoupled attention architecture naturally supports this tiered interface because the modality-specific and cross-modal contributions are computed as separable quantities before final fusion, enabling the system to surface statements such as "The imaging features primarily drove this risk estimate, with moderate additional contribution from the molecular signature."

6. System-Level Deployment and Infrastructure

Translating a cross-modal attention fusion model from a research prototype into a reliable clinical service entails a substantial infrastructure investment that spans data ingestion, model hosting, interoperability standards, and continual monitoring. The first challenge is the acquisition and alignment of radiological and transcriptomic data streams. Radiomic features require standardized image acquisition and reconstruction protocols to limit scanner-dependent variability; adherence to the Quantitative Imaging Biomarkers Alliance profiles is a minimal starting point. Transcriptomic data, especially when analyzed for phase-separation signatures, demand consistent RNA extraction, library preparation, and bioinformatic processing to ensure that condensate-related gene modules are computed comparably across institutions. The system must incorporate a data versioning and provenance layer that tracks exactly which imaging protocol and sequencing pipeline produced each training and inference instance, because model performance is tightly coupled to these preprocessing choices.

Model serving in a hospital network raises further architectural decisions. Running a deep attention model that processes both high-resolution CT volumes and high-dimensional gene expression vectors in real time is computationally intensive, and latency requirements for clinical decision support are stringent. A pragmatic deployment might adopt an asynchronous architecture where the radiomic encoder runs on the picture archiving and communication system server immediately upon CT acquisition, while the transcriptomic encoder processes sequencing data upon availability in the electronic health record, and the lightweight cross-attention fusion layer is invoked only when both modalities are present. This design decouples the lifecycles of the two data sources and aligns with the decoupled attention logic. It also facilitates incremental deployment, starting with radiomic-only risk assessment in sites that lack routine transcriptomic profiling, and upgrading to full cross-modal fusion when molecular data become available, without altering the imaging workflow.

Sustainability considerations demand that the infrastructure supports model updating without catastrophic forgetting. As phase-separation biology advances and new condensate-modulated transcriptional programs are discovered, the transcriptomic encoder and the cross-attention layer will need periodic retraining on expanded datasets that reflect contemporary biological knowledge. The decoupled architecture allows the transcriptomic pathway to be updated without necessarily retraining the radiomic feature extractor, provided that the cross-attention interface remains compatible. This modularity reduces the carbon footprint and computational cost of model maintenance, an increasingly important factor in environmentally sustainable AI for healthcare [19]. It also lowers the barrier for multi-institutional collaborations, since each site can contribute updates to specific components based on its local expertise—radiology departments refining the imaging encoder, molecular pathology labs refining the transcriptomic encoder—without requiring a monolithic retraining of the entire system.

7. Governance, Fairness, and Policy Implications

The deployment of a cross-modal risk stratification system in oncology is inseparable from the governance frameworks that regulate medical software as a device, algorithmic fairness, and data privacy. Algorithmic bias in healthcare has been starkly demonstrated in widely used risk prediction tools, where reliance on healthcare cost as a proxy for health need systematically underestimated the severity of illness in Black patients [17]. A cross-modal fusion model trained predominantly on data from academic medical centers with predominantly White patient populations risks learning radiomic and transcriptomic associations that do not generalize across racial and ethnic groups. Radiomic features are influenced by differences in body habitus, lung density, and prevalence of comorbidities that vary across populations, and transcriptomic phase-separation signatures may be modulated by genetic ancestry and environmental exposures. Mitigating such bias requires both representative data collection and fairness-aware training objectives that penalize disparities in model performance across demographic subgroups. From a governance standpoint, proactive fairness auditing by independent bodies should be mandated before clinical deployment, and continuous post-market surveillance should monitor for emergent disparities as the patient population evolves.

The ethics of algorithms literature provides a mapping of the moral and epistemic considerations that must be addressed when opaque machine learning systems enter high-stakes clinical decisions [18]. The decoupled attention architecture offers structural transparency that aligns with procedural fairness principles: it allows stakeholders to understand which data streams influence decisions and to contest predictions when modality-specific explanations appear inconsistent with clinical judgment. Additionally, the inclusion of transcriptomic signatures that reflect fundamental biophysical processes raises novel regulatory questions. If a phase-separation signature derived from a specific oncogenic fusion drives risk stratification, the system is effectively performing a molecular phenotyping task that may fall under the regulatory definition of a companion diagnostic. Coordination between the U.S. Food and Drug Administration and equivalent international bodies will be necessary to establish an appropriate regulatory pathway that accounts for the dynamic and updatable nature of machine learning models with biological underpinnings.

Robustness and security concerns intersect with policy at the point of adversarial vulnerability. Medical machine learning models have been shown to be susceptible to adversarial attacks that can alter predictions through imperceptible perturbations of input images or structured molecular data [21]. In a cross-modal setting, an attacker could potentially manipulate either the imaging pathway, for instance by introducing adversarial noise into the DICOM files, or the transcriptomic pathway, by altering the reported expression values in the electronic health record. Such attacks could systematically downgrade the risk score of malignant nodules, leading to missed cancers, or inflate the risk of benign nodules, causing unnecessary invasive procedures. Security policies must therefore mandate secure data pipelines with integrity verification, anomaly detection for input data, and robust training methods such as adversarial training and certified defenses. The policy landscape must also address the unique privacy challenges of linking imaging data, which may contain facial features reconstructable from CT scouts, with genomic data, which is inherently identifying. The coexistence of these sensitive data streams under a single analytical platform heightens the need for rigorous privacy-preserving techniques, including federated learning and differential privacy, and for governance structures that give patients meaningful control over how their combined radiomic and molecular profiles are used for model training and ongoing learning [23].

Finally, the biological mechanism of transcriptional phase separation itself has policy implications for model transparency. The incorporation of condensate biology moves the system beyond pure pattern recognition toward a mechanistically informed predictor, which some regulators and professional societies consider a higher level of evidence. Demonstrating that the model's cross-modal attention aligns with known biology—for example, that image textures associated with hypoxia are attended when a HIF-1-alpha-related condensate signature is active—provides a degree of mechanistic corroboration that strengthens the case for clinical adoption [22]. This shifts the governance conversation from "can we trust the black box" to "how can we critically verify that the model's reasoning aligns with pathophysiological knowledge," a more mature and scientifically grounded stance.

8. Conclusion

This paper has presented a systems-level examination of cross-modal attention fusion of radiomic lung nodule features and transcriptomic phase-separation signatures for explainable cancer risk stratification. We argued that the choice of a decoupled attention architecture is not merely a technical optimization but a foundational design decision with far-reaching consequences for explainability, regulatory approval, modular maintenance, and equitable deployment. Radiomics provides a noninvasive window into tissue architecture, while transcriptomic condensate signatures capture the dynamic biophysical underpinnings of oncogenic transcription, and attention mechanisms that keep modality-specific extraction separate from cross-modal interaction afford the transparency necessary for clinical and regulatory trust. The analysis traversed the infrastructure demands of harmonizing cross-institutional imaging and sequencing data, the importance of adversarial robustness and fairness auditing, and the governance frameworks that must evolve to accommodate AI systems that integrate molecular mechanistic insights. By treating the fusion pipeline as a socio-technical infrastructure rather than an isolated algorithm, we underscored that sustainable deployment requires alignment between architectural transparency, clinical workflow integration, and policy safeguards that protect patients while enabling innovation. Future work must focus on prospective validation in diverse populations, development of standardized phase-separation transcriptomic assays, and collaborative governance models that bring together radiologists, pathologists, oncologists, computer scientists, bioethicists, and regulators to guide the responsible translation of cross-modal attention systems into routine cancer care.

References

1. Gillies, R. J., Kinahan, P. E., & Hricak, H. (2016). Radiomics: Images are more than pictures, they are data. Radiology, 278(2), 563–577. https://doi.org/10.1148/radiol.2015151169

2. Lambin, P., Rios-Velazquez, E., Leijenaar, R., Carvalho, S., van Stiphout, R. G. P. M., Granton, P. V., … & Aerts, H. J. W. L. (2012). Radiomics: Extracting more information from medical images using advanced feature analysis. European Journal of Cancer, 48(4), 441–446. https://doi.org/10.1016/j.ejca.2011.11.036

3. Aerts, H. J. W. L., Velazquez, E. R., Leijenaar, R. T. H., Parmar, C., Grossmann, P., Carvalho, S., … & Lambin, P. (2014). Decoding tumour phenotype by noninvasive imaging using a quantitative radiomics approach. Nature Communications, 5, 4006. https://doi.org/10.1038/ncomms5006

4. Chang, C., Fu, M., Chen, X., et al. (2025, November). Research on PDU-Net Lung Nodule Segmentation Algorithm Based on Path Aggregation and Dual Attention. In 2025 4th International Conference on Image Processing, Computer Vision and Machine Learning (ICICML) (pp. 1897-1900). IEEE.

5. Setio, A. A. A., Traverso, A., de Bel, T., Berens, M. S. N., van den Bogaard, C., Cerello, P., … & van Ginneken, B. (2017). Pulmonary nodule detection in CT images: false positive reduction using multi-view convolutional networks. IEEE Transactions on Medical Imaging, 36(5), 1160–1169. https://doi.org/10.1109/TMI.2017.2650162

6. Ardila, D., Kiraly, A. P., Bharadwaj, S., Choi, B., Reicher, J. J., Peng, L., … & Shetty, S. (2019). End-to-end lung cancer screening with three-dimensional deep learning on low-dose chest computed tomography. Nature Medicine, 25(6), 954–961. https://doi.org/10.1038/s41591-019-0447-x

7. Vaswani, A., Shazeer, N., Parmar, N., Uszkoreit, J., Jones, L., Gomez, A. N., … & Polosukhin, I. (2017). Attention is all you need. In Advances in Neural Information Processing Systems 30 (pp. 5998–6008).

8. Baltrusaitis, T., Ahuja, C., & Morency, L.-P. (2019). Multimodal machine learning: A survey and taxonomy. IEEE Transactions on Pattern Analysis and Machine Intelligence, 41(2), 423–443. https://doi.org/10.1109/TPAMI.2018.2798607

9. Shin, Y., & Brangwynne, C. P. (2017). Liquid phase condensation in cell physiology and disease. Science, 357(6357), eaaf4382. https://doi.org/10.1126/science.aaf4382

10. Boeynaems, S., Alberti, S., Fawzi, N. L., Mittag, T., Polymenidou, M., Rousseau, F., … & Fuxreiter, M. (2018). Protein phase separation: A new phase in cell biology. Trends in Cell Biology, 28(6), 420–435. https://doi.org/10.1016/j.tcb.2018.02.004

11. Chung, C.-I., et al. (2024). Phase separation of YAP-MAML2 differentially regulates the transcriptome. Proceedings of the National Academy of Sciences, 121, e2310430121.

12. Cancer Genome Atlas Research Network. (2014). Comprehensive molecular profiling of lung adenocarcinoma. Nature, 511(7511), 543–550. https://doi.org/10.1038/nature13385

13. Gentles, A. J., Bratman, S. V., Lee, L. J., et al. (2015). Integrating tumor and stromal gene expression signatures with clinical indices for survival stratification in lung cancer. Journal of the National Cancer Institute, 107(10), djv168. https://doi.org/10.1093/jnci/djv168

14. Lundberg, S. M., & Lee, S.-I. (2017). A unified approach to interpreting model predictions. In Advances in Neural Information Processing Systems 30 (pp. 4765–4774).

15. Ribeiro, M. T., Singh, S., & Guestrin, C. (2016). "Why should I trust you?": Explaining the predictions of any classifier. In Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (pp. 1135–1144). https://doi.org/10.1145/2939672.2939778

16. Ding, Y., Wang, X., Yuan, H., Qu, M., & Jian, X. (2025). Decoupling feature-driven and multimodal fusion attention for clothing-changing person re-identification. Artificial Intelligence Review, 58(8), 241.

17. Obermeyer, Z., Powers, B., Vogeli, C., & Mullainathan, S. (2019). Dissecting racial bias in an algorithm used to manage the health of populations. Science, 366(6464), 447–453. https://doi.org/10.1126/science.aax2342

18. Mittelstadt, B. D., Allo, P., Taddeo, M., Wachter, S., & Floridi, L. (2016). The ethics of algorithms: Mapping the debate. Big Data & Society, 3(2), 2053951716679679. https://doi.org/10.1177/2053951716679679

19. Esteva, A., Kuprel, B., Novoa, R. A., Ko, J., Swetter, S. M., Blau, H. M., & Thrun, S. (2017). Dermatologist-level classification of skin cancer with deep neural networks. Nature, 542(7639), 115–118. https://doi.org/10.1038/nature21056

20. Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., & Batra, D. (2017). Grad-CAM: Visual explanations from deep networks via gradient-based localization. In Proceedings of the IEEE International Conference on Computer Vision (pp. 618–626). https://doi.org/10.1109/ICCV.2017.74

21. Finlayson, S. G., Bowers, J. D., Ito, J., Zittrain, J. L., Beam, A. L., & Kohane, I. S. (2019). Adversarial attacks on medical machine learning. Science, 363(6433), 1287–1289. https://doi.org/10.1126/science.aaw4399

22. Chung, C.-I. et al. Phase separation of YAP-MAML2 differentially regulates the transcriptome. Proc. Natl Acad. Sci. Usa. 121, e2310430121 (2024).

23. Hnisz, D., Shrinivas, K., Young, R. A., Chakraborty, A. K., & Sharp, P. A. (2017). A phase separation model for transcriptional control. Cell, 169(1), 13–23. https://doi.org/10.1016/j.cell.2017.02.007

24. Price, W. N., & Cohen, I. G. (2019). Privacy in the age of medical big data. Nature Medicine, 25(1), 37–43. https://doi.org/10.1038/s41591-018-0272-7