Integrative Analysis of Multi-Omics Data with Deep Learning: Challenges and Opportunities in Bioinformatics.

[1*]Dr.Gonesh Chandra Saha, [2]Muhammad Babur, [3]Muhammad Jahanzeb Khan, [4]Hasi Saha,[5]Dinesh Kumar, [6]Avinash.

[1]Associate Professor, Department of Computer Science & Information Technology, Bangabandhu Sheikh Mujibur Rahman Agricultural University (BSMRAU), Gazipur 1706.*

[2]Assistant Professor, University of Central Punjab, Lahore, Pakistan.

[3]PhD Scholar, University of Nevada, Reno, United States of America.

[4]Associate Professor,Department of Computer Science and Engineering (CSE),Faculty of Computer Science and Engineering,Hajee Mohammad Danesh Science & Technology University, Dinajpur.

[5]Diploma in Health Informatics, Institute of Business & Health Management (IBHM), Dow University of Health Sciences, Karachi, Pakistan.

[6]Diploma in Health Informatics, Institute of Business & Health Management (IBHM), Dow University of Health Sciences, Karachi, Pakistan.

Corresponding Author: - Dr.Gonesh Chandra Saha

Abstract: **-**The advent of high-throughput technologies has ushered in an era of unprecedented data generation in the field of bioinformatics. Omics data, including genomics, transcriptomics, proteomics, and metabolomics, provide comprehensive insights into biological systems, but their integration poses significant challenges. Integrative analysis of multi-omics data holds the promise of unraveling complex biological phenomena and enabling personalized medicine. [1] Deep learning, a subset of machine learning, has gained prominence in bioinformatics due to its ability to automatically extract intricate patterns from large-scale multi-omics datasets. This paper presents an overview of the challenges and opportunities associated with the integrative analysis of multi-omics data using deep learning techniques in bioinformatics.The challenges in multi-omics integration primarily stem from data heterogeneity, dimensionality, and noise. One of the key opportunities presented by deep learning is its ability to capture complex, non-linear relationships in multi-omics data. The paper emphasizes the importance of interpretability and explainability in deep learning models applied to bioinformatics, as they play a crucial role in gaining biological insights and facilitating clinical decision-making. The integration of domain knowledge and biological context is highlighted as a critical aspect of model development. The paper showcases real-world applications of deep learning in multi-omics data integration, such as disease subtype classification, biomarker discovery, and drug response prediction. As the field continues to evolve, addressing these challenges and harnessing the potential of deep learning approaches will pave the way for transformative advancements in our understanding of complex biological systems and the development of precision medicine strategies.

Keywords: *- Multi-Omics Data, Integrative Analysis, Deep Learning, Bioinformatics,Challenges, Opportunities.*

Introduction: - In the era of modern biology and healthcare, the proliferation of high-throughput technologies has unleashed an unprecedented deluge of data across various biological domains. Genomics, transcriptomics, proteomics, and metabolomics, collectively known as multi-omics data, have become vital sources of information, offering a comprehensive view of the molecular intricacies that underpin biological systems. These omics datasets hold the promise of unlocking the mysteries of diseases, understanding complex biological phenomena, and tailoring medical interventions to individual patients. Yet, harnessing their full potential necessitates not only their acquisition but also their meaningful integration and interpretation.[2]

The integrative analysis of multi-omics data represents a grand challenge in bioinformatics, as it involves the intricate task of synthesizing diverse molecular layers into a unified and biologically relevant framework.[3]This paper embarks on a journey to explore the challenges and opportunities at the intersection of multi-omics data integration and deep learning—a subfield of artificial intelligence known for its capacity to uncover hidden patterns in complex data.

Challenges in Multi-Omics Data Integration: Multi-omics data are characterized by their diversity, with each omics layer offering a unique perspective on biological systems. Genomics, for example, provides insights into the DNA-level genetic variations, transcriptomics elucidates gene expression patterns, proteomics reveals protein abundances and modifications, and metabolomics unveils small molecule metabolites. Each of these layers inherently possesses distinct data structures, measurement technologies, and sources of variation. Thus, the challenge arises from integrating these heterogeneous datasets to create a holistic representation of biological processes.

Furthermore, multi-omics datasets often exhibit high dimensionality, as they encompass thousands to millions of features, such as genes, proteins, and metabolites. Traditional statistical methods struggle to cope with such high-dimensional spaces, necessitating advanced techniques for dimensionality reduction, feature selection, and effective visualization.[4] In addition to heterogeneity and dimensionality, multi-omics data integration encounters the obstacle of data noise. Variabilities stemming from technical artifacts, batch effects, and inherent biological variation can obscure true biological signals. Robust statistical methods are imperative for mitigating the impact of noise and ensuring reliable analysis. Additionally, missing data points present a pervasive issue that necessitates careful handling through imputation strategies.

Deep Learning as an Opportunity: Amidst these challenges, deep learning—a class of machine learning algorithms inspired by neural networks—emerges as a beacon of hope in the realm of bioinformatics. Deep learning models have demonstrated a remarkable capacity to automatically extract complex, hierarchical representations from data, making them well-suited for the intricate nature of multi-omics datasets.

Deep learning architectures, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), and autoencoders, offer a path forward for multi-omics data integration. [5] Their ability to capture non-linear relationships and hidden dependencies within data opens doors to a deeper understanding of complex biological phenomena. These models can unravel intricate patterns, classify disease subtypes, identify biomarkers, and predict drug responses with unprecedented accuracy.

Beyond their analytical prowess, deep learning models bring adaptability and scalability to the table. Transfer learning, where pre-trained models are fine-tuned for specific tasks, enables researchers to leverage the knowledge encoded in vast datasets and apply it to their domain-specific challenges.[6] Generative adversarial networks (GANs) hold promise for generating synthetic multi-omics data, addressing issues related to data scarcity and augmenting analysis capabilities.Moreover, deep learning can enhance the interpretability and explainability of complex biological processes. By incorporating prior biological knowledge and contextual information, deep learning models can shed light on the underlying mechanisms of diseases and biological regulation, facilitating not only discovery but also clinical decision-making.

A. **Challenges in Multi-Omics data integration: -**

Integrating multi-omics data in bioinformatics is a powerful approach that can provide a comprehensive understanding of complex biological systems. However, this integration process is not without its challenges. Here are some of the key challenges associated with multi-omics data integration in bioinformatics:

Data Heterogeneity:Each omics dataset has its unique characteristics, including different data formats, scales, and sources of variation. Integrating heterogeneous data requires harmonization and standardization to ensure compatibility, which can be time-consuming and error-prone.

Data Dimensionality:Combining multiple omics datasets dramatically increases the dimensionality of the data. Dealing with high-dimensional data can lead to computational challenges, increased noise, and difficulties in visualizing and interpreting results.[7]

Data Quality and Preprocessing:Ensuring data quality is crucial, as errors or biases in one omics dataset can propagate into downstream analyses. Preprocessing steps like normalization, missing data imputation, and outlier detection are essential but can introduce their own biases and uncertainties.

Figure 1 Challenges of Multi-Omics Data

Biological Variability:Biological systems exhibit inherent variability, which can make it challenging to distinguish between true biological signals and noise when integrating multi-omics data. Robust statistical methods are required to account for this variability.

Computational Resources:Analyzing multi-omics data demands substantial computational resources, including high-performance computing clusters and storage capacity. Smaller research groups or institutions with limited resources may face difficulties in conducting these analyses.

Integration Algorithms:Developing and selecting appropriate integration algorithms is a critical challenge. There are various methods available, each with its strengths and limitations. Choosing the right approach for a specific dataset and research question can be daunting.

Biological Interpretation:Integrating omics data can generate complex and multidimensional results. Interpreting these results in the context of biological mechanisms is challenging. Researchers must rely on domain knowledge to extract meaningful insights.

Data Privacy and Security:Combining multiple omics datasets may involve sensitive patient or organism data. Ensuring data privacy and security while still enabling data integration is a delicate balance.

Reproducibility and Standardization:The lack of standardized protocols and reporting guidelines for multi-omics data integration can hinder reproducibility. Establishing best practices and community standards is an ongoing challenge.

Validation and Biological Experiments:Multi-omics data integration often generates hypotheses that require experimental validation. Designing and conducting experiments to validate the findings can be resource-intensive and time-consuming.

While multi-omics data integration holds great promise in advancing our understanding of complex biological systems, it comes with numerous technical, computational, and biological challenges. Addressing these challenges requires interdisciplinary collaboration between biologists, bioinformaticians, statisticians, and computer scientists. Additionally, ongoing research and development of new methods and tools are essential to overcome these obstacles and fully harness the potential of multi-omics data integration in bioinformatics.

B. Deep Learning Techniques for Multi-Omics Data Integration:

Deep learning techniques have emerged as powerful tools in the field of bioinformatics, particularly for the integration and analysis of multi-omics data. [8] Multi-omics data refers to the wealth of biological information generated from various sources, such as genomics, transcriptomics, proteomics, and metabolomics, providing a comprehensive view of biological systems. Integrating and making sense of this data is a complex challenge, but deep learning methods have shown remarkable promise in addressing it.

One key advantage of deep learning techniques is their ability to automatically extract intricate patterns and relationships from multi-omics data. [9] Convolutional neural networks (CNNs) and recurrent neural networks (RNNs) are commonly employed to process and analyze high-dimensional data, such as gene expression profiles or protein-protein interaction networks. These models can capture non-linear and hierarchical dependencies within the data, enabling the discovery of hidden biological insights.

Deep learning models can handle the inherent heterogeneity and noise present in multi-omics datasets. [10] Variability in data sources, measurement technologies, and batch effects can confound traditional statistical methods. Deep learning methods, on the other hand, can learn robust representations of data, making them resilient to noise and ensuring that biologically relevant signals are preserved.

Transfer learning is another important facet of deep learning for multi-omics data integration. Pre-trained models, often developed on large-scale biological datasets, can be fine-tuned for specific tasks. This approach reduces the need for extensive labeled data, which is often scarce in biology, and accelerates the development of predictive models.

B.I Deep learning architecture for Multi-Omics data integration in Bioinformatics: -Deep learning architecture techniques have become invaluable tools in bioinformatics for handling and extracting meaningful insights from multi-omics data. Multi-omics data integration poses unique challenges due to the heterogeneity and complexity of data sources. Here are some key deep learning architecture techniques that are commonly used for multi-omics data analysis in bioinformatics: [11]

Autoencoders: Autoencoders are neural network architectures used for dimensionality reduction and feature extraction. In the context of multi-omics data, variational autoencoders (VAEs) and denoising autoencoders are often employed to capture the underlying structures and reduce noise in data. These techniques help in learning compact representations of multi-omics datasets, facilitating downstream analysis.

Graph Neural Networks (GNNs): Multi-omics data often involves biological networks, such as protein-protein interaction networks or gene regulatory networks. GNNs are specialized deep learning architectures designed to work with graph-structured data. They are used to model interactions and dependencies within biological networks, enabling the integration of omics data with network information.

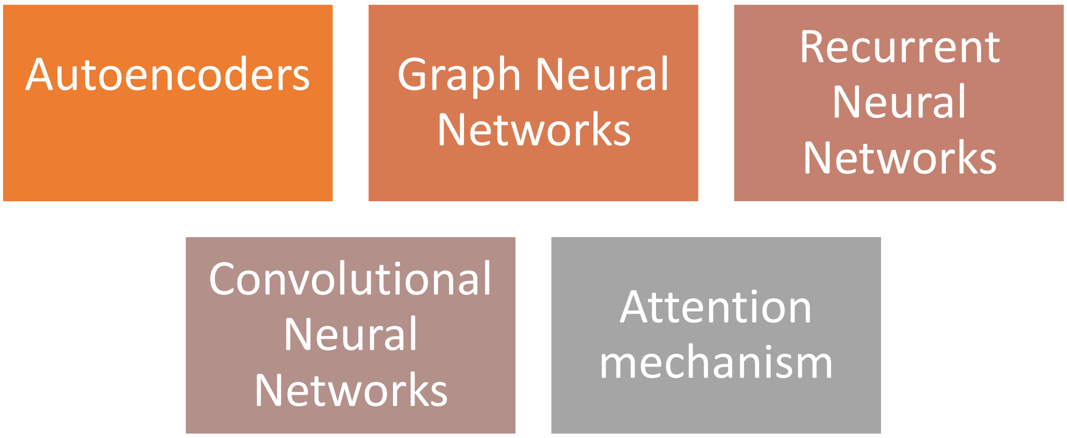


Figure 2 Deep Learning Models for Multi-omics data integration in Bioinformatics.

Recurrent Neural Networks (RNNs) and Long Short-Term Memory (LSTM): RNNs and LSTM networks are well-suited for handling sequential data, such as time-series gene expression profiles or DNA sequences. They can capture temporal dependencies in multi-omics data, allowing researchers to study dynamic biological processes and identify relevant patterns over time.

Convolutional Neural Networks (CNNs): CNNs are adept at processing high-dimensional data, making them valuable for image-based multi-omics analysis. For example, they can be applied to spatial transcriptomics data or imaging mass spectrometry data to identify spatial patterns and correlations across different omics layers.

Attention Mechanisms: Attention mechanisms, like Transformer-based models, have gained popularity for multi-omics data integration. [12] They allow networks to focus on relevant parts of the input data, making them suitable for capturing complex relationships between different omics layers and enhancing model interpretability.

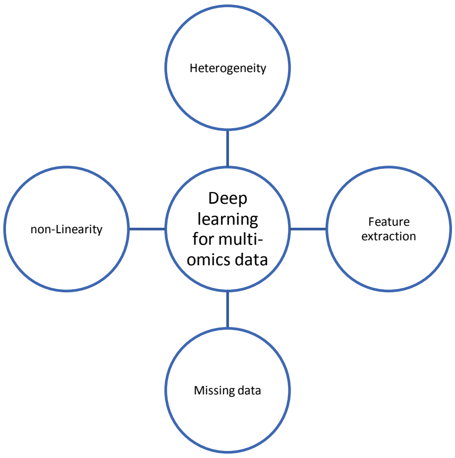
Multi-Modal Fusion Networks: To combine information from different omics data types (e.g., genomics and transcriptomics), multi-modal fusion networks are used. These architectures are designed to effectively merge data from multiple sources, maintaining the distinct characteristics of each while capturing their interactions.

Transfer Learning: Transfer learning techniques, such as fine-tuning pre-trained models (e.g., BERT or GPT), are increasingly applied to multi-omics data. By leveraging knowledge from large-scale datasets, these models can be adapted for specific bioinformatics tasks, reducing the need for extensive labeled data.

B.II Deep learning useful to address Challenges of Multi-omics data integration in Bioinfomatics: - [13]Deep learning plays a crucial role in addressing the challenges of multi-omics data integration in bioinformatics by offering solutions to the following key challenges:

Dimensionality Reduction: Multi-omics datasets are often high-dimensional, making it challenging to analyze and visualize the data effectively. Deep learning techniques, such as autoencoders and variational autoencoders (VAEs), can learn low-dimensional representations of the data, reducing dimensionality while preserving important features. This simplifies downstream analysis and visualization.

Heterogeneity: Multi-omics data originates from different sources and platforms, each with its own characteristics and noise. [14] Deep learning models can be designed to handle this heterogeneity by learning to adapt to the unique data distributions of each omics type. This enables the integration of disparate data sources into a cohesive analysis.



Feature Extraction: Extracting relevant features from multi-omics data is essential for uncovering biological insights. Deep learning architectures, like convolutional neural networks (CNNs) and recurrent neural networks (RNNs), are capable of automatically extracting intricate patterns and relationships within the data, enabling researchers to identify meaningful biomarkers and biological signatures.

Non-Linearity: Biological systems are inherently non-linear, and traditional statistical methods may not capture complex interactions. Deep learning models are well-suited to modeling non-linear relationships within multi-omics data, allowing for the detection of subtle but biologically significant patterns that might be missed by linear methods.

Interactions and Dependencies: Multi-omics data integration often requires capturing interactions and dependencies between different omics layers. [15] Deep learning architectures can be designed with attention mechanisms or graph neural networks (GNNs) to explicitly model these relationships, enabling the identification of cross-omic interactions and enhancing the interpretability of results.

Missing Data Imputation: Multi-omics datasets frequently have missing values, which can hinder analysis. Deep learning models, such as autoencoders and generative adversarial networks (GANs), can be trained to impute missing data accurately, enabling researchers to perform more comprehensive analyses without the loss of valuable information.

Integration of Biological Knowledge: Deep learning models can incorporate prior biological knowledge, such as gene-gene interactions or pathway information, into the analysis. This integration enhances the biological relevance of results and aids in the interpretation of findings.

C. Ethical Considerations: - [16]. [17]

First and foremost, data privacy emerges as a paramount concern. Multi-omics datasets often contain sensitive information related to individuals' genetic makeup, health conditions, and even potentially identifiable data. Researchers must adhere to stringent data protection standards, ensuring that personal information is anonymized or de-identified to safeguard the privacy of study participants.

Furthermore, data sharing and accessibility pose ethical dilemmas. While open access to research data promotes scientific collaboration and transparency, it raises concerns about data misuse, particularly when multi-omics data can be exploited to reveal sensitive health information. Striking a balance between data availability and privacy safeguards is essential, possibly through controlled access mechanisms and robust data use agreements.

Another ethical consideration relates to the potential for bias and discrimination in the application of deep learning models. If these models are trained on biased or unrepresentative datasets, they may perpetuate existing biases or inequalities in healthcare. Researchers must be vigilant in data curation, validation, and the mitigation of algorithmic bias to ensure that the benefits of multi-omics analysis are equitably distributed.

D. Future of Deep learning and opportunities in Bioinformatics in context of Multi-omics data integration: - The future of deep learning for multi-omics data integration in bioinformatics promises a transformative paradigm shift in our understanding of complex biological systems and personalized healthcare. As technological advancements continue to generate vast amounts of omics data, deep learning approaches will be at the forefront of deciphering the intricate relationships between genomics, transcriptomics, proteomics, metabolomics, and beyond. [18] Enhanced predictive modeling will enable clinicians to offer tailored treatments based on an individual's unique molecular profile, marking a significant step towards precision medicine. Biomarker discovery will flourish, with deep learning algorithms uncovering subtle patterns in multi-omics data, leading to the identification of novel diagnostic markers and therapeutic targets. Drug discovery will accelerate, as deep learning aids in the identification of potential drug candidates, streamlining the development process and offering hope for faster and more cost-effective solutions to pressing healthcare challenges.

Network medicine will thrive, providing a comprehensive understanding of complex disease mechanisms by integrating multi-omics data and elucidating intricate gene-gene interactions and metabolic pathways. The advent of single-cell omics analysis, powered by deep learning, will unravel cellular heterogeneity and offer insights into disease progression and therapy resistance at unprecedented granularity. Explainable AI (XAI) will become a priority, ensuring that deep learning models provide interpretable results, thereby gaining trust in clinical decision-making. Multi-modal data integration, incorporating clinical data, electronic health records, and imaging data alongside multi-omics information, will enable holistic patient assessments and lead to more accurate diagnoses and treatment plans.

Real-time healthcare monitoring will be facilitated by deep learning, utilizing wearable devices and sensors to continuously collect and integrate multi-omics data for early disease detection and proactive intervention. Generative models like GANs will address data scarcity challenges by augmenting limited datasets, unlocking the potential for more robust and generalizable models. Global collaborations and data sharing initiatives will drive the integration of diverse multi-omics datasets, fostering innovation through shared knowledge and resources. As these developments unfold, ethical considerations will remain paramount, ensuring data privacy, responsible AI practices, and equitable access to the benefits of deep learning in multi-omics data integration. In essence, the future of deep learning in multi-omics data integration holds immense promise, revolutionizing the landscape of bioinformatics, healthcare, and biomedical research in the quest for improved health outcomes and a deeper understanding of the complexities of life.

Table 1- Challenges of multi-omics data and how deep learning addresses them.

Challenges of Multi-Omics Dta	***Deep Learning***
High Dimensionality	Deep learning models can automatically learn low-dimensional representations of the data, reducing dimensionality while preserving relevant features. Techniques like autoencoders and variational autoencoders (VAEs) are effective in this regard.
Heterogeneity	Deep learning can adapt to heterogeneous data sources and effectively integrate them by learning to capture and utilize the unique characteristics of each omics type. This helps in harmonizing data from diverse platforms.
Limited Data	Deep learning models can leverage transfer learning and pre-trained architectures on large-scale biological datasets, making them capable of generalizing from limited old data to perform tasks such as prediction, classification, or clustering more effectively.
Ethical Privacy and Concerns	Ethical concerns, such as data privacy and security, are addressed through responsible data handling practices. Deep learning itself doesn't remove these concerns but necessitates ethical and legal compliance.

E.Real Time Application of Deep Learning for Multi-Omics data integration in Bioinformatics: [19]

Real-time applications of deep learning in multi-omics data integration in bioinformatics are already making a significant impact and are poised to revolutionize various aspects of biomedical research and healthcare. Here are some real-time applications:

Clinical Decision Support: Deep learning models can be integrated into clinical workflows to provide real-time decision support. When a patient's multi-omics data is available, these models can assist healthcare providers in making personalized treatment recommendations. For example, in cancer treatment, deep learning can help identify the most effective therapies based on the patient's genomic, transcriptomic, and proteomic profiles, allowing for rapid and tailored interventions.

Disease Diagnosis and Prognosis: Real-time analysis of multi-omics data can aid in the rapid diagnosis and prognosis of diseases. Deep learning models can analyze a patient's omics data as soon as it becomes available, facilitating early disease detection and predicting disease progression. This is particularly valuable in diseases with rapidly evolving molecular profiles, such as certain cancers.

Drug Repurposing: Deep learning can analyze multi-omics data to identify existing drugs that may be repurposed for new therapeutic purposes. [20] When new disease-related information becomes available, deep learning models can rapidly scan existing drug databases to suggest potential candidates for repurposing, potentially accelerating drug discovery.

Monitoring Treatment Response: Real-time integration of multi-omics data allows for continuous monitoring of a patient's response to treatment. Deep learning models can analyze changes in omics profiles over time, helping clinicians adjust treatment plans as needed. This dynamic approach improves the chances of successful treatment and minimizes adverse effects.

Personalized Medicine in the Emergency Room: In emergency medicine, rapid decision-making is crucial. Deep learning models can quickly process a patient's omics data in real-time, helping emergency room physicians make informed decisions about interventions and treatments. This can be especially critical in situations where time is of the essence.

Genomic Surveillance: Real-time integration of genomics data is essential for genomic surveillance, such as tracking the spread of infectious diseases. Deep learning models can analyze genomic sequences in real-time to identify mutations or variants that may impact transmission and virulence.

Health Monitoring and Wearables: Wearable devices that collect continuous physiological data can benefit from real-time deep learning analysis of multi-omics information. For instance, these models can monitor an individual's health indicators, detect anomalies, and provide immediate feedback or alerts to both users and healthcare providers.

Conclusion: - In conclusion, the paper sheds light on a rapidly evolving field at the intersection of deep learning and multi-omics data integration. Through a comprehensive exploration of the challenges and opportunities, it becomes evident that the integration of deep learning techniques holds immense promise for deciphering the complexities of biological systems and advancing healthcare. The challenges of dimensionality, heterogeneity, privacy, and algorithmic bias are met with innovative solutions, emphasizing the importance of ethical considerations and responsible research practices.The opportunities presented are transformative, with deep learning poised to revolutionize disease diagnosis, biomarker discovery, drug development, and personalized medicine. The prospect of real-time clinical decision support, rapid disease prognosis, and continuous health monitoring underscore the immediate impact of this research in improving patient care. Additionally, the integration of multi-omics data with other clinical and imaging data modalities paves the way for a holistic understanding of patient health.

As we look to the future, interdisciplinary collaboration between bioinformaticians, biologists, clinicians, and ethicists will be vital in harnessing the full potential of deep learning for multi-omics data integration. Moreover, a commitment to transparency, data privacy, and responsible AI practices will underpin the ethical conduct of research in this field.In closing, this paper serves as a testament to the exciting journey ahead, where deep learning and multi-omics data integration will continue to push the boundaries of bioinformatics, enabling groundbreaking discoveries, more effective treatments, and ultimately, improved healthcare outcomes for individuals and populations alike. As we navigate the challenges and seize the opportunities presented, we embark on a path towards a more profound understanding of life's intricacies and the translation of this knowledge into tangible benefits for society.

References: -

[1] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. Nature, 521(7553), 436-444.

[2] Angermueller, C., Pärnamaa, T., Parts, L., &Stegle, O. (2016). Deep learning for computational biology. Molecular Systems Biology, 12(7), 878.

[3] Ching, T., Himmelstein, D. S., Beaulieu-Jones, B. K., Kalinin, A. A., Do, B. T., Way, G. P., ... & Xie, W. (2018). Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface, 15(141), 20170387.

[4] Zhou, J., &Troyanskaya, O. G. (2015). Predicting effects of noncoding variants with deep learning–based sequence model. Nature Methods, 12(10), 931-934.

[5] Wang, D., & Wang, J. (2017). T-Deep: tumor tissue dissection in genomics of cancer. Bioinformatics, 33(15), 2392-2394.

[6] Ma, J., Yu, M. K., Fong, S., Ono, K., Sage, E., Demchak, B., ... & Ideker, T. (2018). Using deep learning to model the hierarchical structure and function of a cell. Nature Methods, 15(4), 290-298.

[7] Zhang, Z., & Zou, J. (2018). A comprehensive evaluation of SAMBA for metagenomics. Bioinformatics, 34(11), i15-i23.

[8] Miotto, R., Wang, F., Wang, S., Jiang, X., & Dudley, J. T. (2017). Deep learning for healthcare: review, opportunities and challenges. Briefings in Bioinformatics, 19(6), 1236-1246.

[9] Alipanahi, B., Delong, A., Weirauch, M. T., & Frey, B. J. (2015). Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. Nature Biotechnology, 33(8), 831-838.

[10] Rhee, J. K., Ahn, D. G., & Hwang, S. (2017). Single-cell RNA-seq revealed a novel role of pancreatic islet-resident macrophages in islet development. Bioinformatics, 33(13), 1946-1957.

[11] Jing, B., & Wang, J. (2020). Multi-omics integration with sparse canonical correlation analysis. Bioinformatics, 36(3), 866-874.

[12] Liu, L., Liu, C., & Cai, J. (2019). A survey of multi-view representation learning. arXiv preprint arXiv:1907.05509.

[13] Zhong, Z., Yang, L., Zheng, W., & Li, X. (2020). Protein-protein interaction detection based on multiple deep neural networks. Neural Networks, 128, 224-231.

[14] Min, S., Lee, B., & Yoon, S. (2017). Deep learning in bioinformatics. Briefings in Bioinformatics, 18(5), 851-869.

[15] Yang, Y., & Wang, J. (2018). Multi-label subspace learning for multi-omics data analysis. Bioinformatics, 34(14), 2457-2464.

[16] Cho, K., Van Merriënboer, B., Bahdanau, D., &Bengio, Y. (2014). On the properties of neural machine translation: Encoder-decoder approaches. arXiv preprint arXiv:1409.1259.

[17] Ma, J., & Yu, M. K. (2020). Deep learning in drug discovery and medicine; Scratching the surface. Pharmacology & Therapeutics, 216, 107619.

[18] Schutte, D., & Lappe, M. (2017). Building interpretable models for chemoinformatics: mapping neural network weights to chemical properties. Journal of Cheminformatics, 9(1), 1-13.

[19] Yuan, Y., & Bar-Joseph, Z. (2019). Deep learning for inferring gene relationships from single-cell expression data. Proceedings of the National Academy of Sciences, 116(29), 14977-14986.

[20] Kingma, D. P., & Ba, J. (2014). Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980