

The Convergence of Artificial Intelligence and CRISPR-Cas9: A Review of Innovations, Challenges, and Future Vistas in Genome Editing

Muskan Kirar¹, Palak Sobti^{1*}, Akash Sharan², Abhinav Mishra³, Shaweta Sharma⁴

^{1,3,4,5} Nielit University, Ropar, Punjab, India,

² Nielit University, Gorakhpur

Email:muskankirar1@gmail.com Email:palaksobti567@gmail.com

Abstract. CRISPR-Cas9 has become an effective genome-editing methodology, but the practical deployment of this approach is limited due to challenges such as unintended off-target effects, inconsistent editing efficiency, and lower predictability in complex biological systems. Simultaneously, the development of artificial intelligence (AI) has made it possible to model genomic processes based on data and help us to increase the accuracy and reliability of CRISPR-based treatment. This review is a critical assessment of recent applications of AI combined with CRISPR-Cas9, and specifically guides RNA optimisation, off-target prediction, data enhancement methods, and identification of synthetic lethal interactions applicable to cancer therapy. Instead of a chronological summary of existing methods, the paper includes a comparative analysis of AI-CRISPR-based techniques, along with methodologies, strengths, limitations, and readiness to translate into clinical practice. To conclude, a structured comparative table is highlighted to outline the unresolved gaps that should be bridged in order to translate clinical knowledge to practice. Lastly, recent issues, ethical aspects, and future research are addressed in the direction of responsible AI-driven genome editing.

Keywords: Artificial Intelligence (AI), Bioinformatics, Cancer Therapy, CRISPR-Cas9, Deep Learning, Genome Editing, Guide RNA (gRNA) Optimisation, Machine Learning, Off-Target Prediction, On-Target Efficiency, Precision Oncology, Synthetic Lethality.

(Article history: Selected papers from NIELIT's International Conference on Communication, Electronics and Digital Technologies (NICE-DT'26), 29-30Jan, 2026, Gorakhpur, India)

1 Introduction

Genome editing with CRISPR as a prokaryotic adaptive immune system has been repurposed into a programmable genome-wide system to induce targeted changes in sequences of nucleic acids using guide RNA - directing cleavage by Cas proteins [13,14]. Among the different systems of CRISPR-associated, CRISPR-Cas9 has been the most widely used as it is relatively simple and flexible. Even though the CRISPR- Cas9 technology has potential uses in biological research, as well as in a limited number of additional applications in therapeutic applications, the technology is limited to the unresolved challenges of off-target effects, uneven on-target performance, and increased variation in predictability when used in complex biological systems[17].

Simultaneously, artificial intelligence (AI) has emerged as a progressively more significant resource in biomedical research because it has enabled the data-driven analysis of large and multidimensional biological datasets.[18] Deep-learning and machine-learning techniques have proven to be highly effective in modelling genomic sequences and predicting molecular interactions and, accordingly, provide precise decision-making in complex biological systems. The intersection of AI and CRISPR-Cas9 has thus become an effective option that could help overcome the significant technical drawbacks of genome editing, specifically, enhancing specificity, efficiency, and predictability [15].

Recent research has investigated AI-aided methods of optimising guide RNA, predicting off-target activity, augmenting data and identifying novel therapeutic targets. Nevertheless, much of the literature describes these techniques in serial or application-specific ways, making systematic comparison of their methodological assumptions, performance trade-offs, and response to real-world or clinical applications. Consequently, the overall synthesis of AI-CRISPR approaches, especially its critical methods synthesis, remains limited.

To solve this gap, this review will specifically assess the representative AI-based CRISPR-Cas9 methods, but with a focus on cancer research applications where high specificity and reliability of genetic interventions are required. Instead of providing a descriptive overview, the paper critically evaluates the capabilities, shortcomings and translational maturity of current methods and summarises the main trends with the help of a structured comparison table. The review also addresses the unsolved technical and ethical issues, as well as future research avenues, on which AI-assisted genome editing can be developed responsibly.

2 Essential Technology: CRISPR-Cas9 and Artificial Intelligence.

2.1 CRISPR-Cas9 is a platform that is used for genome editing.

CRISPR-Cas9 is a programmable genome-editing system and relies on the adaptive immune response mechanism of bacteria and archaea, with the help of which it is applied to identify and cut invading nucleic acids [1,2]. The system in the context of genome engineering consists of a guide RNA (gRNA), which directs the Cas9 endonuclease to a genomic complementary sequence of DNA, which induces a site-specific, double-stranded break. DNA cleavage causes the activation of intrinsic cell repair mechanisms, the two major ones being homology-directed repair (HDR) or non-homologous end joining (NHEJ). NHEJ is an error-prone process that often induces insertion or deletion of sequences that disrupt genes, whereas HDR precisely alters the sequences of genes in the presence of a donor DNA template [3]. CRISPR-Cas9 has become widely used in functional genomics, disease modelling, and other therapeutic research, especially owing to its efficiency in cancer and genetic disorders [3,7]. Although these benefits exist, the efficiency of CRISPR-Cas9 depends on a number of factors, such as the sequence of gRNA, the genome, and the dynamics of the DNA repair system within the cell. The inconsistency in the editing efficiency and the off-target activity are still two significant issues, which prompt future efforts to generate predictive and optimisation-oriented methods to enhance the safety and reliability of genome editing.

2.2 Artificial Intelligence in Biomedical Research

Artificial intelligence (AI) can be defined as the computational methods that support systems to learn patterns based on data and provide predictions or decisions without being programmed with rules [3]. Machine learning and deep learning methods are finding use in biomedical research to analyse large and complicated datasets that are based on genomics, transcriptomics, proteomics, and medical imaging.

AI-based methods have been applied and used in mutation detection, biomarker identification, drug target discovery and image-based diseases characterization. Such methods can be successfully applied for effective modelling of nonlinear relationships and high-dimensional interactions of features that are hard to model with standard analytical signal processing tools. In the genome editing domain, AI can offer predictive modelling of gRNA activity, evaluation of off-target hazard, and integration of various biological characteristics, complementing CRISPR-Cas9-based technologies and overcoming major limitations of the genome editing workflows.

3 THE MOST CRITICAL CHALLENGES IN ACHIEVING PRECISION AND SAFETY IN GENOME EDITING

Despite the transformative potential of the usage of CRISPR-Cas9, it is still not used on a large-scale basis in clinical and industry settings due to the ongoing issues of accuracy, efficacy, and safety of the technology in question. The inconsistency of the effect of editing and the risk of unintended modifications of the genome undermine the reliability required of therapeutic application. In order to overcome these shortcomings, should focus on CRISPR-Cas9 as a possible solution that can be effectively employed in the laboratory, as well as in clinical practice. The limitation arises from the complicated sequence-based and context-based factors, which are not easy to solve with the traditional design methods, which encourage the utilisation of data-driven and predictive approaches.

3.1 Off-Target Effects

Off-target activity is one of the most crucial safety issues in CRISPR-based genome editing[19],[20]. The effects occur in situations where the guide RNA (gRNA) is utilised to direct the Cas9 nuclease to loci in the genome that are very similar to the target sequence, but have one or more mismatches [4]. Such unintended cleavages may lead to the introduction of harmful mutations, which may disrupt vital genes or regulators, and this is a significant risk when used as therapies. It has been found in experimental studies that such features of a sequence as mismatch position within the protospacer and nucleotide substitution patterns affect off-target susceptibility [4]. These observations indicate that there are latent sequence-dependent motifs that indicate on-target and off-target sites. Nonetheless, these features are hard to identify and assign the appropriate weights manually, which is why it is necessary to have computational models that can learn the sophisticated and subtle patterns of the sequences to forecast off-target risk with greater precision.

3.2 Inconsistent Editing Efficiency

Besides safety considerations, CRISPR-Cas9 has substantial variability in the efficiency of editing on - target, which directly

relates to experimental reproducibility and therapeutic reliability. The process of editing is highly affected by intrinsic gRNA activity, sequence composition and local genomic context [5]. Guides exhibiting low intrinsic activity are also not able to recruit Cas9 efficiently, leading to inconsistent editing behaviours.

This variability makes it complicated to design an experiment and constrains the prediction of therapeutic interventions. In clinical applications, where high editing efficiency is required to be relatively consistent and robust, this lack of reliability in high editing efficiency is a significant obstacle. To overcome this problem, systematic approaches are needed, which would predict and optimise the performance of gRNA prior to experiments.

3.3 Data Scarcity and Class Imbalance

Availability of training data is again a limitation to the development of accurate predictive models to optimise CRISPR. Despite the increasing size of a range of experimental datasets that are scaled by high-throughput screening technologies, there are still a relatively small number of gRNAs that have well-characterised on-target efficiency and experimentally validated off-target phenotypes [5]. Lack of data scarcity limits the generalizability of machine learning models.

In addition, datasets associated with CRISPR are highly unbalanced in classes. The possible off-target sites in the whole human genome are tremendous, and experimentally validated off-target events form a tiny fraction of the available data points [4,6]. This imbalance may bias predictive models in the negative classes, and thus, the models are less sensitive to actual off-target risks. The problem of the lack of information and the imbalance of data is thus important to estimate strong and trustworthy AI-based solutions. The issues with off-target effects, variation in editing effectiveness, and data scarcity all demonstrate the necessity to develop superior computational techniques that can acquire predictive representations. The next section discusses how artificial intelligence-driven approaches were invented to resolve these issues in a systematic manner towards improving the accuracy and safety of CRISPR-Cas9 genome editing.

4 AI-DRIVEN METHODOLOGIES AND SOLUTIONS

The development of artificial intelligence has transformed the paradigm of genome editing into much more of a predictive and design-focused process. AI-based solutions to critical issues related to CRISPR-Cas9, especially precision, safety and efficiency, are achieved by facilitating the derivation of informative patterns of complex biological data. In this section, the review is carried out regarding the key AI-based approaches that have been designed to enhance the guide RNA design, minimise off-target risk, work with limited data, and aid in the discovery of new therapeutic targets.

4.1 Enhancing On-Target Efficacy with AI-Powered gRNA Design

The variability of the on-target editing efficiency is one of the key restrictions of CRISPR-Cas9. And in order to solve this problem, AI-based models have been created to forecast and optimise gRNA activity before it can be experimentally implemented [4,14,21,22]. Such models do not rely on trial-and-error techniques to establish empirical dependencies on the performance of editing, but instead use data to find optimisation using sequence features that are known to have an impact.

The notable predictive characteristics are GC content, nucleotide composition at given positions, specifically at the seed region, and gRNA length, which determines a binding affinity and cleavage activity of Cas9 [2,3]. Deep learning models like DeepSpCas9, CRISPR-Net and CRISPRscan combine these functions to model more complicated dependencies in the sequences and get better accuracy than other traditional heuristic-based frameworks [7].

4.2 Mitigating Risk Through Predictive Off-Target Analysis

Among the most important safety issues of CRISPR-based genome editing is the off-target activity[23,24]. To mitigate this risk, AI-based off-target prediction systems have been developed to recognise high-risk gRNAs prior to experimentation. Models like DL-CRISPR and CAF-Net use deep Learning designs to evaluate possible unintended cleavage incidences by examining the mismatched counts, position and complete nucleotide sequence between the gRNA sequences and possible genomic locations [4,6].

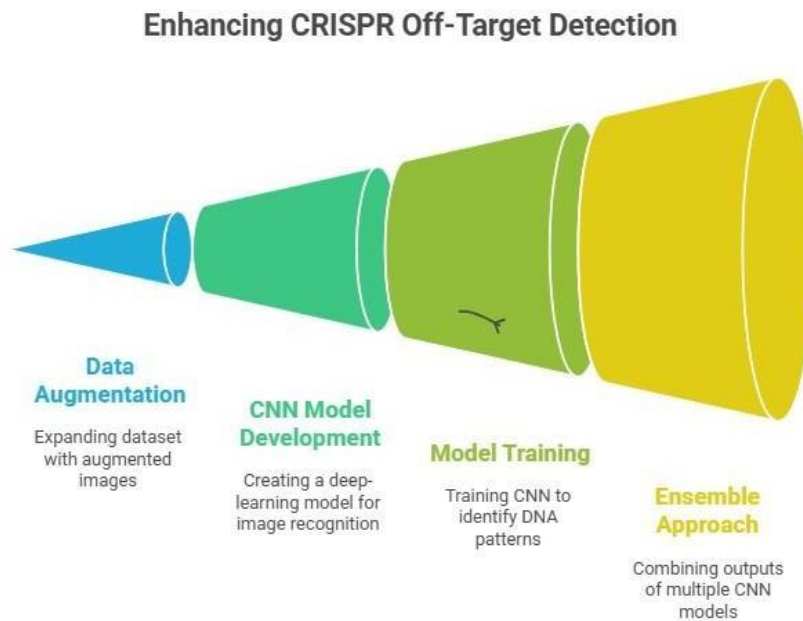


Fig. 2: Enhancing CRISPR off-Target.

These models evaluate a combination of properties and characteristics, such as the amount, location and kind of the mismatches between the sgRNA sequence and the possible off-target DNA sites. They use the results of this analysis to determine the specificity score, which is a quantified measure of the probability of off-target activity [4]. The predictive nature of this capability enables researchers to computationally screen thousands of potential gRNAs and only those with the lowest chances of inducing unwanted mutations, which allows one to create safer and more reliable genome editing strategies.

4.3 Data Scarcity: Dealing with Data Augmentation

The results of the AI-based CRISPR models heavily depend on both the accessibility and the proportion of the training data. In order to address the drawbacks of data scarcity and class imbalance, methods of data augmentation have been proposed to artificially increase training data, especially when using underrepresented positive off-target samples [16].

Some other methods, like autoencoder-based, generate synthetic gRNA sequences through the learning of latent representations of existing samples, thus retaining key sequence features but diversifying samples [5]. Other methods have also been used, such as Borderline-SMOTE, feature space transformations, to improve the representation of minority classes in off-target prediction tasks [4,6]. These methods aid in the building of stronger and more generally predictive models by enhancing the quality and quantity of data and the balance of data.

4.4 Discovery of Novel Therapeutic Strategies with the help of AI

In addition to the optimisation of the functionality of CRISPR-Cas9, AI has facilitated the discovery of new therapeutic opportunities, especially by the introduction of synthetic lethal (SL) interactions. Synthetic lethality is used to define gene pairs where perturbation of one gene is permissible, but simultaneous disruption of both genes causes cell death, which is a targeted approach for cancer therapy [8].

The large-scale analysis of genomic and transcriptomic data based on AI is used to identify cancer-specific SL pairs. Methods like ASTER combine somatic copy number alterations and gene expression data of cancer cohorts and healthy tissues to make inferences about clinically dependent genetic dependencies [8]. These AI-powered plans increase the therapeutic capabilities of CRISPR, as shown in Fig. 3, as they aid in generating hypotheses to target cancer therapies, though a lot more experimental research is needed.

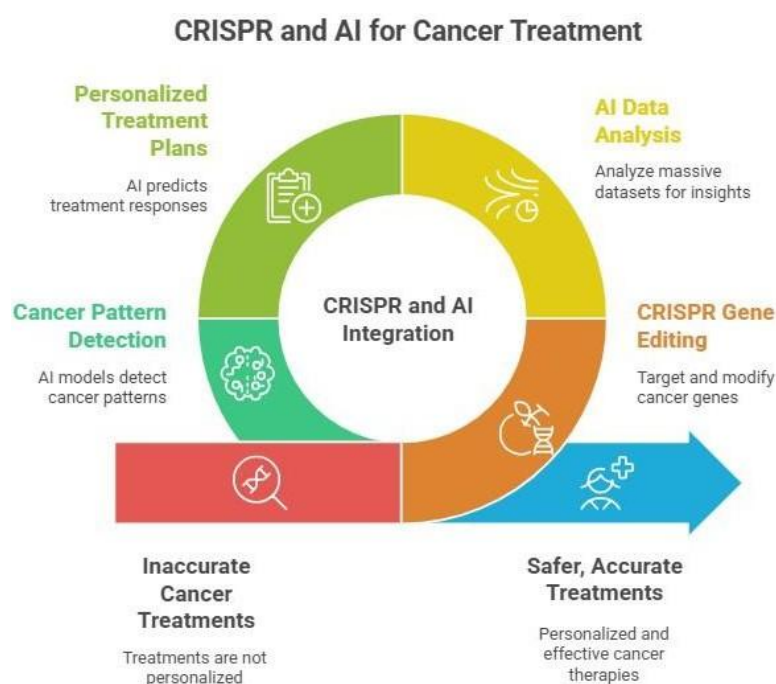


Fig. 3: CRISPR and AI for Cancer Treatment.

4.5 Comparison of AI–CRISPR Approaches

Although a variety of AI-based approaches are suggested to improve precision and efficiency of CRISPR-Cas9, their methodological assumption, data needs and practical preparedness differ significantly. Instead of considering these approaches as unanimous in their performance, strength, and translational capacity, we need to critically compare their performance, strength, and their potential to be translated. Table 1 summarises the methodologies of AI-CRISPR approaches and indicates their comparative advantages, shortcomings, and level of development.

DeepSpCas9 and CRISPR-Net AI-based models of gRNA design are largely based on the sequence-derived features in predicting the on-target efficiency. Although these models can be very predictive with respect to benchmark data, they rely on inputs that are only sequence-based, and this restricts their generalizability to a variety of cellular and genomic contexts. The more complex architectures are better performing, with the disadvantage of adding both computational cost and interpretability. As such, these methods are best applied in experimental design at an early phase and not necessarily in a clinical trajectory [2,7].

The off-target prediction systems, like the DL-CRISPR and CAF-Net, overcome the safety issues of CRISPR-Cas9 with completely different modelling approaches. The main source of class imbalance alleviated by DL-CRISPR is the use of data augmentation strategies to learn using limited populations of positive off-target sites, whereas CAF-Net uses a two-step learning architecture to independently identify datasets of candidate off-target site followed by specificity score prediction. Both methods report increased predictive performance in benchmark sets; however, the effectiveness of the two methods is highly contingent on the quality and representativeness of the training data. There is therefore still concern on the robustness of models in varied genomic settings and the necessity to have increased validation before translating the models [4,6].

Unlike AI-focused optimisation-based AI-CRISPR methods, AI-based interaction discovery to identify synthetic lethality interactions is a discovery-based application in search of new drug targets. Techniques like ASTER use the benefits of large-scale genomic and transcriptomic data to draw deductions about cancer-specific genetic dependencies by analysing them using statistical and machine learning techniques. Although they increase the therapy of CRISPR, beyond guide optimisation, their predictions are based on an indirect inference and population-wide information, thus requiring massive amounts of experimental data. Consequently, synthetic lethality models have now become a hypothesis-generating instrument as opposed to a therapeutic intervention that can be taken at face value [8].

Table1: Comparative overview of AI–CRISPR methods

AI Method	CRISPR Task	Model Type	Input Features	Key Strengths	Key Limitations	Validation
DeepSpCas9	gRNA efficiency prediction	Deep learning (CNN)	Sequence-derived features	High predictive accuracy on benchmark datasets	Limited generalizability; sequence-only context	In silico benchmarks
CRISPR-Net	gRNA efficiency & specificity	Deep learning	Sequence and positional encoding	Improved modelling of mismatch effects	Increased computational complexity; limited interpretability	In silico benchmarks
DL-CRISPR	Off-target prediction	Deep learning with data augmentation	Sequence features; augmented data	Addresses class imbalance; improved specificity	Performance is sensitive to data quality	In silico benchmarks
CAF-Net	Off-target prediction	Two-stage deep learning	Sequence and mismatch features	Explicit separation of detection and scoring	High model complexity; curated data dependence	In silico benchmarks
ASTER	Synthetic lethality discovery	Statistical / ML-based	Genomic & transcriptomic data	Identifies cancer-specific genetic dependencies	Indirect inference requires extensive validation	Retrospective data analysis

5 CHALLENGES AND FUTURE DIRECTIONS

Despite the fact that great advances have been achieved in combining artificial intelligence with CRISPR-Cas9 technology, there are still enablers before the full therapeutic and scientific potential of the technology is achieved. Addressing these challenges, technical innovation will be required, but with consideration of ethical, regulatory implications and societal effects. This section will explain technical constraints that have remained, ethical issues that are likely to emerge, and future research areas that will probably define the next phase of AI-assisted genome editing.[25]

5.1 Persistent Technical and Computational Hurdles

Although there are recent developments, AI -CRISPR technologies still have a number of technical and computational impediments on the way to translating laboratory discoveries to clinical application. Deliberate and secure delivery of CRISPR-Cas9 components to a certain target tissue in vivo is one of the primary challenges. Therapeutic genome editing still requires the requirement of reliable delivery, which limits clinical adoption [7].

The other significant limitation is the model's generalizability. Most AI-based CRISPR models need data of a particular cell type or experimental condition, which may restrict their application in a variety of biological settings [5,7]. Models that have been shown to behave well on benchmark datasets might not be able to define new genomic or cellular environments.

The interpretability of models is also a major challenge. Deep learning architecture is frequently a black-box system, and it is frequently challenging to understand the biological rationale behind its predictions. This lacks experimental trust, clinical adoption, and regulatory approval, which explains why explainability and reproducibility are fundamental [7].

5.2 The Ethical, Safety, and Regulatory Issues

The growing power of genome editing technologies prompts significant ethical and safety issues that should be handled in a responsible manner. Among them, there is the possibility of heritable germline alterations, and the unintended genetic outcomes in the long-term and fair access to high-technology gene-editing procedures [7]. These issues require proper

management and open governance systems.

The aspect of introducing AI also complicates the ethical environment. Even though AI is able to make safety more predictive and assess risks, it presents a range of issues like data privacy, bias in algorithms, and AI-assisted decision-making accountability. It will be necessary to set up strong regulatory frameworks to consider both genome editing and AI-based practices to provide safe and ethical clinical deployment [7].

5.3 Future Projections and Future Opportunities

To sum up, it can be stated that some research directions will further contribute to AI-assisted genome editing. Further growth of AI models, such as transformer networks and reinforcement learning, can allow making more precise predictions and optimising genome-editing programs more adaptively in the dynamic environments of the biological field [7].

- The other direction of significance is the integration of multi-omics data. With the incorporation of epigenomic, transcriptomic, proteomic and chromatin accessibility data, the next generation AI might be more context-aware and predictive with higher accuracy than sequence-based approaches [7].
- Outside of the immediate uses, preliminary studies have already commenced to observe the convergence of AI and CRISPR technologies on newer technology platforms like single-cell sequencing and bio-nanotechnologies. All the mentioned genome engineering fields, like DNA-based data storage and bio-cyber interfaces, demonstrate the broadening of the genome engineering research scope and the necessity of interdisciplinary cooperation and responsible evaluation[9,11].

6 CONCLUSION

The combination of artificial intelligence and CRISPR-Cas9 has introduced a paradigm shift in the understanding of genome editing studies by enhancing precision, efficiency, and safety. As this review has described, AI methodologies have been used in response to some of the major shortcomings that still limit the clinical translation of CRISPR technologies, such as prediction of off-target risk, optimisation of guide RNA, and identification of novel therapeutic targets. AI can provide more systematic and reliable plans for genome editing by converting complex and high-dimensional biological data into predictive and actionable insights.

However, the effective translation of AI-assisted CRISPR technologies into clinical practice extends beyond successful work on a technical level. Further progress in this field will be based on the long-term interdisciplinary collaboration between the biologists, computational scientists, clinicians, ethicists, and regulatory organisations. The challenges related to the strength of models, their explainability, data control, and morals will have to be resolved in order to guarantee that the development and implementation are accountable.

Overall, AI-CRISPR convergence is an emerging and promising system of genome engineering. More methodological development, strict experimentation validation and free interaction among the controlling bodies will be of paramount importance in making sure that this potential is fully realised in biomedical research and application in the therapeutic field.

References

- [1] J. Shukla, B. Pant, N. Tufchi, G. C. Paul, K. Pant, M. Pant, and S. Sihna, "CRISPR Cas9 genome editing approach for gRNA of the genes associated with schizophrenia: A computational approach," in *Proc. 2022 2nd Int. Conf. Advance Computing and Innovative Technologies in Engineering (ICACITE)*, 2022, pp. 1957–1962.
- [2] N. Kumar S, A. Hariharan, A. B. H, B. K., and P. Basavaraj, "CRISPR-Cas9 guide RNA designer using Python," in *Proc. 2023 7th Int. Conf. Trends in Electronics and Informatics (ICOEI)*, 2023, pp. 1350–1356.
- [3] D. N. Kumar, D. L. Chowdhary, T. Pathuri, P. Katta, and L. Arya, "AI enhanced smart genome editing: Integration of CRISPR-Cas9 with artificial intelligence for cancer treatment," in *Proc. 2024 2nd Int. Conf. Power, Energy, Control and Transmission Systems (ICPECTS)*, 2024, pp. 1–7.
- [4] Y. Zhang, Y. Long, R. Yin, and C. K. Kwok, "DL-CRISPR: A deep learning method for off-target activity prediction in CRISPR/Cas9 with data augmentation," *IEEE Access*, vol. 8, pp. 78013–78022, 2020.
- [5] H. Chen and Z. Jiang, "CrisprDA: A data augmentation method enhancing the efficiency of sgRNA on-target activity prediction," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, pp. 1–11, 2025.
- [6] T. Zhang, B. Jiang, G. Yang, Z. Bao, and J. Peng, "Predicting CRISPR-Cas9 off-target activities with a two-stage deep learning framework," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, pp. 1–5, 2025.
- [7] G. L. Sagar, R. A. H., and R. Chakravaradhan, "Revolutionizing genetic manipulation: Artificial intelligence-powered precision in CRISPR-Cas9 gene editing," *IEEE Access*, vol. 11, pp. 98765–98778, 2023.
- [7] G. L. Sagar, R. A. H., and R. Chakravaradhan, "Revolutionizing genetic manipulation: Artificial intelligence-powered precision in CRISPR-Cas9 gene editing," in *Proc. 2023 Int. Conf. Disruptive Technologies (ICDT)*, 2023, pp. 412–421.
- [8] H. Liany *et al.*, "ASTER: A method to predict clinically relevant synthetic lethal genetic interactions," *IEEE J. Biomed. Health Inform.*, vol. 28, no. 3, pp. 1391–1402, 2024.
- [9] U. A. K. Chude-Okonkwo and A. V. Vasilakos, "CRISPR-enabled graphene-based bio-cyber interface model for in vivo monitoring of non-invasive therapeutic processes," *IEEE Trans. Nanobiosci.*, vol. 23, no. 2, pp. 300–313, 2024.
- [10] [10] M. Chen, "Hiding secret messages in living cells in the age of next-generation sequencing," *IEEE Access*, vol. 6, pp. 12115–12127, 2018.
- [11] M. R. Abbasy and B. Shanmugam, "Enabling data hiding for resource sharing in cloud computing environments based on DNA sequences," *IEEE Access*, vol. 7, pp. 112345–112356, 2019.
- [12] G. L. Sagar, R. A. H., and R. Chakravaradhan, "Revolutionizing genetic manipulation: Artificial intelligence-powered precision in CRISPR-Cas9 gene editing," *IEEE Access*, vol. 11, pp. 98765–98778, 2023.
- [11] N. N. Orlova, M. G. Gladkova, G. A. Ashniev, and A. V. Orlov, "Fluorescently controlled investigation of super-enhancers with CRISPR interference and CRISPR prime-editing systems," *IEEE Access*, vol. 10, pp. 118945–118955, 2022.
- [12] G. Chai, M. Yu, L. Jiang, Y. Duan, and J. Huang, "HMMCAS: A web tool for the identification and domain annotations of CAS proteins," *IEEE Access*, vol. 8, pp. 221345–221355, 2020.
- [15] Z. S. Ameen, A. U. Ibrahim, A. S. Mubarak, and F. Al-Turjman, "An improved CNN-LSTM deep model for classification of guide RNA in CRISPR-Cas12 system," *IEEE Access*, vol. 9, pp. 165432–165444, 2021.
- [16] H. Chen and Z. Jiang, "CrisprDA: A data augmentation method enhancing the efficiency of sgRNA on-target activity prediction," *Bioinformatics*, vol. 38, no. 4, pp. 1–9, 2022..
- [17] J. A. Doudna and E. Charpentier, "The new frontier of genome engineering with CRISPR-Cas9," *Science*, vol. 346, no. 6213, 2014.
- [18] J. Zou *et al.*, "A primer on deep learning in genomics," *Nature Methods*, vol. 16, no. 4, pp. 337–348, 2019.
- [19] Y. Fu *et al.*, "High-frequency off-target mutagenesis induced by CRISPR-Cas nucleases in human cells," *Nature Biotechnology*, vol. 31, no. 9, pp. 822–826, 2013.
- [20] S. Q. Tsai *et al.*, "GUIDE-seq enables genome-wide profiling of off-target cleavage by CRISPR-Cas nucleases," *Nature Biotechnology*, vol. 33, no. 2, pp. 187–197, 2015.
- [21] H. K. Kim *et al.*, "SpCas9 activity prediction by DeepSpCas9," *Genome Biology*, vol. 20, article 191, 2019.
- [22] G. Chuai *et al.*, "DeepCRISPR: optimized CRISPR guide RNA design by deep learning," *Bioinformatics*, vol. 34, no. 1, pp. 1–9, 2018.
- [23] X. Zhang *et al.*, "DL-CRISPR: a deep learning method for off-target activity prediction," *Bioinformatics*, vol. 35, no. 18, pp. 1–8, 2019.
- [24] J. Lin *et al.*, "CRISPR-Net: a recurrent convolutional network quantifies CRISPR off-target activities," *Bioinformatics*, vol. 34, no. 1, pp. 1–8, 2018.
- [25] National Academies of Sciences, Engineering, and Medicine, *Human Genome Editing: Science, Ethics, and Governance*, National Academies Press, Washington, DC, 2017.