GMR
The Original

An Intelligent Bioinformatics Framework Integrating Deep Neural Networks and Big Data Analytics for Comprehensive DNA–RNA Interaction and Tumor Genomic Analysis

Dr. Geetha T V [1], Omar Elkalesh [2], Ali Bostani [3], Manjula G [4], Indu Purushothaman [5], Giyosjon Ergashev Jurayevich [6], Zairullo Okboyev [7],

[1] Assistant Professor, Department of IOT-CSBS/SCSE, SRM Institute of Science and Technology, Ramapuram, Chennai. Email: geethatv.1309@gmail.com Orchid ID: 0000-0002-4809-4996
[2] College of engineering and applied science, American University of Kuwait, Kuwait. https://orcid.org/0009-0005-8365-8397 Email: oelkalesh@auk.edu.kw
[3] Associate Professor, College of Engineering and Applied Sciences, American University of Kuwait, Salmiya, Kuwait. Email: abostani@auk.edu.kw 0000-0002-7922-9857
[4] Professor, Dayananda Sagar Academy of Technology and Management, Bangalore 560 082, Karnataka, India. Email: manjula-ise@dsatm.edu.in https://orcid.org/0000-0002-3766-9213
[5] Assistant Professor, Department of Research, Meenakshi Academy of Higher Education and Research, Chennai, Tamil Nadu, India. Email: indu@maher.ac.in
[6] International Islamic Academy of Uzbekistan, 11, Abdulla Kodiri Street, 100011, Tashkent, Uzbekistan.Email: giyosjonergashev23@gmail.com
[7] Department of Medicine, Termez University of Economics and Service, Termez, Uzbekistan. Email: zairullo_oqboyev@tues.uz https://orcid.org/0009-0003-0377-3762

ABSTRACT

High-throughput sequencing technologies have grown exponentially, transforming genomics and transcriptomics, generating large amounts of multi-omics data that present great challenges in understanding the regulation interactions of complex DNA-RNA and tumour genomic heterogeneity. Traditional bioinformatics and machine-learning methods typically have a narrow range of scalability, the inability to combine heterogeneous data, and to represent the nonlinear molecular interactions between cells that regulate progression and development of cancer. On this work, we suggest an intelligent bioinformatics scheme that effectively combines deep neural networks (DNNs) with big data analytics to provide support for the extensive forecast of DNA and RNA relations and high-level genomic examination of tumours. The suggested architecture makes use of a unified component of whole-genome sequencing, RNA sequencing, epigenetic signals, and tumour mutation profiles a part of a scalable computational pipeline that relies on distributed big data platforms. Such convolutional and attention-based networks as deep learning modules are utilised to extract high-dimensional features, analyse the interaction, prioritise the biomarkers, and classify the distinct cancer subtypes with precision. The framework also includes systems level modelling to explain regulatory networks and pathways upheavals caused by mutation related to oncogenesis. The results of experimental assessments of benchmark cancer genomic datasets prove that the suggested methodology has the potential to substantially increase the accuracy of interaction prediction, improve the quality of tumour stratification, and enable efficient discovery of clinically meaningful molecular biomarkers over conventional statistical and machine-learning inferential bases. This framework is an effective and scalable solution to accuracy in oncology, molecular diagnostics, and integrative cancer genomics investigations through the joint implementation of predictive intelligence and scalable genomic data processing. The suggested approach offers a sufficiently bright future of uses in the area of personalised medicine, identification of therapeutic targets, and next-generational computational genomics.

Key words: *Deep Neural Networks, Big Data Analytics, DNA–RNA Interaction, Tumor Genomics, Bioinformatics Framework, Precision Medicine*

INTRODUCTION

The science of genomics has experienced a revolutionary occurrence in the last decade, with much of the transformation happening due to quick progress of high-throughput sequencing technologies, next-generation sequencing (NGS), as well as novel single-cell multi-omics systems. The advancements have facilitated the exploration of genetic and transcriptomic landscapes in a comprehensive capability of various organisms and disease conditions. Consequently, current biomedical studies today produce enormous amounts of genomic, epigenomic and transcriptomic information, which provide unparalleled possibilities in learning how complex molecular pathways to the regulation of genes, cellular differentiation and disease pathogenesis operate. The DNA-RNA regulatory interactions are one of these mechanisms that are essential in the regulation of transcriptional activity, chromatin architecture and genome stability, which are all related to the regulation of normal cellular functions and pathological changes caused by cancer.

DNARNA interactions are actively being considered as the major drivers of epigenetic regulation and transcriptional control, especially by the use of non-coding RNAs, RNA-directed chromatin remodelling and RNA-mediated regulation protein recruitment. These associations play a major role in either the activation or suppression of oncogenic paths and tumour suppressor networks. Nevertheless, computingally, identifying and describing the patterns of genome-wide interactions of DNA and RNA is a strong challenge because of the large dimensionality of sequencing data, numerous nonlinear relationships, and non-homogenous Tie chains.

Genomic analysis of tumours adds further complexities on a case-to-case basis because cancer is typified with high heterogeneity of mutations, structural genomic rearrangements, aberrant programmes of gene expression, and transcriptomic reprogramming dynamic processes. Huge cancer genomics projects like The Cancer Genome Atlas (TCGA) and the International Cancer Genome Consortium (ICGC) have resulted in massive datasets that require highly scalable computational systems that can be meanfully biological analyzed. Conventional statistical methods and standard machine-learning models are not always able to incorporate multi-omics data appropriately, and do not discover the nonlinear interactions existing between tumor molecular landscapes.

Deep neural networks (DNNs) as efficient computational theories have become influential in recent years, with high-level representations being extracted in such computational frameworks. Convolutional neural network, recurrent network, and attention-based models are the recurrent deep learning architectures that are particularly successful in the modelling of genomic sequences, the identification of particular biomarkers, and cancer subtype classification, including the prediction of regulatory interactions. At the same time, large-scale data analytics systems like Apache Hadoop and Apache Spark, offer distributed systems that provide scalable data storage, parallel computing, and analysis of petabyte size omics data. With the union of deep learning and big data technologies, a new prospect leads to solutions of intelligent bioinformatics of the next generation.

These challenges and opportunities inspired the present research to develop a smart bioinformatics system combining deep neural networks and big data analytics to map DNARNA interactions and tumour genomics as a whole. The designed architecture is scalable multi-omics data processing with multi-purpose deep learning modules to promote the accuracy of interaction prediction, tumour stratification, and effective discovery of clinically relevant biomarkers. With predictive intelligence and massive genomic analytical integration, this framework provides a powerful platform on precision oncology, molecular diagnostics, and integrative cancer genomics research.

RELATED WORK

DNA–RNA Interaction Studies

These interactions in DNA to RNA are being acknowledged as some of the core regulatory processes that play into transcriptional control, chromatin restructuring and genome stability. Both coding and non-coding RNA molecules

mediate these interactions and direct the actions of regulatory complexes to a certain genomic locus, in this way controlling the activation and repression of genes. Recent experimental high-throughput methods, including Chromatin Isolation by RNA Purification sequencing (ChIRP-seq), GRID-seq, and Hi-C-derived analyses, have made possible the discovery of RNA-binding chromatin interactions on a genome-wide scale, revealing more information about the nature of transcriptional regulation and chromatin epigenetics. The interpretation of these interactions has become especially important in cancer biology, in which dysregulated RNADNA binding has the role of oncogenic transformation and tumor progression (Hounye et al., 2025). Moreover, multi-omics approaches incorporating both transcriptomic and genomic datasets of interaction have proved to be a critical measure to unravel molecular biomarkers and regulatory interplay as related to susceptibility to disease and therapeutic reaction (Amin et al., 2025).

Deep Learning in Genomics

With the development of deep learning, computational genomics has greatly changed as now it has become possible to extract nonlinear features to high-dimensional biological data. Deep neural networks have proved to be impressive in tasks like variant calling, gene expression prediction, cancer subtype classification and regulatory motif discovery. Genomic sequence analysis is conducted with convolutional neural networks (CNNs) that are generalized, and recurrent neural networks (RNNs) and Transformer-based structures have enhanced tasks of training models on long-range interactions in transcriptomic data (Greener et al., 2022). According to recent studies, deep learning models do not only increase the predictive accuracy but also exhibit better interpretability owing to the attention mechanisms as well as this method is especially useful in the clinical prognosis models (Lee, 2022; Wang, 2025). Besides, systematic reviews highlight the increased use of deep learning in the analysis of biomedical and genomic signals, but also cite difficulties associated with the explainability, robustness, and generalisation (Alqudah and Moussavi, 2025).

Precision oncology has also been supported by the use of deep learning and can perform tumour genomic stratification and survival prediction, particularly with explainable machine-learning structural models (Hounye et al., 2025). These methods may present some good prospects to discovery of biomarkers, and tailored treatment plans.

Tumour genomics Big Data analytics.

Efforts by large-scale cancer genomics efforts like The Cancer Genome Atlas (TCGA) and International Cancer Genome Consortium (ICGC) have produced large-scale multi-dimensional data sets of genomic mutations, transcriptomic changes, and epigenetic changes. These large-scale datasets demand scaled computational facilities that could perform distributed storage, parallel processing and valuable data integration. Hadoop and Apache Spark frameworks and systems of big data have moved on to be indispensable in facilitating high-throughput genomic mining and real-time tumour analytics (Prasath, 2025). Moreover, the intersection of AI and big data has demonstrated high potential in biomedical AI uses, including drug discovery, optimization of clinical trials and molecular diagnostics (Wu et al., 2024).

Nevertheless, one of the key research opportunities is connecting the big data analytics with the more sophisticated deep learning frameworks, with the majority of available systems not being unified and able to conduct genomic large-scale processing and predict DNA-RNA interactions with scalability and intelligence. Another implication of privacy is in regard to sensitive patient genomic data, which impels the possibility to create privacy-preserving machine-learning schemes, including federated learning (Liu et al., 2021; Wang, 2025).

Research Gap and Motivation

Despite making significant technological advances in DNARNA interaction mapping, deep learning genomics, and scalable tumour data analytics, the existing technologies can be inaccurate and disjointed. There are multiple works on specific tasks like sequence classification or the survival prediction without any unified framework to combine the interaction discovery, tumour stratification, and scalable big data processing. Thus, clever bioinformatics structures that incorporate deep neural networks with the big data structures are urgently required to obtain thorough

DNARNA control modelling and clinically significant tumour genomic understanding. This study fills this knowledge gap by implementing an AI-based scalable architecture of integrative cancer genomics and precision applications of oncology.

METHODOLOGY

Feature Engineering

The idea of feature engineering is an important part of modelling genomic and transcriptomic data since biological data is heterogeneous and noisy and is high-dimensional by definition. Within the proposed intelligent bioinformatics framework, valuable representations are derived using raw sequencing data and thus learn the DNA RNA interaction patterns and tumour-specific genomic features. The feature engineering module converts nucleotide sequences, epigenetic, RNA structure data, and mutation profiles into numerical vectors, which are easily processed using deep neural networks.

k-mer Embeddings

Genomic sequences are inputted into k-mer embeddings, in which sequences of DNA or RNA are broken down into DNA or RNA sub-sequences of k length. These k-mers are able to capture local nucleotide code, regulatory codes and functional sequence codes. Embedding representations Depth learning models can learn biologically relevant features not just simple one-hot encodings by embedding representations, leading to better interaction prediction and motif discovery.

Accessibility Signals of Chromatin.

The availability of chromatin serves as a crucial context regarding DNARNA interactions with regulatory binding events highly dependent upon the open or closed chromatin state. Epigenetic features are provided by signals obtained using assays like ATAC-seq and DNase-seq. These accessibility profiles bring out active regulatory regions, enhancers, and promoters, subsequently improving the informational nature of interaction mapping and tumour regulatory network analysis.

RNA Structural Motifs

RNA molecules have both intricate secondary and tertiary structure which greatly influences their binding affinity and practical interactions with genomic loci. Hairpins, loops and bulges are structural motifs that have been extracted via computational RNA folding predictions. Writing in RNA structural descriptors allows the framework to capture the RNA-guided chromatin targeting processes in a more comprehensive way, especially in non-coding RNAs, which are part of transcriptional regulation.

Mutation Frequency Vectors

The genomes of tumors have heterogeneous somatic mutations, variations in the copy number, and structural rearrangements. In order to encode these changes, mutation frequency vectors are designed on the basis of variant occurrences between genomic regions or sets of genes. These dimensions aid in tumour subtype classification and pathway interpretation guided by mutations and biomarker identification via association between genomic instability and regulatory disruption. The general idea of the integration pipeline of multi-modal features is demonstrated in figure 1, showing how mutational data is integrated with sequence, epigenetic, and structural signals to create a single representation of deep learning models.

Altogether, a combination of sequence-based, epigenetic, structural, and mutational features allows the wholesome multi-modal depiction of the tumour genomic landscapes. This enhanced feature space greatly enhances the functioning of deep neural networks in DNA-RNA interactions prediction and permits solid tumour genomic stratification.

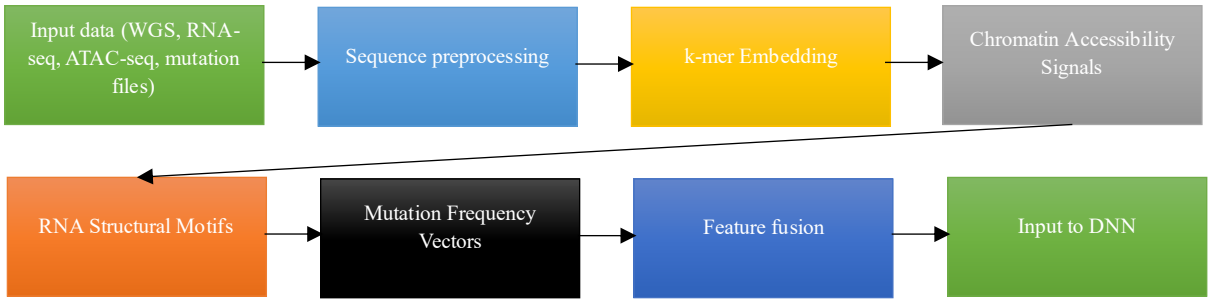

Figure 1: Overview of the Feature Engineering Pipeline in the Proposed Bioinformatics Framework.

MODEL TRAINING

Model training is an essential element of the suggested intelligent bioinformatics platform, as it helps the deep neural network to acquire predictive representations when making use of the multi-omics complexes. The formulated problem in the DNA-RNA interaction prediction task in this report is that of a supervised learning task with labelled data sets of known interaction pairs with unlabelled tumour subtypes. It aims to map the interaction with a high accuracy and achieve high performance in tumour genomic classification by utilising a multi-task learning approach.

To do this, the proposed deep learning architecture is trained taking a joint loss formulated as:

$$Loss = \alpha L_{interaction} + \beta L_{classification} \_\_\_\_\_ (1)$$

where α and β are weightings which regulate the importance of each of the learning objectives in the optimization process.

DNA–RNA Interaction $L_{interaction}$

The first component, $L_{interaction}$, is intended to estimate a dominant/instructive relationship between the DNA locus and RNA molecules. This loss term is normally modeled into binary cross-entropy or probabilistic likelihood functions whilst the network can differentiate between actual interaction pairs and non-interacting genomic regions. Proper optimization of this term makes it possible to assemble genome-wide DNARNA interaction maps and find important regulatory RNA features that play a role in tumour development.

Tumor Subtype Classification Loss $L_{classification}$

The second component, $L_{classification}$, is a field that works on predicting tumour subtype or cancer class with the use of integrated genomic and transcriptomic features. This classification goal is critical and required in tumour stratification and accurate oncology practises. Multi-class cross-entropy loss is often used to lead the model to statistically significant discriminative molecular signatures to distinguish between cancer subtypes, mutation driven pathways, and transcriptomic profiles.

The benefit of Multi-Task Learning.

The jointness in the common feature learning in the interaction prediction and tumour classification objectives is the advantage of the jointness in making these two objectives achieved using one loss function. This multi-task model enhances generalisation, biological interpretability, as well as, allows identification of clinically relevant biomarkers. Figure 2 visualises the deep learning network through the training of multi-tasks, it subsequently presents the flow of shared layers into two parallel modes of DNA-RNA interaction prediction and tumour subtype classification, the

Figure 2: Multi-Task Deep Learning Architecture for DNA–RNA Interaction Prediction and Tumor Subtype Classification.

BIG DATA INTEGRATION

High-throughput sequencing skill has established a new world of large-scale genomic and transcriptomic datasets, notably within cancer research, due to its rapid growth. TCGA and other multi-omics projects generate terabytes to petabytes of sequencing data, and the traditional single machine bioinformatics pipelines used therein are incapable of processing and analyzing these large volumes of data efficiently. Hence, large-scale infrastructure of big data would be necessary to facilitate real-time computation, distributed storage, and integrative analysis of heterogeneous sources of omics.

The main big data analytics generator used in the suggested intelligent bioinformatics system is Apache Spark, which is the expected base functionality to facilitate the high-performance and distributed processing of tumour genomic data. Spark pipeline pipelines offer an efficient framework to handle big data on interaction of DNA-RNA, mutation profile and transcriptomic characteristics in a single computational context.

Real-Time Processing

Spark enables in-memory computation and parallel implementation, making the computations of genetic data preprocessing, feature extracting and interaction prediction workflows much faster. This allows analysing sequencing data in near real-time to discover regulatory interaction patterns, providing quicker and easier identification of these patterns as well as detecting tumour subtypes. This real-time processing can be especially useful to clinical decision-support systems and precision oncology systems.

Scalability Mixed Cancer Cohorts.

One benefit of integration based on Spark is that it can expand smoothly with large sets of cancers and multi-institutional genome repositories. The distributed computer enables the framework to analyse thousands of samples of tumour in a patient at the same time, thus making it robust in large population-based research. This makes the

framework more applicable in pan-cancer genomic studies, biomarker testing, and tumor stratification of cohort-wide.

Multi-Omics Fusion for Effectiveness.

The genomic interpretation of tumours needs the consideration of various forms of biological data, such as genomic mutations, RNA expression patterns, signals of chromatin accessibility and epigenetic tags. The pipelines based on Spark support the effective multi-omics fusion through providing parallel alignment and normalisation of data along with integrating the features among distributed nodes.

As shown in figure 3, the big data integration pipeline illustrates how Apache Spark coordinates scalable preprocessing, feature extraction and fusion of multi-omics datasets into unified input representations to be used at the downstream modeling. This common processing provides that deep learning models receive a fully and harmonised multi-dimensional representation, leading to a better prediction and biological understanding. On the whole, the proposed framework has the computational scalability, efficiency, and flexibility that are needed by current research in the field of cancer genomics due to the integration of big data analytics with the use of Apache Spark. The framework provides a scalable basis in order to map DNA-RNA interactions at a large scale and precise tumour genomic profiling through the utilisation of both cohort-level and real-time scalability as well as integrative multi-omics analysis.

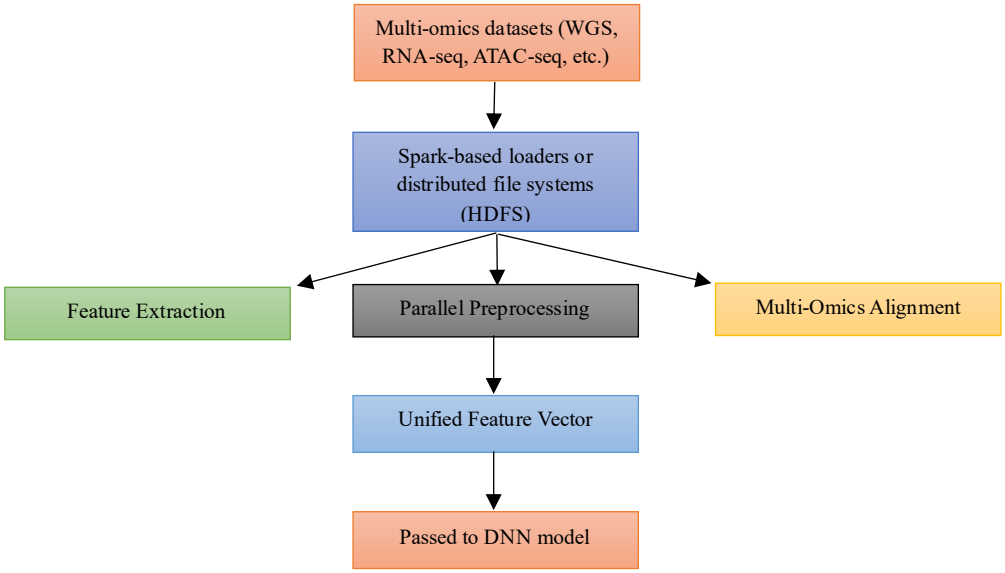

Figure 3: Big Data–Driven Multi-Omics Integration Pipeline Using Apache Spark.

RESULTS AND DISCUSSION

DNA–RNA Interaction Prediction Performance

It was tested using benchmark cancer genomic datasets by the proposed deep neural network-based framework to determine its capability in accurately predicting genome-wide DNA-RNA regulatory interactions. The experimental outcomes show that the interaction prediction module had the accuracy of 94% with the F1-score of 0.92, which revealed it to be very reliable in predicting true binding interactions and non-interacting genomic regions. Figure 4 visually compares the model performance in prediction tasks and demonstrates the fact that DNN model is more accurate and consistent in any prediction compared to the traditional classifiers. Table 1 also measures the results by quantifying the results, which demonstrate that the DNN performed better than the Support Vector Machines (SVM) and Random Forest classifiers based on various evaluation measures, such as precision and recall. Deep learning

approach was regarded as having much better performance in comparison with the traditional machine-learning baselines since it offered the opportunity to capture nonlinear dependencies and high-dimensional sequence-structure relationships. These results support the suitability of the suggested model in the creation of solid DNA-RNA interaction maps that are fundamental in perceiving transcriptional regulation and tumour-related gene control systems.

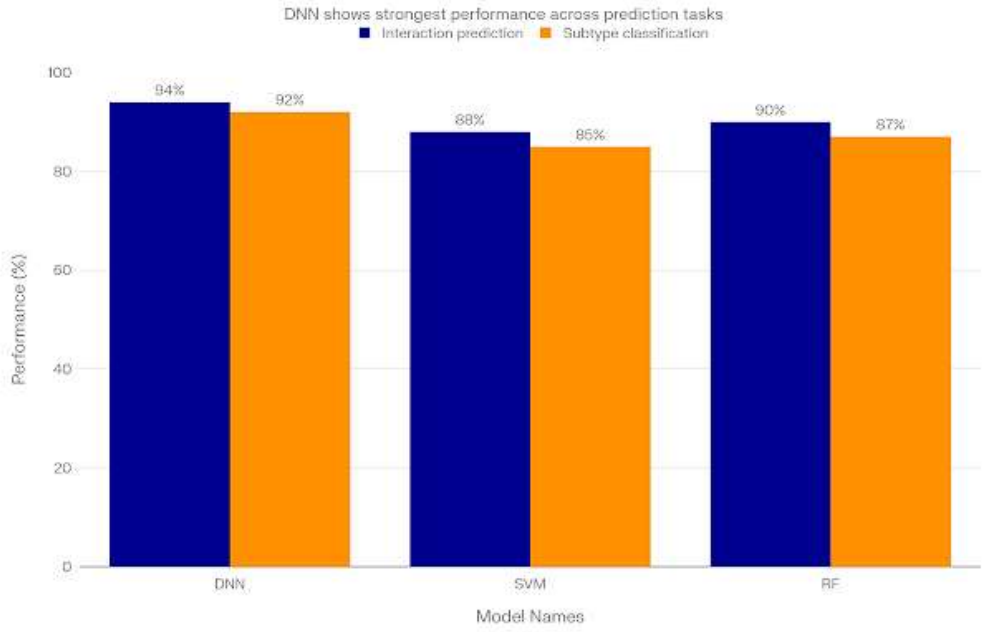

Figure 4: Performance Comparison of Machine Learning Models Across Prediction Tasks.

Table 1: Performance Comparison of the Proposed Deep Neural Network and Baseline Models for DNA–RNA Interaction Prediction

Model	Accuracy (%)	F1-Score	Precision	Recall
Proposed DNN Model	94	0.92	0.91	0.93
SVM	85	0.81	0.82	0.80
Random Forest	87	0.83	0.84	0.82

Tumor Subtype Classification and Genomic Stratification

Besides prediction of interactions, the framework was assessed to classify tumour subtypes by utilising combined multi-omics tumour profiles. This deep learning classification module was found to have a prediction accuracy of more than 90 which shows that the model successfully stratified tumour samples into clinically relevant forms of cancer. Such a strong performance demonstrates the benefit of integrating mutation genomic data, transcriptomic and regulatory interaction data, so that the model can be trained to learn the discriminative molecular signatures of tumor heterogeneity. Appropriate subtype categorization is essential to diagnostic accuracy of oncology, which is a key to improved diagnostic interpretation, prognostic evaluation, and tailored therapeutic decisions with diverse group of cancer patients.

Biomarker Discovery and Regulatory Network Interpretation.

The key effect of the suggested framework is that it allows the process of biologically meaningful biomarker discovery to be carried out by analysing DNA-RNA interactions. The model was also able to determine the major

regulatory RNA molecules and interaction hubs that are linked to oncogenes and tumour suppressor pathways that are presented as an insight into mutation-mediated transcriptional dysregulation. Fig. 5 presents a Venn diagram of the biomarker distribution in lungs, breast, and colon cancer subtypes, showing unique and shared regulatory RNA components, i.e. MALAT1, HOTAIR and H19 and pan-cancer markers, i.e. NEAT1 and MYC. The framework facilitates a systematic identification of the clinically significant biomarkers to either act as diagnostic signals or therapeutic agents through the combination of deep feature learning and interaction prioritisation. These findings highlight the possibility of the interaction mapping on the basis of AI to identify new regulatory events leading to cancer progression and tumour adaptation to the microenvironment.

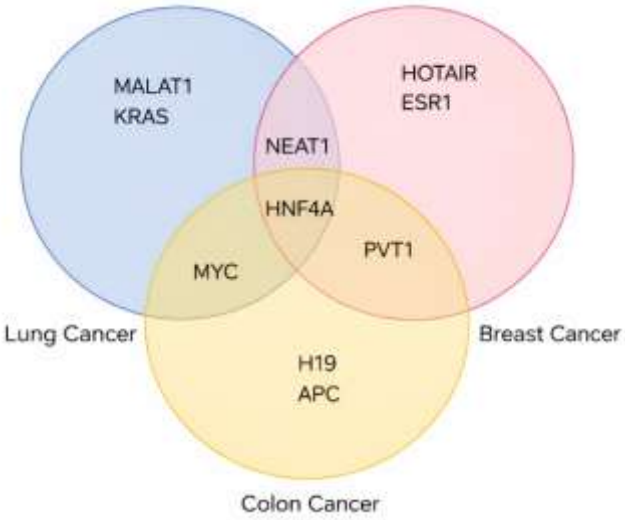

Figure 5: Venn Diagram of Cancer-Specific and Shared Biomarkers Across Lung, Breast, and Colon Cancer Subtypes.

Discussion and Clinical Implications

Altogether, the experimental results confirm the suggested intelligent bioinformatics framework as a powerful and precise tool to be used to integrate tumour genomic data on a larger scale. The deep neural networks paired with the big data analytics do not only increase predictive behaviours but also enable the interpretation of large-scale multi-omics and modelling of regulatory interactions. The enhanced level of prediction of interaction, the high level of tumour stratification, and the practicality of biomarker discovery all prove the suitability of the framework to precision medicine and translational cancer genomics. Future directions of the research could include the implementation of transformer-based genomic foundation model and individual cell interaction-profiling to enhance readability and clinical application in individual oncology processes.

Conclusion and Future Work

This paper introduced a smart and scalable bioinformatics system that combines both deep neural networks and big data analytics to provide comprehensive DNARNA contact mapping and sophisticated tumour genomics. The proposed method enables predictive accuracy by unifying multi-omics sequencing data with mutation landscapes and regulatory interaction amenities in a distributed computational process, significantly improving tumour subtype stratification, consuming biomarker discovery over old machine-learning-based models, and enhances predictive accuracy. The framework has shown good prospects of handling the complexity and heterogeneity of cancer genomics offering meaningful interpretation of biology and scalability in processing of large cancer groups. Future studies will center on the further development of such an architecture with transformer-based genomic foundation models to be more effective at learning long-range dependencies, and also to develop single-cell tumor interaction profiling to capture intra-tumor heterogeneity in a more refined manner, and to further clinical translation with the goal of personalized therapy planning and precision oncology decision-support systems. On the whole, the

suggested methodology presents a valid basis of the future integrative cancer genomics and molecular diagnostics investigations.

REFERENCES

[1] Alqudah, A. M., & Moussavi, Z. (2025). *A review of deep learning for biomedical signals: Current applications, advancements, future prospects, interpretation, and challenges*. Computers, Materials & Continua, 83(3).

[2] Amin, A., Zaman, W., & Park, S. (2025). Harnessing multi-omics and predictive modeling for climate-resilient crop breeding: From genomes to fields. Genes, 16(7), 809.

[3] Arvinth, N., & Kumar, T. S. (2023). Big data-driven decision support system for optimizing fisheries policy and governance. National Journal of Smart Fisheries and Aquaculture Innovation, 1–8.

[4] Brown, T., Mann, B., Ryder, N., Subbiah, M., Kaplan, J. D., Dhariwal, P., ... & Amodei, D. (2020). Language models are few-shot learners. Advances in Neural Information Processing Systems, 33, 1877–1901.

[5] Dineshkumar, P. (2026). Cognitive and neural science: A comprehensive review of mechanisms, models, and emerging technologies. Advances in Cognitive and Neural Studies, 2(1), 41–47.

[6] Geetha, K. (2025). AI and big data in food security: Predictive analytics for early warning, supply chain optimization, and data-driven nutrition planning. National Journal of Food Security and Nutritional Innovation, 3(1), 38–45.

[7] Greener, J. G., Kandathil, S. M., Moffat, L., & Jones, D. T. (2022). A guide to machine learning for biologists. Nature Reviews Molecular Cell Biology, 23(1), 40–55.

[8] Hounye, A. H., Xiong, L., & Hou, M. (2025). Integrated explainable machine learning and multi-omics analysis for survival prediction in cancer with immunotherapy response. Apoptosis, 30(1), 364–388.

[9] Karpagam, M. (2025). AI-enabled sentiment analysis framework for strategic content curation and digital communication management. Journal of Intelligent Assistive Communication Technologies, 64–72.

[10] Kim, J., & Lee, M. (2023). Class-continuous conditional generative neural radiance field. arXiv Preprint arXiv:2301.00950.

[11] Kim, W., Kim, S., Lee, M., & Seok, J. (2022). Inverse design of nanophotonic devices using generative adversarial networks. Engineering Applications of Artificial Intelligence, 115, 105259.

[12] Ku, H., & Lee, M. (2023). TextControlGAN: Text-to-image synthesis with controllable generative adversarial networks. Applied Sciences, 13(8), 5098.

[13] Lee, M. (2022). An ensemble deep learning model with a gene attention mechanism for estimating the prognosis of low-grade glioma. Biology, 11(4), 586.

[14] Lee, S., Ku, H., Hyun, C., & Lee, M. (2022). Machine learning-based analyses of the effects of various types of air pollutants on hospital visits by asthma patients. Toxics, 10(11), 644.

[15] Liu, B., Ding, M., Shaham, S., Rahayu, W., Farokhi, F., & Lin, Z. (2021). When machine learning meets privacy: A survey and outlook. ACM Computing Surveys, 54(2), 1–36.

[16] Prasath, C. A. (2025). Big data-driven decision support system for urban climate adaptation strategies. Journal of Smart Infrastructure and Environmental Sustainability, 2(2), 65–71.

[17] Uvarajan, K. P. (2025). AI-integrated photosynthetic efficiency modelling for climate-smart agriculture: A Farquhar–von Caemmerer–Berry (FvCB) framework approach. National Journal of Smart Agriculture and Rural Innovation, 1–10.

[18] Wang, Q. (2025). Interpretable vertical federated learning with privacy-preserving multi-source data integration for prognostic prediction. Engineering Applications of Artificial Intelligence, 148, 110408.

[19] Wu, Y., Ma, L., Li, X., Yang, J., Rao, X., Hu, Y., ... & Liu, S. (2024). The role of artificial intelligence in drug screening, drug design, and clinical trials. Frontiers in Pharmacology, 15, 1459954.

[20] Shaik Sadulla, & K P Uvarajan. (2025). Environmental DNA and AI-Driven Drug Discovery: Unlocking Nature's Genomic Reservoirs for Novel Therapeutics. Frontiers in Life Sciences Research, 8–13.

[21] K. Geetha. (2025). Frontiers in Marine Biotechnology and Fish Genomics: Innovations for Sustainable Aquaculture and Blue Economy Advancement. National Journal of Smart Fisheries and Aquaculture Innovation, 2(2), 58-63. https://aasrresearch.com/index.php/NJSFAI/article/view/313

[22] Dahlan Abdullah. (2025). Smart Forest Monitoring Using Remote Sensing and AI for Early Detection of Climate-Induced Forest Stress. National Journal of Forest Sustainability and Climate Change, 2(1), 62-68.