Proceedings of the 4th International Conference on Biological Engineering and Medical Science
DOI: 10.54254/2753-8818/73/2024.19820

The future of protein assembly: A deep learning paradigm for efficient and accurate data processing

Haodi Chen

School of Biotechnology, Northeast Forestry University, Harbin, China

hd2475622099@gmail.com

Abstract. Protein assembly is critical to understanding biomolecular functions and biological processes. In recent years, with the large increase in protein sequence data, the demand for data integration has also increased, and deep learning has made significant progress in protein structure prediction and functional analysis. Deep learning is an efficient data processing method that helps traditional experiments improve the efficiency, speed, and accuracy of data processing. Combined with deep learning, the development of new proteins is no longer limited by experimental conditions, which is of great significance for drug target prediction, material design, and new drug discovery. The purpose of this research was to investigate a deep learning-driven approach for protein assembly and assess its effectiveness in forecasting the three-dimensional conformation and functional properties of proteins. By using the latest deep learning algorithms and large-scale protein databases, the potential of this method to improve the accuracy and efficiency of protein assembly is demonstrated.

Keywords: deep learning, protein assembly, protein structure prediction, bioinformatics, artificial intelligence.

1. Introduction

Proteins are essential molecules in life, with their function determined by their three-dimensional structure. The amino acid sequence of each protein is determined by the nucleotide sequence of the gene that encodes it. Proteins typically adopt a specific three-dimensional shape, referred to as the native state. While many proteins are capable of folding autonomously based on their amino acid sequences, several proteins require assistance from chaperones to achieve proper folding. Traditional protein structure prediction methods depend on experimental approaches, which are both expensive and time-consuming. However, with advancements in computing technology, computational prediction methods have become a prominent area of research. As a key branch of artificial intelligence, Deep learning has made remarkable advancements in the fields of image recognition and natural language processing. Utilizing deep learning in protein assembly is expected to significantly improve both the precision and efficiency of predicting protein structures.

© 2025 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/73/2024.19820

2. Literature Review

2.1. Traditional methods for protein structure prediction

2.1.1. Electron Microscopy

With the continuous development of science and technology, the study of biological science is gradually deepening, and the study of the structure and function of protein, as the most critical molecule in life activities, is of great significance. In recent years, cryo-EM technology, as a scientific research tool, has provided strong support for the analysis of protein structures, enabling scientists to unravel the mystery of proteins and further understand the nature of life phenomena[1] Cryo-EM, short for Cryo-Electron Microscope (Cryo-EM),is a high-resolution three-dimensional microscope whose greatest feature is the ability to observe samples at low temperatures to maintain their native state[1].Cryo-EM has a wide range of applications, including biology, biophysics, chemistry, and other fields. In bioscience research, cryo-EM offers unprecedented possibilities for protein structure elucidation. New advances in the synthesis of green nanoparticles conserve natural resources and reduce environmental pollution. The current study uses Anastatica hierochuntica L. and mugwort absinthe extract to prepare silver nanoparticles (AgNPs), detect phytochemicals by FTIR and GC-MS, evaluate their antimicrobial and cytotoxic effects on bacteria and cancer cells, and investigate synergistic effects with antibiotics[1]. The conversion of silver ions by plant extracts was examined using dynamic light scattering, zeta potential, and transmission electron microscopy, and the morphological changes of microorganisms were observed by scanning electron microscopy. The findings of the study indicated that AgNPs exhibited substantial antimicrobial activity and cytotoxicity against a broad spectrum of microorganisms. It is postulated that the biosynthesis of these AgNPs may involve the participation of various biomolecules, such as polysaccharides, proteins, and phenolic compounds. The study confirms that the use of AgNO3 to prepare plant-based AgNPs is a cost-effective method and that the combination with antibiotics has shown potential for the treatment of bacterial infections[1]. Besides sperm cells, semen contains small membrane vesicles, such as prostate bodies, which may influence immune cell activity in the female reproductive tract and enhance sperm motility and function. How the prostatic body mediates diverse functions is unclear. Studies have shown that vesicles in seminal plasma are actually heterogeneous mixtures produced by different glands and secretory mechanisms. Two vesicles of different sizes but similar buoyancy densities were isolated from seminal plasma obtained from vasectomized men. GLIPR2 was enriched in smaller vesicles, annexin A1 was enriched on the surface of larger vesicles, and PSCA protein was present in both vesicles but differed between individuals. By electron microscopy analysis, the characteristics and distribution of different vesicles were compared using a variety of antibody labeling and imaging techniques, and it was found that the characteristics of the prostate body had significant individual differences[2].

2.1.2. Nuclear Magnetic Resonance, NMR

Nuclear magnetic resonance (NMR) spectroscopy, also called nuclear magnetic resonance, is an analytical technique and theory that utilizes nuclear magnetic resonance phenomena to determine the microstructure of substances[3]. Matter is made up of atoms, quantum mechanics research has found that the nucleus of some atoms has both nuclear magnetic moment and nuclear spin brought about by angular momentum, so under a strong static magnetic field, nuclear magnetic resonance phenomenon will occur with radio frequency electromagnetic waves, and produce radio frequency electromagnetic spectrum feedback that can reflect its internal structure, that is, nuclear magnetic resonance spectrum. The resulting NMR spectra are analyzed and can be adjusted for sample preparation, selection, or design of specific RF pulse sequences to obtain specific information. A total of 162 metabolomics measurements were analyzed in T2D patients in four cohort studies and one replication cohort using an NMR-based approach. Linear and logistic regression analyses were conducted to account for potential confounding factors, followed by meta-analyses to examine the relationship between these metabolic parameters and hemoglobin A1c levels, six categories of glucose-lowering medications, and insulin

DOI: 10.54254/2753-8818/73/2024.19820

initiation (n=698) over the course of the 7-year follow-up period[4]. Reactive oxygen and nitrogen species (ROS and RNS) play essential roles in cell signaling, yet their kinetics are challenging to study due to the intricacies of spatial and temporal regulation. Although several techniques exist to assess ROS, many are inadequate in directly pinpointing its subcellular localization. Electron paramagnetic resonance (EPR) spectroscopy is a highly effective method for examining ROS dynamics across various biological samples and cellular compartments, especially in muscle function research. This paper explores spin trapping of different ROS, EPR detection of nitric oxide, and EPR's capacity to identify stable free radical environments. Despite EPR's distinct ability to offer insights into free radicals, its use in muscle physiology remains limited[5].

2.1.3. Principles and applications of X-ray crystallography

X-ray crystallography is a robust analytical method for determining the atomic and molecular structures of crystals[6]. By measuring the angle and intensity of diffracted X-rays, a three-dimensional electron density map of the crystal is generated. This data is then used to determine the atomic positions and chemical bonds within the crystal, and various other structural parameters[7]. X-ray crystallography has played an important role in the development of numerous scientific fields such as chemistry, biology, and materials science. This technique involves irradiating X-rays onto crystals and analyzing the patterns of diffracted rays to gain insight into the structure of the crystal. This method's development began in the early 20th century and has significantly advanced since then, enabling the detailed study of complex biomolecules such as proteins and nucleic acids[8]. The process of X-ray crystallography can be broken down into several key steps, the first step is crystal preparation, and obtaining pure, ordered crystals is essential for accurate analysis. In the second step, real-time PCR is used to determine the specific data of the DNA[8]. The third step involved purification and expression, using a cDNA library from mRNA of mouse bone marrow-derived macrophages as the PCR amplification template, cloning and expressing a His-tagged mouse latex protein, and purifying it via metal affinity chromatography to validate the selenomethionine (SeMet)-tagged protein. In the fourth step, latex and SeMet latex crystals, obtained via the hanging drop vapor diffusion method, were cryoprotected, and X-ray diffraction data were collected at an advanced light source, the data were processed and the structure was solved by HKL2000, SOLVE and ARP/WARP, and finally the structure of the latex protein was refined by the maximum likelihood method and a single B factor, and a model containing latex residues 1-217 and some tag residues was obtained. Finally, modeling analysis was carried out, CLUSTALX was used for protein alignment, and MODELLER was used to construct the C-terminal region model of latex protein sequence of rat, mouse and human TIG1, followed by the highest refinement in HOMOLOGY, the stereochemical quality of the model was verified by PROCHECK, and the docking analysis of pentaccharide binding was performed by GRID and GOLD, and finally the binding site was optimized according to the GoldScore fitness function[8]. Visible X-ray crystallography remains a cornerstone technique for structural analysis, providing key insights into the molecular structure of various substances. Continuous advancements in technology and methods are expected to further enhance its functionality and applications.

2.2. Application of deep learning in drug-target interaction prediction

2.2.1. Drug-Target Interaction (DTI)

Predicting Drug-Target Interaction (DTI) is a crucial part of the drug discovery process. Accurately predicting the interactions between drugs and their potential targets can significantly accelerate the development of new drugs and reduce development costs[9]. Traditional experimental methods, while accurate, are costly and time-consuming. With the development of computer technology, the application of computational methods, especially deep learning, in DTI prediction has attracted extensive attention[10]. Deep learning uses complex neural network models to learn the implicit relationships between targets and drugs from large amounts of data, demonstrating powerful predictive capabilities.

DOI: 10.54254/2753-8818/73/2024.19820

2.2.2. Benefits

Traditional methods rely on hand-designed features and often struggle to cover complex biological information. Deep learning models can automatically extract higher-order features from raw data, capture complex nonlinear relationships, and improve prediction accuracy. Deep learning algorithms are capable of efficiently processing massive amounts of data. With the explosive growth of biomedical data, deep learning methods can fully utilize these data for training and enhance model generalization. These models come in various forms, including convolutional neural networks (CNNs), recurrent neural networks (RNNs), graph neural networks (GNNs), and more. which can be selected and adjusted according to different task requirements to achieve the best prediction results.

2.3. Application of deep learning in bioinformatics

2.3.1. Deep Neural Networks, DNNs

The eukaryotic genome contains essential regions and signals, such as promoters, enhancers, transcription factor binding sites, translation start sites, splice sites, and polyadenylation signals (PAS), all critical for gene regulation. While numerous machine learning (ML) and deep learning (DL) models have been developed to predict these signals, there is still room for improvement in accuracy. In this study, researchers focused on the computational identification of human PAS in genomic DNA and introduced a novel method. This approach integrates 12 deep neural networks (DNNs) with logistic regression models (LRM) into a hybrid model named HybPAS. By combining signal processing, statistical techniques, and DNA sequence features, HybPAS achieves an average accuracy of 91.22%, surpassing the performance of current state-of-the-art models (Omni-PolyA, DeepGSR, DeeReCT-PolyA), significantly enhancing PAS prediction accuracy. DNNs can also predict mutational status by analyzing H&E-stained cancer slides, enabling cost-effective and timely precision oncology studies. Using weakly supervised learning, we trained DNNs to predict BRAF V600E mutation in thyroid cancer without regional annotation. The area under the receiver operating characteristic curve in the independent external cohort was 0.98, surpassing the result from strongly supervised training. We've also developed visualization technology that automatically highlights key areas. The t-test confirmed the differences in histological characteristics between mutant and wild-type patients. Weakly supervised learning shows great potential in DNN model training[12].

2.3.2. Convolutional Neural Network (CNN)

Convolutional neural networks (CNNs) are deep learning models built to process mesh-like data structures, such as images. Introduced by Yann LeCun in the late 1980s, CNNs revolutionized fields like computer vision, natural language processing, and drug discovery. They excel in tasks involving image and sequence data by automatically and adaptively learning the spatial hierarchy of features from input data. Drug molecules and proteins can be represented as graph structures, and GNNs are able to efficiently capture the information of nodes and edges in graph structures[13]. In DTI prediction, GNNs predict interactions by extracting features from molecular and protein maps. In plant evolutionary history, alterations in cis-regulatory elements (CREs) have played a key role in driving the diversification of gene expression, contributing to the evolution of lineage-specific traits. However, predicting the behavior of CRE patterns remains a complex task. This study employed a cistrome dataset along with an interpretable convolutional neural network (CNN) framework to forecast genome-wide expression patterns based on tomato (Solanum lycopersicum) DNA sequences. By applying single-cell spatiotemporal transcriptome data, a prediction model for key expression patterns during early tomato fruit ripening was developed. CNN analysis identified critical nucleotide residues for each gene expression pattern, which were experimentally validated. This method not only enhances understanding of CRE-regulatory networks and transcription factor interactions but also offers a strategy to optimize gene expression design[14].

DOI: 10.54254/2753-8818/73/2024.19820

2.3.3. Recurrent Neural Networks (RNNs)

Recurrent neural networks (RNNs) are a class of artificial neural networks specifically designed to recognize patterns in sequential data, including time series, natural language, and video frames. Their natural aptitude for managing sequences makes them particularly well-suited for tasks that involve ordered or time-dependent data[15]. Unlike traditional neural networks, RNNs have directed loop connections that allow them to retain memory of previous inputs, making them particularly useful for tasks where context and order matter. Neural networks in machine learning began with Rosenblatt's perceptron (1958) and Minsky and Papert's work (2017). Traditional learning methods struggle with raw data due to the need for manual feature design. In contrast, deep learning can automatically discover effective features from data, eliminating the need for complex feature engineering. Despite this, neural networks have not been the preferred method for machine learning for more than half a decade and have been surpassed by many alternatives. With the rise of powerful hardware and numerous fast processors (e.g., GPUs), along with access to abundant training data, deep neural networks (DNNs) have recently achieved outstanding performance in various machine learning applications, particularly in computer vision, natural language processing, and complex board games like Go. The importance of protein-RNA binding in biology has made it a key area of research for experimentalists and machine learning researchers. High-throughput measurement techniques, methods like CLIP and its derivatives are used to measure protein-RNA binding in vivo on a transcriptome-wide scale. However, these experiments are affected by multiple factors, leading to high noise and predominantly binary outcomes (binding present or absent). Learning protein-RNA binding preferences from this data remains challenging due to the complexity of the in vivo environment and experimental noise[16].

2.4. Current research on protein structure prediction based on deep learning

2.4.1. AlphaFold

AlphaFold, an AI system developed by DeepMind, a subsidiary of Alphabet, aims to solve the complex problem of protein structure prediction. Proteins are vital biomolecules in living organisms with various functions, all dictated by their three-dimensional structure. Accurately predicting protein structure from its amino acid sequence has long been a significant challenge in computational biology. AlphaFold has made significant advances in this area, achieving breakthrough results that have far-reaching implications for biology and medicine. AlphaFold utilizes a deep neural network that takes a protein's amino acid sequence as input and predicts its 3D structure. Trained on known structures from databases like the Protein Database (PDB), AlphaFold employs an attention mechanism, allowing the network to focus on specific parts of the protein sequence, which aids in capturing long-distance amino acid interactions necessary for accurate folding. Immune receptor proteins are crucial in the immune system and have significant potential in biotherapeutics, making structural understanding key to determining antigen-binding properties. This article presents ImmuneBuilder, a suite of deep learning models designed to precisely predict the structures of antibodies (ABodyBuilder2), nanobodies (NanoBodyBuilder2), and T cell receptors (TCRBuilder2). Research demonstrates that ImmuneBuilder achieves state-of-the-art accuracy while offering significantly faster performance compared to AlphaFold2[17]. For example, ABodyBuilder2 predicted an RMSD of 2.81Å for the CDR-H3 loop in a benchmark of 34 antibodies, which is 0.09Å higher than AlphaFold-Blender and more than 100 times faster. NanoBodyBuilder2 predicted an average RMSD of 2.89 Å for the nanoantibody CDR-H3 loop, which was 0.55 Å higher than that of AlphaFold2, and similar results were achieved for the prediction of T cell receptors. In addition, ImmuneBuilder gives an error estimate of the final prediction for each residue by predicting a set of structures. ImmuneBuilder is a collection of deep learning models that predict the structures of antibodies, nanobodies, and T-cell receptors with top-tier accuracy, while being significantly faster than AlphaFold2 and AlphaFold-Multimer[17].

DOI: 10.54254/2753-8818/73/2024.19820

2.4.2. RoseTTAFold

Oomycete and fungal interactions with plants can range from neutral to symbiotic or pathogenic, influencing plant health and overall fitness. They colonize hosts by producing effector proteins that modify stress pathways, developmental processes, and immune responses, benefiting the microorganisms. Bioinformatics and experimental approaches help investigate the roles of these effectors in plant-microbial interactions. RoseTTaffold and AlphaFold2 have advanced protein 3D structure prediction from amino acid sequences using machine learning, though both depend on supercomputers. Google Colabfold provides a more user-friendly alternative. This article explores the structural biology, sequence motifs, and domain knowledge of filamentous microbial effectors, and discusses AlphaFold2 and RoseTTafold's applications in effector biology[18]. The results indicated that PDB ID 6hug.1.B had the highest sequence homology (42.17%) with RmGABACl, but its amino acid sequence was 210 residues shorter, complicating homology modeling. As a result, a complete RmGABACl model was constructed using de novo design and compared with the homologous model. The models produced by SWISS-MODEL, RoseTTAFold, and TrRosetta were as follows: SWISS-MODEL yielded a Prosaweb Z score of −3.27, an ERRAT value of 77.79, with 80.6% of residues in the favorable region. RoseTTAFold achieved a Prosaweb Z score of −7.06, ERRAT value of 92.218, and 88.7% of residues in the favorable region, though with accuracy defects. TrRosetta had a Prosaweb Z score of −4.92, ERRAT value of 91.82, and 94.2% of residues in the favorable region. Overall, the TrRosetta model proved most reliable for further study[19].

3. Results

X-ray crystallography provides high-resolution three-dimensional structural information and has been widely used in chemistry, biology, and materials science. Nuclear magnetic resonance (NMR) is used to determine the microstructure of matter, particularly proteins and complex molecules. Cryo-electron microscopy (Cryo-EM) enhances the ability to elucidate protein structures by preserving samples at low temperatures in their native state. Deep learning models automatically extract higher-order features from large datasets, improving the accuracy of drug-target interaction predictions. While traditional methods are precise, they are expensive and time-consuming, whereas deep learning offers efficiency and cost-effectiveness. Deep neural networks (DNNs) are also used to predict key genome signals, such as promoters, enhancers, and transcription factor binding sites, significantly improving the prediction accuracy. Convolutional neural networks (CNNs) efficiently process image and sequence data and can be applied to predict gene expression patterns and drug molecule-protein interactions. Recurrent neural networks (RNNs) are ideal for processing time series and sequential data, showing strong potential in learning protein-RNA binding preferences. AlphaFold, using a deep neural network, accurately predicts protein 3D structures, with profound implications in biology and medicine. ImmuneBuilder is both more accurate and faster than AlphaFold2 for predicting the structures of antibodies, nanobodies, and T-cell receptors. RoseTTAFold provides an easy-to-use alternative for predicting protein structure through machine learning algorithms, especially for the study of filamentous microbial effector proteins.

4. Discussion

Traditional methods like X-ray crystallography, NMR, and cryo-EM have been successful in determining protein structures, but they are often expensive and time-consuming. Deep learning approaches, including DNNs, CNNs, and RNNs, address these challenges by automatically extracting features from large datasets, enhancing both prediction accuracy and efficiency. These models show significant potential in predicting drug-target interactions, significantly improving prediction performance by automatically extracting higher-order features and efficiently processing massive amounts of data[20]. In addition, the application of deep learning in bioinformatics, such as genomic signal prediction, gene expression pattern prediction, and protein-RNA binding preference learning, further demonstrates its advantages in processing complex biological data. In current deep learning-based protein structure prediction research, AlphaFold, ImmuneBuilder, and RoseTTAFold represent state-of-the-art technologies. AlphaFold accurately predicts protein structure through deep neural

DOI: 10.54254/2753-8818/73/2024.19820

networks, while ImmuneBuilder and RoseTTAFold demonstrate greater precision and speed in specific application areas such as immune receptor proteins and filamentous microbial effector proteins.

5. Conclusion

In closing, applying deep learning to protein structure prediction, drug-target interaction prediction and bioinformatics has greatly promoted the development of life science research. Traditional methods are still important tools for structural elucidation, but deep learning methods offer a more efficient, cost-effective, and accurate alternative. Advanced models such as AlphaFold, ImmuneBuilder, and RoseTTAFold demonstrate the great potential of deep learning in protein structure prediction, and future research and applications will continue to benefit from the advancements of these innovative technologies.

References

[1] Aabed, K., & Mohammed, A. E. (2021). Synergistic and Antagonistic Effects of Biogenic Silver Nanoparticles in Combination With Antibiotics Against Some Pathogenic Microbes. Frontiers in Bioengineering and Biotechnology, 9. https://doi.org/10.3389/fbioe.2021.652362

[2] Aalberts, M., van Dissel-Emiliani, F. M. F., van Adrichem, N. P. H., van Wijnen, M., Wauben, M. H. M., Stout, T. A. E., & Stoorvogel, W. (2012). Identification of Distinct Populations of Prostasomes That Differentially Express Prostate Stem Cell Antigen, Annexin A1, and GLIPR2 in Humans1. Biology of Reproduction, 86(3). https://doi.org/10.1095/biolreprod.111.095760

[3] A Keniry M;Smith R. (2024). A 13C NMR spin-lattice relaxation study of the interaction of myelin proteins with lipid vesicles. Biophysical Chemistry, 12(1). https://pubmed.ncbi.nlm.nih.gov/17000147/

[4] Leen, Vogelzangs, N., Mook-Kanamori, D. O., Brahimaj, A., Nano, J., Amber, Dijk, van, Slieker, R. C., Steyerberg, E. W., M. Arfan Ikram, Beekman, M., Boomsma, D. I., Cornelia, P. Eline Slagboom, Coen, Schalkwijk, C. G., Arts, W., Dekker, J. M., Dehghan, A., & Muka, T. (2018). Blood Metabolomic Measures Associate With Present and Future Glycemic Control in Type 2 Diabetes. The Journal of Clinical Endocrinology & Metabolism, 103(12), 4569–4579. https://doi.org/10.1210/jc.2018-01165

[5] A. Abdel-Rahman, E., Mahmoud, A. M., Khalifa, A. M., & Ali, S. S. (2016). Physiological and pathophysiological reactive oxygen species as probed by EPR spectroscopy: the underutilized research window on muscle ageing. The Journal of Physiology, 594(16), 4591–4613. https://doi.org/10.1113/jp271471

[6] Aalbergsjø, S. G., & Einar Sagstuen. (2015). New Evidence for Hydroxyalkyl Radicals and Light- and Thermally Induced Trapped Electron Reactions in Rhamnose. Radiation Research, 184(2), 161–161. https://doi.org/10.1667/rr14081.1

[7] Aalbers, F., Turkenburg, J. P., Davies, G. J., Dijkhuizen, L., & Lammerts van Bueren, A. (2015). Structural and Functional Characterization of a Novel Family GH115 4-O-Methyl-α-Glucuronidase with Specificity for Decorated Arabinogalactans. Journal of Molecular Biology, 427(24), 3935–3946. https://doi.org/10.1016/j.jmb.2015.07.006

[8] Aagaard, A., Listwan, P., Cowieson, N., Huber, T., Ravasi, T., Wells, C. A., Flanagan, J. U., Kellie, S., Hume, D. A., Kobe, B., & Martin, J. L. (2005). An Inflammatory Role for the Mammalian Carboxypeptidase Inhibitor Latexin: Relationship to Cystatins and the Tumor Suppressor TIG1. Structure, 13(2), 309–317. https://doi.org/10.1016/j.str.2004.12.013

[9] You, J., McLeod, R. D., & Hu, P. (2019). Predicting drug-target interaction network using deep learning model. Computational Biology and Chemistry, 80, 90–101. https://doi.org/10.1016/j.compbiolchem.2019.03.016

[10] Abbasi Mesrabadi, H., Faez, K., & Pirgazi, J. (2023). Drug–target interaction prediction based on protein features, using wrapper feature selection. Scientific Reports, 13(1). https://doi.org/10.1038/s41598-023-30026-y

[11] Albalawi, F., Chahid, A., Guo, X., Albaradei, S., Magana-Mora, A., Jankovic, B. R., Uludag, M. , Van Neste, C., Essack, M., Laleg-Kirati, T.-M., & Bajic, V. B. (2019). Hybrid model for efficient prediction of poly(A) signals in human genomic DNA. Methods, 166, 31–39. https://doi.org/10.1016/j.ymeth.2019.04.001

[12] Anand, D., Yashashwi, K., Kumar, N., Rane, S., Gann, P. H., & Sethi, A. (2021). Weakly supervised learning on unannotated H&E-stained slides predicts BRAF mutation in thyroid cancer with high accuracy. The Journal of Pathology, 255(3), 232–242. https://doi.org/10.1002/path.5773

[13] El_Rahman, S. A., & Ala Saleh Alluhaidan. (2024). Enhanced multimodal biometric recognition systems based on deep learning and traditional methods in smart environments. PloS One, 19(2), e0291084–e0291084. https://doi.org/10.1371/journal.pone.0291084

[14] Akagi, T., Masuda, K., Kuwada, E., Takeshita, K., Taiji Kawakatsu, Tohru Ariizumi, Kubo, Y., Ushijima, K., & Uchida, S. (2022). Genome-wide cis-decoding for expression design in tomato using cistrome data and explainable deep learning. The Plant Cell, 34(6), 2174–2187. https://doi.org/10.1093/plcell/koac079

[15] Abbaspour, S., Fotouhi, F., Sedaghatbaf, A., Fotouhi, H., Vahabi, M., & Linden, M. (2020). A Comparative Analysis of Hybrid Deep Learning Models for Human Activity Recognition. Sensors, 20(19), 5707. https://doi.org/10.3390/s20195707

[16] Ben-Bassat, I., Chor, B., & Orenstein, Y. (2018). A deep neural network approach for learning intrinsic protein-RNA binding preferences. Bioinformatics, 34(17), i638–i646. https://doi.org/10.1093/bioinformatics/bty600

[17] Brennan Abanades, Wing Ki Wong, Boyles, F., Georges, G., Bujotzek, A., & Deane, C. M. (2023). ImmuneBuilder: Deep-Learning models for predicting the structures of immune proteins. Communications Biology, 6(1). https://doi.org/10.1038/s42003-023-04927-7

[18] Amoozadeh, S., Johnston, J., & Meisrimler, C.-N. (2021). Exploiting Structural Modelling Tools to Explore Host-Translocated Effector Proteins. International Journal of Molecular Sciences, 22(23), 12962. https://doi.org/10.3390/ijms222312962

[19] Ayub, S., Malak, N., Cossío-Bayúgar, R., Nasreen Nasreen, Khan, A., Niaz, S., Khan, A., Alanazi, A. D., & Mourad Ben Said. (2023). In Vitro and In Silico Protocols for the Assessment of Anti-Tick Compounds from Pinus roxburghii against Rhipicephalus (Boophilus) microplus Ticks. Animals, 13(8), 1388–1388. https://doi.org/10.3390/ani13081388

[20] Abbasi, K., Poso, A., Ghasemi, J., Amanlou, M., & Masoudi-Nejad, A. (2019). Deep Transferable Compound Representation across Domains and Tasks for Low Data Drug Discovery. Journal of Chemical Information and Modeling, 59(11), 4528–4539. https://doi.org/10.1021/acs.jcim.9b00626