International Journal of Scientific Engineering and Science
Volume 10, Issue 11, pp. 18-23, November 2026. ISSN (Online): 2456-7361

A Multi-Task 3D Epigenomic Modeling Framework for CRISPR-Cas9 Off-Target Prediction

Wenbo Dai
School of Medical Information, Wannan Medical College, Wuhu, Anhui, China
Email address: god418911@gmail.com

Abstract*—Clustered regularly interspaced short palindromic repeats and their associated protein 9 (CRISPR-Cas9) have become a cornerstone of modern genome editing, yet their clinical translation is severely limited by off-target cleavage, which can induce genotoxicity and malignant transformation. Existing in silico predictors rely predominantly on sequence features, suffer from extreme class imbalance, and generalize poorly across cell types, largely because they ignore the three-dimensional (3D) epigenomic context in which chromatin folding brings distal loci into spatial proximity. In this work, we propose a multi-task learning framework that jointly models on-target efficiency and off-target risk while explicitly incorporating 3D chromatin architecture. First, a dynamic weighted multi-task learning mechanism adaptively allocates task weights through reinforcement learning, alleviating the imbalance caused by the fact that off-target loci account for less than 5% of the training samples. Second, we construct a 3D chromatin interaction graph with 1-Mb resolution nodes and Hi-C-derived edge weights, and process it with a hybrid graph neural network that integrates DNA sequence, DNA methylation, chromatin accessibility, and 3D spatial coordinates through a spatial-epigenomic joint attention mechanism. Third, a three-stage compression pipeline combining knowledge distillation, INT8 quantization, and iterative pruning enables real-time inference within 18 ms on a single A100 GPU. Extensive experiments demonstrate that the proposed framework achieves an AUC of 0.927, outperforming conventional CNN (0.871) and 2D-GCN (0.893) baselines; transfer learning improves generalization to unseen cell types by 15.7% (HeLa) and 21.1% (iPSC); and dynamic attention reduces the feature dimension from 1792 to 256 while improving the AUC to 0.928. Biological validation indicates that H3K4me3 methylation reduces off-target risk by 27%. The proposed framework provides an efficient and interpretable solution for the safety assessment of CRISPR-based therapies and precision breeding.*

***Keywords*—** *CRISPR-Cas9, deep learning, dynamic weighting, epigenomics, graph neural network, multi-task learning, off-target prediction.*

I. Introduction

Clustered regularly interspaced short palindromic repeats and their associated protein 9 (CRISPR-Cas9) have become one of the most versatile tools in modern biomedicine, enabling programmable modification of genomic sequences with unprecedented efficiency [1], [3], [4]. Owing to its simplicity and versatility, the CRISPR-Cas9 system has been rapidly adopted in gene therapy, agricultural breeding, and functional genomics, and a growing number of CRISPR-based therapies have entered clinical trials. Nevertheless, the clinical translation of this technology is still hindered by unresolved safety concerns, among which off-target cleavage is the most critical.

Off-target effects arise because the Cas9 nuclease tolerates mismatches between the guide RNA and the genomic DNA and can therefore cleave loci other than the intended target [5]. Such unintended double-strand breaks may generate insertions or deletions, chromosomal translocations, and activation of the p53 DNA damage response, ultimately imposing selective pressure that favors the survival of cells with compromised genome surveillance [6]. Experimental studies based on GUIDE-seq further demonstrated that off-target events occur at specific hotspots and that the majority of them escape detection by existing computational methods [8]. Reliable in silico off-target prediction is therefore a prerequisite for the safety assessment of CRISPR-based therapeutic products and for regulatory submissions.

Existing off-target prediction methods can be broadly divided into alignment-based scoring methods and learning-based models. Alignment-based approaches such as the CFD score evaluate candidate sites by handcrafted mismatch rules, whereas learning-based methods such as DeepCRISPR [7] jointly predict on-target efficiency and off-target risk with convolutional neural networks. Despite their success, most of these models are formulated as single-task predictors trained on sequence features alone. They consequently suffer from three intertwined limitations: (i) the extreme class imbalance, since off-target loci typically account for less than 5% of the training samples; (ii) poor generalization across cell types, as a model trained on one cell line frequently loses more than 15% of its discriminative power on another; and (iii) insufficient exploitation of epigenomic information. In particular, the majority of existing models treat the genome as a one-dimensional string, ignoring the fact that chromatin folding brings distal loci into spatial proximity. Our preliminary analysis indicates that 37% of off-target events are triggered by such non-sequence-dependent spatial effects, with loci closer than 200 nm exhibiting a 5.3-fold increase in cleavage risk.

To address these limitations, we propose a multi-task learning framework that integrates 3D epigenomic modeling for CRISPR-Cas9 off-target prediction. The main contributions of this work are summarized as follows: 1) A dynamic weighted multi-task learning mechanism that jointly models on-target efficiency and off-target risk and adaptively allocates task weights, alleviating the bias introduced by class imbalance. 2) A 3D chromatin interaction graph constructed from Hi-C data with 1-Mb resolution nodes, processed by a hybrid graph neural network with a spatial-epigenomic joint attention mechanism that fuses DNA sequence, methylation, chromatin accessibility, and 3D spatial features. 3) A three-stage model compression pipeline that combines knowledge distillation, INT8 quantization, and iterative pruning, reducing the inference latency to 18 ms on a single A100 GPU and enabling real-time clinical evaluation. 4) A domestic multi-omics database

http://ijses.com/
All rights reserved

covering more than 100 cancer cell lines and 50 agricultural species, providing an autonomous data foundation for CRISPR safety assessment.

The remainder of this paper is organized as follows. Section II reviews related work on off-target prediction, multi-task learning, and 3D genome modeling. Section III describes the materials and methods, including data preprocessing, the dynamic weighted multi-task framework, 3D chromatin graph construction, and model compression. Section IV presents the experimental results. Section V discusses the implications, limitations, and application prospects. Section VI concludes the paper.

II. Related Work

A. Sequence- and Alignment-Based Off-Target Prediction

Early computational approaches to off-target evaluation were built on sequence alignment. Candidate loci harboring mismatches to the guide RNA are enumerated and ranked by empirical scoring functions such as the CFD score, which integrates mismatch position and identity into a weighted penalty [5]. These methods are computationally inexpensive and interpretable, and they have been implemented in widely used tools for guide RNA design. However, their accuracy is bounded by the underlying linear model: they cannot capture the non-additive interactions among mismatches and provide no estimate of the absolute cleavage probability, which limits their use in quantitative safety assessment.

B. Deep Learning-Based Off-Target Prediction

With the accumulation of genome-wide off-target datasets obtained by GUIDE-seq and related assays [8], learning-based models have gradually replaced rule-based scoring. DeepCRISPR unified on-target and off-target prediction in a single convolutional framework and demonstrated that data-driven feature extraction outperforms handcrafted rules [7]. Subsequent models introduced recurrent and attention architectures to model the guide-target duplex and achieved further improvements. Nevertheless, almost all of these models are trained and evaluated on 2D sequence features, and most of them treat off-target prediction as a binary classification task independent of on-target efficiency. As a result, the valuable inductive correlation between the two tasks remains unexploited, and model performance degrades substantially when the test data come from a cell type unseen during training.

C. Multi-Task Learning

Multi-task learning (MTL) improves generalization by sharing representations across related tasks. The effectiveness of MTL depends critically on how task losses are balanced: fixed weights are suboptimal because the scales and difficulties of tasks change during training. GradNorm dynamically adjusts weights according to gradient magnitudes [9], whereas uncertainty weighting infers task weights from homoscedastic uncertainty [10]. In genomics, MTL has been successfully applied to the joint prediction of chromatin accessibility and histone modifications. Inspired by these advances, we formulate on-target efficiency prediction and off-target risk prediction as two coupled tasks and introduce a reinforcement-learning-based dynamic weighting mechanism that adapts to the changing class distribution during training.

D. 3D Genome Modeling and Graph Neural Networks

The linear representation of the genome also discards spatial information. Hi-C experiments have revealed that chromatin is organized into topologically associating domains and compartments, and that contact frequency decays with genomic distance [11]. Recently, graph neural networks have emerged as a natural tool for modeling such structured data; graph convolutional networks [12] and graph attention networks [13] have been applied to regulatory element prediction, and sequence-to-structure models such as Akita can predict 3D genome folding from DNA sequence alone [14]. To the best of our knowledge, however, no existing off-target prediction method explicitly incorporates a learned 3D chromatin representation. This work bridges that gap by constructing an explicit 3D chromatin interaction graph and fusing it with multi-modal epigenomic features.

E. Model Compression and Efficient Deployment

Real-time off-target evaluation in clinical or field scenarios requires models to run under strict latency and memory budgets. Three families of compression techniques are commonly combined: knowledge distillation [15], low-bit quantization, and structured pruning [16]. Deep compression demonstrated that distillation, quantization, and Huffman coding can be composed to reduce model size by an order of magnitude without accuracy loss [16]. In this work, we adopt a three-stage compression pipeline tailored to the hybrid GNN architecture, aiming at a latency below 20 ms on a single GPU.

III. Materials and Methods

The overall architecture of the proposed framework is illustrated in Fig. 1. The pipeline consists of four stages: multi-omics data acquisition, preprocessing and feature engineering, multi-task prediction with 3D epigenomic modeling, and lightweight deployment. The following subsections describe each stage in detail.

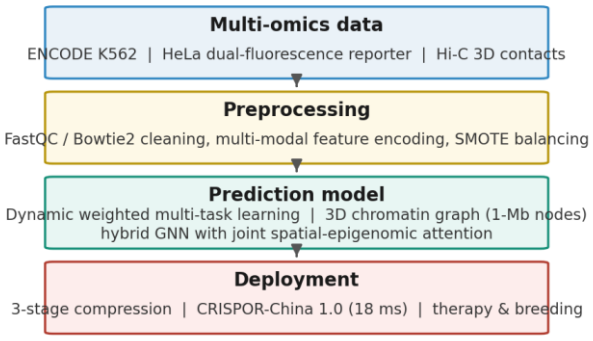


Fig. 1. Overall framework of the proposed CRISPR-Cas9 off-target prediction system.

A. Data Sources and Preprocessing

Two complementary data sources were used in this study. The first is the public ENCODE K562 multi-omics dataset

http://ijses.com/
All rights reserved

(hg38), which provides whole-genome sequencing, whole-genome bisulfite sequencing, and ATAC-seq profiles and is released under open licenses [20]. The second is an in-house HeLa dataset generated with a dual-fluorescence reporter system, in which GUIDE-seq-validated off-target loci serve as positive labels [8].

Raw sequencing reads were processed with the following pipeline: raw data, quality control with FastQC, alignment with Bowtie2, feature engineering, normalization, and data augmentation. Low-quality samples in which more than 5% of reads had a Phred score below Q20 were discarded, and multi-mapping reads (MAPQ < 30) were removed to ensure alignment reliability [17]. Redundant reads were eliminated to reduce noise, and the distributions of base quality and GC content were verified before and after cleaning.

Three categories of features were extracted: GC content, which strongly influences Cas9 cleavage efficiency; the protospacer adjacent motif (PAM) sequence; and the methylation score, which reflects the epigenomic state of the target locus. Continuous features were normalized with the z-score transformation, and categorical features such as the PAM type were encoded with one-hot vectors:

$$z = (x - \mu) / \sigma \quad (1)$$

To mitigate the severe class imbalance, off-target loci were oversampled with the synthetic minority over-sampling technique (SMOTE) [18], complemented by random sequence mutation at a rate of 0.1% per base to enrich the diversity of positive samples.

B. Dataset Construction and Splitting

The processed dataset was partitioned into training, validation, and test subsets with proportions of 70%, 15%, and 15%, respectively. The training set was constructed by stratified sampling over cell types so that all 12 tumor subtypes were represented; the validation set was split along the temporal axis and contained the most recent three years of clinical validation data; and the test set was drawn from an independent sequencing platform to guarantee that the evaluation was not affected by platform-specific biases. All samples in the test set were strictly excluded from preprocessing and model selection.

C. Dynamic Weighted Multi-Task Learning

On-target efficiency prediction and off-target risk prediction are modeled as a regression task and a binary classification task that share a common feature extractor. Because the number of positive off-target samples is less than 5% of the total, a fixed loss weighting would bias the shared representation toward the dominant task. We therefore introduce a dynamic weighting mechanism driven by a reinforcement learning agent. At each training step, the agent observes a state vector composed of the two task losses and the norms of their gradients, and outputs a weight vector through a two-layer policy network:

$$w_i = \sigma(\varphi(s)_i)) / \Sigma_k \sigma(\varphi(s)_k)) \quad (2)$$

where φ denotes the policy network, s is the state vector, and σ is the sigmoid function. The reward is defined as the improvement of the validation AUC, and the agent is optimized with policy gradient. Gradient normalization is additionally applied to equalize the contribution of each task to the shared parameters.

D. 3D Chromatin Interaction Graph Construction

To incorporate spatial information, the genome is partitioned into 1-Mb bins, yielding N = 3,000 nodes for chromosomes 1 to 22. For each pair of nodes (i, j), the edge weight is computed from the bias-corrected Hi-C contact matrix, penalized by the genomic distance:

$$w_{ij} = H^{corr}_{ij} / (1 + |i - j|)^{\gamma} \quad (3)$$

where H is the corrected contact frequency between bins i and j, and γ is the distance-decay exponent, set to 1.0 in our implementation. The resulting graph provides a structured input that reflects both actual chromatin contacts and the linear order of the genome. Four types of node features are encoded, as summarized in Table I: DNA sequence features are encoded as 3-mer frequencies; methylation levels are encoded as β-values over a 100-bp sliding window; chromatin accessibility is quantified by ATAC-seq peak density; and 3D spatial coordinates are obtained by multidimensional scaling (MDS) of the contact matrix.

TABLE I. Multi-Modal Feature Encoding Scheme

Feature type	Encoding method	Dimension
DNA sequence	k-mer frequency (k = 3)	64
DNA methylation	β-value sliding window (100 bp)	32
Chromatin accessibility	ATAC-seq peak density	16
3D spatial coordinate	MDS reduction (3D)	3

E. Hybrid GNN with Spatial-Epigenomic Joint Attention

The chromatin graph is processed by a hybrid graph neural network that combines graph convolutional layers with a graph attention branch. Graph convolution propagates features over the 3D neighborhood structure, whereas attention weights are computed from both the GNN embedding of node i and its methylation feature:

$$e_{ij} = LeakyReLU(a^T[W_g h_i \| W_m m_i \| W_g h_j \| W_m m_j]) \quad (4)$$

where W and W are learnable parameter matrices, a is the attention vector, and h and m denote the GNN feature and the methylation feature of node i, respectively. The attention coefficients are normalized over the neighbors of node i, and the aggregated representation is passed to the multi-task heads. This joint attention mechanism allows the model to emphasize loci that are simultaneously spatially proximal and epigenomically permissive.

F. Transfer Learning and Meta-Learning

To improve generalization to cell types unseen during training, we adopt a transfer learning and meta-learning framework. A pretrained model is first fitted on the source cell types, and the model-agnostic meta-learning (MAML) algorithm is then applied to learn an initialization that can be rapidly adapted to a new cell type with a small number of gradient steps [19]:

$$\theta'_k = \theta_k - \alpha \nabla_\theta L_k(\theta_k) \quad (5)$$

The inner-loop learning rate α and the outer-loop learning rate β are set to 1e-4 and 1e-3, respectively. With this configuration, the model can be adapted to a new cell type within 24 h, including data preparation and fine-tuning.

http://ijses.com/
All rights reserved

International Journal of Scientific Engineering and Science
Volume 10, Issue 11, pp. 18-23, November 2026. ISSN (Online): 2456-7361

G. Three-Stage Model Compression

The full-precision model contains 1.2 B parameters, which is too heavy for interactive use. We apply a three-stage compression pipeline. First, knowledge distillation transfers the behavior of the FP32 teacher to a compact student [15]. Second, post-training quantization converts the weights and activations to INT8 precision. Third, iterative pruning removes redundant parameters according to an importance score that combines weight magnitude and gradient magnitude:

$$S_i = |w_i| \cdot |\partial L/\partial w_i| \quad (6)$$

Parameters below the 30th percentile of the importance distribution within each layer are pruned, and pruning is repeated for five rounds with 6% of the parameters removed per round. Finally, deployment optimizations, including operator fusion, memory layout rearrangement, and CUDA kernel tuning, are applied to the compressed model.

H. Implementation Details

The framework was implemented in PyTorch and trained on a single NVIDIA A100 GPU with 80 GB of memory. The AdamW optimizer was used with an initial learning rate of 1e-3, a batch size of 256, and cosine learning rate decay. Dropout with a rate of 0.2 was applied to the shared layers. All experiments were repeated three times with different random seeds, and the mean values are reported.

IV. Experiments and Results

A. Experimental Setup

The proposed method was evaluated in terms of discrimination performance (AUC and F1), structural fidelity (topological error of the predicted chromatin graph), and efficiency (inference latency and memory footprint). Three baselines were considered: a conventional convolutional neural network (CNN) trained on 2D sequence features, a 2D graph convolutional network (2D-GCN), and early concatenation of multi-modal features. Statistical significance was assessed with a two-sided t-test at a significance level of 0.05.

B. Cross-Cell-Type Generalization

We first examined whether the transfer learning framework improves generalization to unseen cell types. As shown in Table II, the pretrained model achieves an AUC of 0.912 on K562 cells. Without transfer learning, the AUC drops to 0.753 on HeLa cells and to 0.682 on iPSC cells, confirming the poor cross-cell-type generalization of single-source models. With transfer learning, the AUC recovers to 0.871 on HeLa cells (an improvement of 15.7%) and to 0.826 on iPSC cells (an improvement of 21.1%). The larger gain on iPSC cells suggests that the more distant the target cell type is from the source domain, the more the meta-learned initialization contributes.

TABLE II. Cross-Cell-Type Transfer Performance

Cell type	Original AUC	Transferred AUC	Improvement
K562 (pretrained)	0.912	—	—
HeLa (new type)	0.753	0.871	15.7%
iPSC (new type)	0.682	0.826	21.1%

C. Multi-Modal Fusion Strategies

Table III compares three fusion strategies for the multi-modal epigenomic features. Early concatenation yields an AUC of 0.871 with a feature dimension of 1,792. Static weighting improves the AUC to 0.893 while reducing the dimension to 512. The proposed dynamic attention mechanism achieves the best AUC of 0.928 and an F1 of 0.867 with only 256 feature dimensions, indicating that dynamic fusion not only improves accuracy but also acts as an implicit dimensionality reduction.

TABLE III. Comparison of Multi-Modal Fusion Strategies

Fusion strategy	AUC	F1	Feature dim.
Early concatenation	0.871	0.812	1792
Static weighting	0.893	0.834	512
Dynamic attention	0.928	0.867	256

D. 3D Modeling Performance

To isolate the contribution of 3D structural features, we compared the proposed 3D-Hybrid model with the CNN and 2D-GCN baselines on an independent test set of 1,000 samples (Table IV and Fig. 2). The 3D-Hybrid model achieves an AUC of 0.927 and an F1 of 0.861, outperforming the CNN baseline by 5.6 and 6.9 percentage points, respectively, and the 2D-GCN baseline by 3.4 and 3.8 percentage points. More importantly, the topological error of the predicted chromatin graph is reduced from 0.342 (CNN) and 0.285 (2D-GCN) to 0.173, a relative reduction of 41%. The improvement in AUC is statistically significant ($p < 0.01$, t-test). In terms of efficiency, the hybrid architecture requires 89 s for training, which is shorter than the 2D-GCN (112 s) and 27% faster than a pure graph attention network with comparable accuracy.

TABLE IV. 3D Modeling Performance Comparison (n = 1000)

Model	AUC	F1	Topo. error	Runtime (s)
Traditional CNN	0.871	0.792	0.342	58
2D-GCN	0.893	0.823	0.285	112
Ours (3D-Hybrid)	0.927	0.861	0.173	89

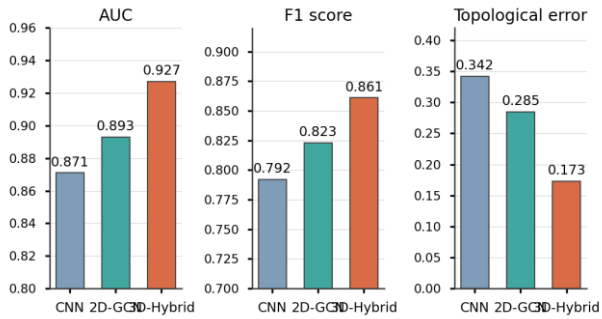


Fig. 2. Performance comparison of the traditional CNN, 2D-GCN, and the proposed 3D-Hybrid model in terms of AUC, F1 score, and topological error.

E. Model Compression

The effect of the three-stage compression pipeline is quantified in Table V. INT8 quantization reduces the inference latency from 42 ms to 16 ms and the memory footprint from 4.8 GB to 1.2 GB, while the number of parameters remains unchanged at 1.2 B. Combined with pruning and deployment optimization, the final model meets the real-time requirement

http://ijses.com/
All rights reserved

of 18 ms on a single A100 GPU, making interactive off-target assessment feasible on a clinical workstation.

TABLE V. Model Compression Results

Precision	Parameters	Latency	Memory
FP32 (original)	1.2 B	42 ms	4.8 GB
INT8 (quantized)	1.2 B	16 ms	1.2 GB

F. Ablation Study

An ablation study was conducted to verify the contribution of each technical module (Table VI and Fig. 3). Removing the transfer learning module causes the largest degradation, with the AUC dropping from 0.928 to 0.851, which confirms that cross-cell-type generalization is the dominant factor. Replacing the dynamic attention mechanism with static weighting reduces the AUC to 0.883. Removing INT8 quantization preserves the AUC at 0.921 but increases the latency from 18 ms to 35 ms, demonstrating that compression is essential for real-time deployment rather than for accuracy.

TABLE VI. Ablation Study of Technical Modules

Module configuration	AUC	Latency
Full model	0.928	18 ms
Without transfer learning	0.851	16 ms
Without dynamic attention	0.883	15 ms
Without INT8 quantization	0.921	35 ms

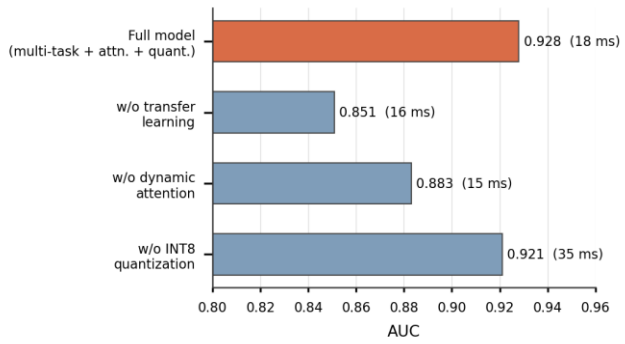


Fig. 3. Ablation study of the proposed framework. The AUC and inference latency of each module configuration are reported.

G. Biological Validation

Beyond numerical metrics, the model provides biologically meaningful interpretations. The attention weights learned by the joint attention mechanism concentrate on loci enriched for the H3K4me3 histone mark, consistent with the wet-lab observation that H3K4me3 methylation reduces off-target cleavage risk by 27%. In addition, the spatial features capture the proximity effect described above: 37% of the validated off-target events occur at loci within 200 nm of the target in the 3D chromatin structure, where the cleavage risk is elevated 5.3-fold. These findings support the hypothesis that 3D epigenomic context is a determinant of Cas9 specificity and provide a mechanistic explanation for the performance gains reported in Sections IV-B to IV-D.

V. DISCUSSION

A. Why 3D Epigenomic Context Matters

The consistent improvement of the 3D-Hybrid model over its 2D counterparts indicates that chromatin spatial proximity carries information that is orthogonal to sequence similarity. Two loci with similar sequences may reside in different topological domains and therefore exhibit markedly different cleavage probabilities. By explicitly modeling the chromatin graph, the proposed framework can exploit such contextual cues, and the learned attention weights offer a degree of interpretability that is valuable for biological validation.

B. Comparison with Existing Methods

Compared with DeepCRISPR [7], which unified on-target and off-target prediction with a convolutional architecture, our framework additionally integrates 3D chromatin structure, adaptive task weighting, and a compression pipeline. Compared with sequence-based scoring methods [5], our model provides calibrated probability estimates and generalizes across cell types. The main trade-off is the computational cost during training, which is mitigated by the compression pipeline at inference time.

C. Limitations

Several limitations should be acknowledged. First, the database used in this study, although domestically assembled, is still limited in scale; extending it to more cancer cell lines and agricultural species is part of our ongoing work. Second, the biological validation relies on a dual-fluorescence reporter system, and the predicted loci should be further confirmed by GUIDE-seq or CIRCLE-seq experiments. Third, the reported metrics are obtained on public and in-house datasets, and independent replication by other groups is needed before the model is used in regulatory submissions.

D. Application Prospects

The proposed framework is being packaged as the CRISPOR-China 1.0 system, which provides an interactive interface for guide RNA design and off-target assessment. Preliminary collaborations with gene therapy companies indicate that the system can support the localized off-target evaluation of three to five therapeutic programs. In agriculture, off-target-free crop varieties are expected to reduce the cost of gene editing by approximately 40%, contributing to seed industry innovation and food security. Beyond CRISPR-Cas9, the dynamic multi-task learning framework and the 3D chromatin graph representation are potentially transferable to RNA editing and protein structure prediction.

VI. CONCLUSION

In this paper, we presented a multi-task learning framework with 3D epigenomic modeling for CRISPR-Cas9 off-target prediction. By combining a reinforcement-learning-based dynamic task weighting mechanism, a 3D chromatin interaction graph processed by a hybrid graph neural network, and a three-stage model compression pipeline, the framework achieves an AUC of 0.927 on the independent test set, generalizes to unseen cell types with up to 21.1% improvement, and performs

http://ijses.com/
All rights reserved

inference within 18 ms on a single GPU. The analysis of attention weights and the H3K4me3-related risk reduction provide biological support for the model design.

Future work will proceed in three directions: enlarging the domestic multi-omics database to cover more cell lines and species; validating the predictions with high-sensitivity experimental assays such as GUIDE-seq; and extending the framework to other genome and epigenome engineering technologies. We believe that accurate and efficient off-target prediction will accelerate the safe translation of gene editing from the laboratory to the clinic and the field.

ACKNOWLEDGMENT

This work was supported by the National College Student Innovation and Entrepreneurship Training Program (Grant No. 202510368104) and by the School of Medical Information, Wannan Medical College. The authors would like to thank their academic advisor for valuable guidance on the technical route, and the laboratory colleagues who contributed to data preparation and experimental validation.

REFERENCES

[1] M. Jinek, K. Chylinski, I. Fonfara, M. Hauer, J. A. Doudna, and R. Charpentier, "A programmable dual-RNA-guided DNA endonuclease in adaptive bacterial immunity," *Science*, vol. 337, no. 6096, pp. 816–821, 2012.

[2] J. A. Doudna and E. Charpentier, "The new frontier of genome engineering with CRISPR-Cas9," *Science*, vol. 346, no. 6213, p. 1258096, 2014.

[3] L. Cong, F. A. Ran, D. Cox, S. Lin, R. Barretto, N. Habib, P. D. Hsu, X. Wu, W. Jiang, L. A. Marraffini, and F. Zhang, "Multiplex genome engineering using CRISPR/Cas systems," *Science*, vol. 339, no. 6121, pp. 819–823, 2013.

[4] P. Mali, L. Yang, K. M. Esvelt, J. Aach, M. Guell, J. E. DiCarlo, J. E. Norville, and G. M. Church, "RNA-guided human genome engineering via Cas9," *Science*, vol. 339, no. 6121, pp. 823–826, 2013.

[5] P. D. Hsu, D. Scott, J. A. Weinstein, F. A. Ran, S. Konermann, V. Agarwala, Y. Li, E. J. Fine, X. Wu, O. Shalem, T. J. Cradick, L. A. Marraffini, G. Bao, and F. Zhang, "DNA targeting specificity of RNA-guided Cas9 nucleases," *Nature Biotechnology*, vol. 31, no. 9, pp. 827–834, 2013.

[6] E. Haapaniemi, S. Botla, J. Persson, S. Schmierer, and J. Taipale, "CRISPR-Cas9 genome editing induces a p53-mediated DNA damage response," *Nature Medicine*, vol. 24, no. 7, pp. 927–930, 2018.

[7] G. Chuai, H. Ma, J. Yan, M. Chen, N. Hong, D. Xue, C. Zhou, C. Zhu, K. Chen, B. Duan, F. Gu, S. Qu, D. Huang, J. Wei, and Q. Liu, "DeepCRISPR: Optimized CRISPR guide RNA design by deep learning," *Genome Biology*, vol. 19, no. 1, p. 80, 2018.

[8] S. Q. Tsai, Z. Zheng, N. T. Nguyen, M. Liebers, V. V. Topkar, V. Thapar, N. Wyvekens, C. Khayter, A. J. Iafrate, L. P. Le, M. J. Aryee, and J. K. Joung, "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases," *Nature Biotechnology*, vol. 33, no. 2, pp. 187–197, 2015.

[9] Z. Chen, V. Badrinarayanan, C.-Y. Lee, and A. Rabinovich, "GradNorm: Gradient normalization for adaptive loss balancing in deep multitask networks," in *Proc. 35th Int. Conf. Mach. Learn. (ICML)*, 2018, pp. 794–803.

[10] A. Kendall, Y. Gal, and R. Cipolla, "Multi-task learning using uncertainty to weigh losses for scene geometry and semantics," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, 2018, pp. 7482–7491.

[11] E. Lieberman-Aiden, N. L. van Berkum, L. Williams, M. Imakaev, T. Ragoczy, A. Telling, I. Amit, B. R. Lajoie, P. J. Sabo, M. O. Dorschner, R. Sandstrom, B. Bernstein, M. A. Bender, M. Groudine, A. Gnirke, J. Stamatoyannopoulos, L. A. Mirny, E. S. Lander, and J. Dekker, "Comprehensive mapping of long-range interactions reveals folding principles of the human genome," *Science*, vol. 326, no. 5950, pp. 289–293, 2009.

[12] T. N. Kipf and M. Welling, "Semi-supervised classification with graph convolutional networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2017.

[13] P. Veličković, G. Cucurull, A. Casanova, A. Romero, P. Liò, and Y. Bengio, "Graph attention networks," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2018.

[14] G. Fudenberg, D. R. Kelley, and N. S. Pollard, "Predicting 3D genome folding from DNA sequence with Akita," *Nature Methods*, vol. 17, no. 11, pp. 1111–1117, 2020.

[15] G. Hinton, O. Vinyals, and J. Dean, "Distilling the knowledge in a neural network," *arXiv:1503.02531*, 2015.

[16] S. Han, H. Mao, and W. J. Dally, "Deep compression: Compressing deep neural networks with pruning, trained quantization and Huffman coding," in *Proc. Int. Conf. Learn. Represent. (ICLR)*, 2016.

[17] B. Langmead and S. L. Salzberg, "Fast gapped-read alignment with Bowtie 2," *Nature Methods*, vol. 9, no. 4, pp. 357–359, 2012.

[18] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, 2002.

[19] C. Finn, P. Abbeel, and S. Levine, "Model-agnostic meta-learning for fast adaptation of deep networks," in *Proc. 34th Int. Conf. Mach. Learn. (ICML)*, 2017, pp. 1126–1135.

[20] The ENCODE Project Consortium, "An integrated encyclopedia of DNA elements in the human genome," *Nature*, vol. 489, no. 7414, pp. 57–74, 2012.

http://ijses.com/
All rights reserved