Proceedings of CONF-MLA 2026 Symposium: Explainable Computing, Modeling & Data Science in Complex Systems
DOI: 10.54254/2755-2721/2026.GF35574

The Role of Artificial Intelligence in Modern Biomedical Research: From Data to Drugs

Shuyue Yang

Wuhan Cogdel Cranleigh High School, Wuhan, China
yangshuyue09@126.com

Abstract. The ability to efficiently analyze large-scale and complex biological data helps artificial intelligence to rapidly revolutionize biomedical research. This paper summarizes the value of artificial intelligence in three biomedical domains: multi-omics data integration for disease prediction, medical image analysis for clinical diagnose and drug discovery. In multi-omics research, artificial intelligence optimizes integration outcomes of heterogeneous biological datasets, deepening cognition of disease mechanisms and polishing predictive models. In medical imaging, deep learning performs excellently in various diagnosis tasks, yet the lack of interpretability brings out limitations. To address the problem, explainable artificial intelligence was generated to effectively increase model transparency and clinical validation. In drug discovery, artificial intelligence significantly accelerates main processes, including virtual screening, drug repurposing and protein structure prediction. In particular, AlphaFold is one of the most paramount breakthroughs. Despite these advancements, data heterogeneity, lack of clinical validation samples and model interpretability shortage restrain the practical usage of relevant techniques. This paper sorts out the potential and limitations of artificial intelligence in biomedical research, pointing out that future research should build more stable and interpretable models.

Keywords: Artificial intelligence, clinical diagnose, drug discovery

1. Introduction

Recently, rapid development of biomedical technologies prompts plenty of unprecedented complicated biological data, including multi-omics profiles, medical imaging and molecular structure [1, 2]. Artificial intelligence becomes a powerful tool to analyze and interpret this kind of high-dimensional and heterogeneous biologic data [3]. Artificial intelligence (AI) mainly applies to three directions within biomedical field: integrating multi-omics data to predict diseases, interpreting medical image to help clinical diagnosis and contributing to drug discovery to boost therapeutic regimen [1, 3]. This paper explores how artificial intelligence propel biomedical research to achieve the transformation from data driven analysis to practical drug development, in the meanwhile, revealing the main drawback about this technique in data integration, model interpretability and experiment verification. Although artificial intelligence develops significantly, issues still remain in the ensurance of model interpretability, data quality and clinical reliability [4-6].

DOI: 10.54254/2755-2721/2026.GF35574

2. Artificial intelligence for multi-omics data integration in disease prediction

Multi-omics means systematically integrating genomics, transcriptomics, proteomics, metabolomics and other multiple-layer biological data in order to more comprehensively parse biological systems and mechanisms of disease action. Multi-omics is increasingly considered a necessary research method in precision medicine since relying on a single type of molecular data cannot completely explain biological processes. Genomics offers information about genetic variation and DNA mutations: transcriptomics reveals gene expression patterns, proteomics uncovers protein abundance and interactions, and metabolomics captures metabolic activities within tissues and cells. Each omics layer interprets disease mechanisms from different perspectives. The process of disease progression can be understood comprehensively, and biomarkers that cannot be found by single omics can be identified after integrating these complementary datasets [7]. Compared with single omics methods, multi-omics provides more abundant biological information, yet high dimensionality, heterogeneity and noise bring enormous challenges to researchers at the same time [1]. For the sake of these problems, artificial intelligence techniques, such as machine learning and deep learning have been widely applied to integrate multi-omics data to improve the effectiveness of disease prediction. These methods have the ability to dig mutual and complementary characteristic patterns between different omics layers, then build more precise models for complex biological processes. Besides, representation learning and data fusion techniques have become the primary tool to manage multimodal biological data [8]. Recently, advancements in deep learning promote modeling nonlinear relationships in various omics datasets. Traditional statistic methods always need manual feature selection; however, deep learning models can automatically learn hierarchical representations from heterogeneous biological data. These abilities optimize disease categorization, patient stratification and biomarker identification, allowing artificial intelligence to become an increasingly essential research tool in precision medicine [1, 8]. Related techniques gain processes, though multi-omics integration research still has notable shortage, including data quality imbalance between different platforms, data missing and data incongruity decrease model reliability. In addition, the interpretability of AI-based integration model limited its practical usage in clinical situations [8].

3. Explainable artificial intelligence in medical imaging and clinical decision support

Model interpretability is definitely essential in high-risk medical decision-making situations. Unlike recommendation systems and online advertisements, diagnostic errors in the medical field directly imperil patients' safety and therapeutic effect. Therefore, clinical medical staff always require an artificial intelligence system that provides not only precise prediction outcomes but also convincing evidence. If explanations can mark susceptible areas in medical images or identify key characters, diagnosis confidence will increase, and communication between artificial intelligence systems and medical practitioners will be promoted [4, 5]. Deep learning successfully analyzes medical images, especially remarkable in fields such as tumor detection, organ segmentation and disease categorization. Convolutional neural networks and other deep learning algorithms are capable of precisely resolving complicated image data; these techniques always perform better than conventional machine learning algorithms. These technical advancements significantly optimize the chance of practice of artificial intelligence in radiology and pathology departments [3, 9]. Current explainable artificial intelligence approaches can be categorized as two types: post-hoc explanation approaches and inherently interpretable models. Post-hoc methods generate explanations by visualization and feature attribution techniques after the model finishes prediction. However,

DOI: 10.54254/2755-2721/2026.GF35574

inherently interpretable models can output transparent decision-making procedures from the beginning. Two types of techniques all increase confidence in artificial intelligence auxiliary diagnosis, yet limitations exist in precision, accuracy and clinical practicability [4, 5, 9].

Nevertheless, deep learning models often suffer severe criticism because of the paucity of interpretability despite their excellent performance. Doctors and nurses need not only precise prediction outcomes from deep learning models but also clarification of the derivation process of conclusions in clinical practice. Many deep learning systems have black box characteristics, which significantly decrease the reliability of the model and hamper the practical usage in hospitals [9]. In order to address the problem, researchers invented explainable artificial intelligence methods to parse the decision processes of the model. Techniques such as saliency maps, feature attribution and attention mechanisms allow researchers to navigate the area in the image that has the most important influence on the prediction outcomes. This kind of method aims to minimize the huge gap between high-performance artificial intelligence systems and clinical interpretability requirements [4, 5]. Explainable artificial intelligence methods still face enormous challenges despite already obtaining some improvements. Part of explainable techniques' performance is unsteady in different datasets, which also leads to inconsistent outcomes; in the meanwhile, technical interpretation is always disjointed from clinical interpretation. Besides, no one knows how to judge various interpretations and whether these interpretations are valuable and feasible for doctors and nurses. To that end, boosting the reliability of explainable artificial intelligence and clinical effectiveness is one of the most popular research directions [5, 6].

4. Artificial intelligence in drug discovery from protein structure prediction to virtual screening

Traditional drug discovery, including target recognition, compound screening, preclinical experiments and clinical trials, entails a time-consuming and costly process. This research and development pipeline always consumes more than ten years and an enormous sum of funding; in the meantime, the failure rate remains definitely high because of low efficiency in filtering out effective candidate drugs. Traditional research methods are limited to highly depending on trial-and-error experiments, which not only consume plenty of resources but also respond slowly when facing new diseases [2, 3]. Besides, conventional drug discovery methods have trouble dealing with complex characteristics of biological systems, especially in diseases that are triggered by multiple factors, such as cancer. The paucity of integration between different biological information leads to imprecise prediction of a drug's efficacy and safety in early research [3].

Artificial intelligence significantly revolutionizes the procedure of drug discovery, which allows to predict molecular characteristics, drug-target interactions and biological activity based on data. Compared to traditional methods, machine learning models are able to analyze large-scale chemical and biological datasets and efficiently filter out potential candidate drugs. These methods shorten the early stage of the research and development cycle of drugs while increasing the accuracy of virtual compound screening [2, 3].

The latest development in the deep learning field strengthens the ability of artificial intelligence systems in modeling complicated molecular structures and interactions. To be specific, generative models and predictive algorithms have been largely applied to accelerate drug design and new usage for old drugs, offering an alternative plan that is more systematized and scalable than conventional experimental screening.

AlphaFold, a system based on deep learning and precise prediction of protein structure, is a paramount breakthrough in this field. AlphaFold can predict the three-dimensional structure of a

DOI: 10.54254/2755-2721/2026.GF35574

protein only based on the amino acid sequence, which shows up as unprecedented excellent performance and tackle an unsolved difficulty in structural biology. This technical advancement has profound meaning for drug discovery since analyzing protein structure is a prerequisite of the identification of drug binding sites and the design of effective drug molecules. AlphaFold significantly improves accuracy and calculation speed of protein structure, resolving the main problem in drug discovery procedures based on structure. Relying on this technology, researchers are able to obtain protein structure information that is unpretertable before, greatly widening the potential target range applicable to drug development research [10].

Excluding structure prediction, artificial intelligence also propels virtual screening and compound optimizing. Virtual screening methods based on machine learning techniques judge large-scale chemical molecules to filter out a molecule with high binding affinity to the target protein. Compared with traditional high-throughput screening, virtual screening based on artificial intelligence accelerates speed and decreases costs remarkably, contributing to a smaller range of candidate compounds at the initial stage of research [2, 3]. Besides, a deep learning model helps analyze potential therapeutic uses of existing drugs. Rapid identification of feasible therapy is paramount in emergency situations, including infectious disease. The drug repurposing technique driven by artificial intelligence can accelerate the valuable drug discovery procedure by scale [2].

Although drug discovery powered by artificial intelligence gained many breakthroughs, limitations still remain. One of the most difficult problems is the paucity of high-quality and standerdized biological data, which declines the stability and generalizability of the model. Plenty of artificial intelligence models belong to the black box model, which has difficulties in explaining prediction outcomes properly, hampering practical application in clinical situations and regulatory systems [2, 6]. In the meantime, AlphaFold cannot fully revivify dynamic biological processes such as protein interactions, despite AlphaFold significantly improving the prediction level of protein structure. Therefore, the experimental verification phase is indispensable. In other words, artificial intelligence can only play the role of assistant rather than fully replace traditional methods.

5. Conclusion

In conclusion, artificial intelligence plays an important role in research, promoting biomedical research to transform from data-intense analysis to practical drug discovery. In multi-omics data integration, biomedical image analysis and drug discovery fields, artificial intelligence shows excellent performance to deal with complex and high-dimensional biological data and boost efficiency in biomedical research [1, 3]. However, despite these developments, troubles still remain. Model interpretability, data heterogeneity and clinical validation continuously restrain large-scale practical usage of artificial intelligence in real-world medical systems. Individuals are concerned about medical decision transparency and reliability, especially because of the black box nature [4, 6]. In order to achieve deep integration between biological data and computational techniques, future research can focus on more explainable, stable and clinically validated artificial intelligence systems. Advancements in explainable artificial intelligence and multimodal learning may break current limitations to comprehensively boost the reliability of biomedical applications [5, 8]. After solving these issues, artificial intelligence may break barriers between massive biological data and feasible medical insights, finally accelerating research in safer and more effective therapeutic regimens.

DOI: 10.54254/2755-2721/2026.GF35574

References

[1] Liu, Y., Zhu, K., Peng, W., Liu, Z., & Mao, X. (2026). Multi-omics and artificial intelligence for precision drug discovery and potential clinical applications. *Signal Transduction and Targeted Therapy, 11*(1), 210.
[2] Askr, H., Elgeldawi, E., Aboul Ella, H., Elshaier, Y. A., Gomaa, M. M., & Hassanien, A. E. (2023). Deep learning in drug discovery: An integrative review and future challenges. *Artificial Intelligence Review, 56*(7), 5975-6037.
[3] Vamathevan, J., Clark, D., Czodrowski, P., Dunham, I., Ferran, E., Lee, G., ... & Zhao, S. (2019). Applications of machine learning in drug discovery and development. *Nature Reviews Drug Discovery, 18*(6), 463-477.
[4] Prentzas, N., Kakas, A., & Pattichis, C. S. (2023). Explainable AI applications in the medical domain: A systematic review. *arXiv preprint arXiv:2308.05411.*
[5] Hou, J., Liu, S., Bie, Y., Wang, H., Tan, A., Luo, L., & Chen, H. (2024). Self-explainable AI for medical image analysis: A survey and new outlooks. *arXiv preprint arXiv:2410.02331.*
[6] Jiménez-Luna, J., Grisoni, F., & Schneider, G. (2020). Drug discovery with explainable artificial intelligence. *Nature Machine Intelligence, 2*(10), 573–584.
[7] Liu, F., Beck, S., Yang, L., Luo, H., & Zhang, K. (2026). Advancing AI for multi-omics and clinical data integration in basic and translational cancer research. *Nature Reviews Cancer*, 1–16.
[8] Hussein, A., Prasad, M., & Braytee, A. (2024). Explainable AI methods for multi-omics analysis: A survey. *arXiv preprint arXiv:2410.11910*.
[9] Jin, W., Li, X., & Hamarneh, G. (2022). Evaluating explainable AI on a multi-modal medical imaging task: Can existing algorithms fulfill clinical requirements? In *Proceedings of the AAAI Conference on Artificial Intelligence, 36*(11), 11945–11953.
[10] Jumper, J., Evans, R., Pritzel, A., Green, T., Figurnov, M., Ronneberger, O., ... & Hassabis, D. (2021). Highly accurate protein structure prediction with AlphaFold. *Nature, 596*(7873), 583-589.