Deep Learning for Protein-Protein Interaction

Subjects: Mathematical & Computational Biology
Contributor: Minhyeok Lee

Deep learning is steadily leaving its transformative imprint across multiple disciplines. Within computational biology, it is expediting progress in the understanding of Protein–Protein Interactions (PPIs), key components governing a wide array of biological functionalities.

Keywords: deep learning ; protein–protein interactions ; computational biology

1. Introduction

In the current era, Artificial Intelligence (AI) forms a transformative underpinning of our scientific progress [1][2][3]. Leveraging advancements in generative deep learning architectures, such as Generative Adversarial Networks (GANs) [4][5][6][7][8], Neural Radiance Fields (NeRF) [9][10][11][12][13][14], and models such as the Generative Pre-training Transformer (GPT) [15][16][17][18], we are facing the proposition that creative intuition, once perceived as an exclusive human trait, may potentially be replicated or even surpassed within an algorithmic framework.

Deep learning has demonstrated exceptional prowess in uncovering complex patterns within high-dimensional data, resulting in ground-breaking applications across various domains [19][20][21]. By exploiting multiple layers of non-linear processing units for feature extraction and transformation, deep learning models can learn hierarchical representations from vast and complex datasets, a characteristic that has found utility in computational biology [22][23][24], and in particular, the prediction of Protein–Protein Interactions (PPIs).

PPIs, pivotal elements in cellular processes, play an instrumental role in various biological functions [25][26][27][28]. These interactions enable proteins to form complex, dynamic networks, which in turn govern biological phenomena spanning from signal transduction to enzymatic activity. Understanding these interactions is crucial, not only for deciphering the complex orchestration of biological systems but also for the identification of novel therapeutic targets for disease intervention. PPIs can be classified into several categories, each with unique characteristics and functional implications. This classification includes direct (physical) and indirect (functional) interactions, permanent and transient interactions, as well as homomeric and heteromeric interactions. Each of these types of PPIs has distinct attributes and implications, necessitating a thorough understanding for successful prediction and analysis.

One groundbreaking application of deep learning in protein studies is embodied by AlphaFold [29], a remarkable AI system developed by DeepMind. AlphaFold stands as a prime example of the confluence of computational prowess and biological understanding, demonstrating the transformative power of AI in deciphering complex biological systems.

AlphaFold utilizes a deep-learning-based approach to predict protein structure, a problem of profound significance in biology. The AI model has been meticulously trained on a wealth of data derived from the Protein Data Bank, integrating a vast multitude of known protein structures into its learning framework. The system leverages this training to predict the arrangement of amino acids within a protein, generating a comprehensive three-dimensional model that illuminates the protein's spatial conformation.

2. Historical Deep Learning Methods for Protein–Protein Interaction Analysis

The emergence and development of historical deep learning methodologies for PPI analysis have significantly facilitated the comprehensive understanding of complex cellular processes. They have been instrumental in enabling thorough investigation and prediction of these interactions. In this section, two representative frameworks (PIPR and DPPI) and their limitations are discussed.

The PIPR framework [30] introduces an innovative approach for PPI prediction centered around amino acid sequences. This method is anchored in a Siamese architecture, leveraging a deep residual recurrent convolutional neural network (RCNN). The integration of recurrent and convolutional layers allows PIPR to accurately capture fundamental local and sequential attributes inherent in protein sequences. To further augment the feature extraction process, PIPR employs an automatic multi-granular feature selection mechanism. This assists PIPR in identifying and giving precedence to the most informative and distinguishing features within the sequences. In addition to this, PIPR amalgamates diverse aspects of PPI data, which includes sequence similarity, evolutionary preservation, and domain-domain interactions, to establish a comprehensive and thorough predictive model. The DPPI model addresses both homodimeric and heterodimeric protein interactions. It can also replicate binding affinities. The creation of the RCNN employed bidirectional gated recurrent units (i.e., bidirectional-GRU), yet GRUs have demonstrated limited learning efficiency and slow convergence [31].

The DPPI method [32] introduces a distinct approach for PPI prediction by harnessing deep learning techniques. The use of deep Siamese-like CNNs, combined with random projection and data augmentation, allows DPPI to deliver accurate sequence-based PPI predictions. This method concentrates on capturing critical aspects of a protein pair's composition, which includes the amino acid sequence and the co-occurrence of overlapping sequence motifs. DPPI employs PSI-BLAST to generate probabilistic sequencing profiles for each protein to extract pertinent features, offering a holistic description. The convolutional module, made up of multiple layers, identifies sequence patterns within each protein's profile. Furthermore, DPPI applies random projection to the representations sourced from the convolutional module, projecting them into two unique spaces. The Siamese-based learning architecture captures the reciprocal influence of protein pairings, allowing for generalization in addressing diverse PPI prediction problems without the necessity for predefined features. However, based on 5-fold cross-validation, DPPI's performance in terms of PPI prediction accuracy on the S.cerevisiae core dataset was found to be inferior to that of PIPR [30].

3. Graph Neural Networks for Protein–Protein Interactions

Graph Neural Networks (GNNs) [33][34][35][36] have emerged as a versatile and powerful class of methods in the computational prediction of PPIs. They represent a specific form of deep learning architecture specially designed for dealing with data structured as graphs. Given the complex nature of biomolecular data, such as proteins, which can be naturally represented as graphs, GNNs provide a unique opportunity to capture intricate patterns and relationships within these datasets.

In essence, a graph can be seen as a collection of nodes and edges, where nodes represent entities (e.g., proteins), and edges denote relationships or interactions (e.g., PPIs). GNNs take advantage of this structured data format by applying various forms of convolutions directly on the graph, enabling them to learn from both local node features and the broader network topology. This ability is particularly useful in the study of PPIs, where the biological significance of an interaction often depends not only on the properties of the interacting proteins but also on their position and role within the larger protein network.

The unique capacity of GNNs to exploit the underlying structure of graph data is achieved through several key mechanisms. Firstly, GNNs use neighborhood aggregation or message-passing frameworks, wherein each node in the graph gathers information from its local neighbors to update its state. This allows GNNs to incorporate local context into node representations, thereby capturing the immediate interaction dynamics in PPIs. Secondly, through multiple rounds of these aggregations, GNNs can learn increasingly abstract representations of nodes, thereby modeling higher-order interaction effects and uncovering complex interaction patterns.

Various types of GNNs have been employed in the study of PPIs, with each offering unique advantages. Graph Convolutional Networks (GCNs) [37][38][39], for instance, are particularly adept at learning from homophily in networks, wherein nodes that are connected or nearby in the graph have similar features. Graph Attention Networks (GATs) [40][41][42] add another level of sophistication by introducing attention mechanisms that allow different weights to be assigned to different neighbors during the aggregation process. These and other variants of GNNs provide a flexible and robust toolset for tackling the challenging task of PPI prediction.

Research leveraging GNNs for PPI prediction spans a wide range of applications, from identifying specific interaction sites on proteins, predicting the existence of interactions between protein pairs, to classifying proteins based on their interaction profiles. These studies typically involve formulating the PPI problem as a graph-based learning task, such as node classification, link prediction, or graph classification, and employing suitable GNN architectures to solve it.

Recent studies have witnessed a prominent trend in utilizing GNNs for PPI predictions. These studies have explored various models and techniques, aiming to enhance the accuracy and efficiency of PPI prediction tasks. Notably, researchers have focused on leveraging GNNs, such as augmented GATs and GCNs, to capture structural invariance, learn graph representations, and improve prediction performance. Additionally, the integration of multimodal data sources, biological features, and prior knowledge has emerged as a significant aspect of recent research efforts. These studies have demonstrated remarkable advancements in predicting PPIs and utilizing PPI information for various predictive tasks, reinforcing the critical role of deep learning methods, particularly GNNs and GCNs, in advancing our understanding of PPIs and their implications in biological systems. Continued research and methodological advancements are expected to drive further progress in this field. The summary of recent studies can be observed in **Table 1**.

Table 1. Summary of Contributions in Studies on Graph Neural Networks for Protein–Protein Interactions. Note that each study employed varied datasets, cross-validation methods, and simulation settings for evaluation, making direct comparisons potentially inconclusive. The highest reported accuracy is presented when models were assessed using multiple datasets.

Author	Metrics and Results	Contributions
Albu et al. [43]	AUC: 0.92 AUPRC: 0.93	Developed MM-StackEns, a deep multimodal stacked generalization approach for predicting PPIs.
Azadifar and Ahmadi [44]	AUC: 0.8847	Introduced a semi-supervised learning method for prioritizing candidate disease genes.
Baranwal et al. [45]	ACC: 0.9889 MCC: 0.9779 AUC: 0.9955	Presented Struct2Graph, a GAT designed for structure-based predictions of PPIs.
Dai et al. [46]	MSE: 0.2446 PCC: 0.8640	Formulated a method for predicting protein abundance from scRNA-seq data.
Gao et al. [47]	ACC: 0.778	Developed the Substructure Assembling Graph Attention Network (SA-GAT) for graph classification tasks.
Hinnerichs and Hoehndorf [48]	AUC: 0.94	Devised DTI-Voodoo, a method combining molecular features and PPI networks to predict drug-target interactions.
Jha et al. [49]	ACC: 0.9813 MCC: 0.9520 AUC: 0.9828 AUPRC: 0.9886	Proposed the use of GCN and GAT to predict PPIs.
Kim et al. [50]	Precision: 0.60 F1: 0.52 NMI: 0.404	Proposed DrugGCN, a GCN for drug response prediction using gene expression data.
Kishan et al. [51]	AUC: 0.936 AUPRC: 0.941	Developed a higher-order GCN for biomedical interaction prediction.
Mahbub and Bayzid [52]	ACC: 0.715 MCC: 0.27 AUC: 0.719 AUPRC: 0.405	Introduced EGRET, an edge aggregated GAT for PPI site prediction.
Quadrini et al. [53]	ACC: 0.731 MCC: 0.054 AUC: 0.588	Explored hierarchical representations of protein structure for PPI site prediction.
Reau et al. [54]	AUC: 0.85	Developed DeepRank-GNN, a graph neural network framework for learning interaction patterns.
Saxena et al. [55]	ACC: 0.9113 F1: 0.90	Proposed a network centrality based approach combined with GCNs for link prediction.
Schapke et al. [56]	AUC: 0.9043 AUPRC: 0.7668	Developed EPGAT, an essentiality prediction model based on GATs.
St-Pierre Lemieux et al. [57]	ACC: 0.84 MCC: 0.94	Presented several geometric deep-learning-based approaches for PPI predictions.
Strokach et al. [58]	Spearman's *R*: 0.62	Described ELASPIC2 (EL2), a machine learning model for predicting mutation effects on protein folding and PPI.

Author	Metrics and Results	Contributions
Wang et al. [59]	ACC 0.9365 MCC 0.4301 AUC 0.6068	Developed SIPGCN, a deep learning model for predicting self-interacting proteins.
Wang et al. [60]	ACC: 0.413	Introduced PLA-GNN, a method for identifying alterations of protein subcellular locations.
Williams et al. [61]	AUC: 0.85	Developed DockNet, a protein–protein interface contact prediction model.
Yuan et al. [62]	ACC: 0.776 MCC: 0.333 AUC: 0.786 AUPRC: 0.429	Proposed GraphPPIS, a deep graph-based framework for PPI site prediction.
Zaki et al. [63]	F1: 0.616	Developed a method for detecting protein complexes in PPI data using GCNs.
Zhou et al. [64]	AUC: 0.5916 AP: 0.85	Conducted a comparative study on various graph neural networks for PPI prediction.
Zhou et al. [65]	ACC: 0.856 F1: 0.569 AUC: 0.867 AUPRC: 0.574	Presented AGAT-PPIS, an augmented graph attention network for PPI site prediction.

3.1. Pairwise PPI Prediction

Albu et al. [43] presented MM-StackEns, a deep multimodal stacked generalization approach for predicting PPIs, employing a Siamese neural network and graph attention networks, with superior performance on Yeast and Human datasets. Similarly, Jha et al. [49] used Graph Convolutional Network (GCN) and Graph Attention Network (GAT) for PPI prediction, yielding superior results on Human and S. cerevisiae datasets.

3.2. PPI Network Prediction

Baranwal et al. [45] offered Struct2Graph, a graph attention network for structure-based PPI predictions, potentially identifying residues contributing to protein–protein complex formation. Gao et al. [47] designed the Substructure Assembling Graph Attention Network (SA-GAT) for graph classification tasks, including potential applications in PPI networks. Zaki et al. [63] proposed a method for detecting protein complexes in PPI data using GCNs, formulating protein complex detection as a node classification problem and implementing the Neural Overlapping Community Detection (NOCD) model.

3.3. PPI Site Prediction

Quadrini et al. [53] used Graph Convolutional Networks for PPI site prediction, exploring a novel abstraction of protein structure termed as hierarchical representations. Mahbub and Bayzid [52] introduced EGRET, an edge aggregated graph attention network for PPI site prediction, reporting significant improvements in performance. Yuan et al. [62] proposed GraphPPIS, a deep graph-based framework for PPI site prediction that delivered significantly improved performance over other methods.

3.4. Docking

Williams et al. [61] developed DockNet, a high-throughput protein–protein interface contact prediction model utilizing a Siamese graph-based neural network. Reau et al. [54] developed DeepRank-GNN, a graph neural network framework that converts protein–protein interfaces into graphs to learn interaction patterns.

3.5. Auxiliary PPI Prediction Tasks

Azadifar and Ahmadi [44] introduced a semi-supervised learning method based on GCNs for prioritizing candidate disease genes. Dai et al. [46] formulated PIKE-R2P, a graph neural network method incorporating PPIs for predicting protein abundance from scRNA-seq data. Hinnerichs and Hoehndorf [48] developed DTI-Voodoo, a method combining molecular features and PPI networks to predict drug-target interactions. Kim et al. [50] proposed DrugGCN for drug response prediction using gene expression data. Wang et al. [59] developed SIPGCN, a GCN-based model for predicting self-interacting proteins (SIPs) from sequence information.

The range and depth of these studies underscore the crucial role deep learning methods, particularly GNNs and GCNs, continue to play in advancing PPI predictions. With ongoing research and methodological enhancements, the future promises continued progress in understanding and predicting PPIs and their influence on biological systems.

4. Multi-task or Multi-modal Deep Learning Models for Protein–Protein Interactions

The utilization of multi-task and multi-modal deep learning models [66][67] has been increasingly recognized as an efficient approach to deal with the complexity and heterogeneity of PPI prediction problems. These models are designed to leverage multiple related tasks or multiple sources of information to improve predictive performance, offering a promising direction for the exploration and prediction of PPIs.

Multi-task learning models are designed to improve learning efficiency and predictive performance by learning multiple related tasks concurrently [66]. The fundamental concept behind multi-task learning is the sharing of representations among tasks, which can improve the generalization performance by leveraging the commonalities and differences across tasks. In a standard multi-task learning framework, each task has its own specific layers (task-specific layers), while some layers (shared layers) are shared among all tasks. During training, each task's loss function is typically optimized, and the overall objective is a weighted sum of these individual loss functions. The shared layers learn a representation that captures the common features among tasks, while the task-specific layers learn the unique features for each task.

Multi-modal deep learning models [67], on the other hand, aim to integrate information from multiple sources or modes. The basic principle of multi-modal learning is to construct a joint representation that leverages the complementarity and correlation among different modalities to improve prediction performance. In a standard multi-modal learning framework, the model first learns a representation for each modality using modality-specific layers and then integrates these representations using shared layers. The modalities can be different types of data (e.g., sequence data, structure data), each of which provides a unique perspective on the problem.

In the context of PPI prediction, these methodologies offer several advantages. Multi-task learning models can learn from multiple related tasks (e.g., predicting different types of protein interactions), thereby leveraging the shared information among tasks to improve prediction performance. Similarly, multi-modal models can integrate information from multiple sources (e.g., sequence data, structural data, functional data), thereby leveraging the complementarity among different types of data to obtain a more comprehensive understanding of the protein interaction mechanisms.

Given their potential for dealing with complex and heterogeneous PPI prediction problems, multi-task and multi-modal deep learning models have found broad applications in the PPI field. They have been used to leverage multiple related tasks or multiple sources of information, improving prediction performance and providing a more comprehensive understanding of the protein interaction mechanisms.

Recent studies have focused on the development of multi-task or multi-modal deep learning models to enhance the prediction of PPIs. These models aim to leverage multiple sources of information, such as protein sequences, structural annotations, gene features, multiomics data, and GO information, to improve the accuracy and robustness of PPI predictions. By incorporating various tasks or modalities into the learning process, these models have demonstrated superior performance compared to single-task methods. Additionally, efforts have been made to enhance the interpretability of deep learning models by incorporating explainable features or methodologies. These advancements in multi-task and multi-modal deep learning approaches have opened up new possibilities for predicting PPIs and expanding our understanding of complex biological interactions in diverse areas, including disease research and infectious disease studies. **Table 2** outlines the main points from recent research.

Table 2. Summary of Contributions in Studies on Multi-task or Multi-modal Models for Protein-Protein Interactions. Note that each study employed varied datasets, cross-validation methods, and simulation settings for evaluation, making direct comparisons potentially inconclusive. The highest reported accuracy is presented when models were assessed using multiple datasets.

Author	Metrics and Results	Contributions
Capel et al. [68]	**AUC: 0.7632 AUPRC: 0.3844**	**Proposed a multi-task deep learning approach for predicting residues in PPI interfaces.**

Author	Metrics and Results	Contributions
Li et al. [69]	AUC: 0.895 AUPRC: 0.899	Developed EP-EDL, an ensemble deep learning model for accurate prediction of human essential proteins.
Linder et al. [70]	AUC: 0.96	Introduced scrambler networks to improve the interpretability of neural networks for biological sequences.
Pan et al. [71]	ACC: 0.8947 MCC: 0.7902 AUC: 0.9548	Proposed DWPPI, a network embedding-based approach for PPI prediction in plants.
Peng et al. [72]	AUC: 0.9116 AUPRC: 0.8332	Introduced MTGCN, a multi-task learning method for identifying cancer driver genes.
Schulte-Sasse et al. [73]	AUPRC: 0.76	Developed EMOGI, integrating *MULTIOMICS* data with PPI networks for cancer gene prediction.
Thi Ngan Dong et al. [74]	AUC: 0.9804 F1: 0.9379	Developed a multitask transfer learning approach for predicting virus-human and bacteria-human PPIs.
Zheng et al. [75]	AUPRC: 0.965	Developed DeepAraPPI, a deep learning framework for predicting PPIs in Arabidopsis thaliana.

4.1. Pairwise PPI Prediction

A range of models have been proposed to predict pairwise PPIs. For instance, Capel et al. [68] proposed a multi-task learning strategy to predict residues in PPI interfaces from protein sequences. Similarly, Li et al. [69] developed EP-EDL, an ensemble deep learning model, to predict human essential proteins using protein sequence information. Thi Ngan Dong et al. [74] employed a multitask transfer learning approach for predicting PPIs between viruses and human cells, showing the effectiveness of this method across multiple PPI prediction tasks.

4.2. PPI Network Prediction

Several models have been developed to predict PPI networks. Peng et al. [72] introduced MTGCN, a multi-task learning method based on the Graph Convolutional Network, to identify cancer driver genes using gene features from the PPI network. Schulte-Sasse et al. [73] developed EMOGI, which utilizes graph convolutional networks to integrate multiomics pan-cancer data with PPI networks for cancer gene prediction. Finally, Pan et al. [71] proposed DWPPI, a network embedding-based approach that integrates deep neural networks for PPI prediction in plants, demonstrating superior performance across multiple datasets.

4.3. PPI Site Prediction

In the PPI site prediction, Capel et al. [68] have demonstrated a promising approach, utilizing a multi-task learning strategy to predict residues in PPI interfaces from protein sequences, outperforming single-task methods significantly.

4.4. Auxiliary PPI Prediction Tasks

A variety of models have been proposed for auxiliary PPI prediction tasks. Linder et al. [70] introduced scrambler networks, a feature attribution method designed for discrete sequence inputs, to improve the interpretability of neural networks for biological sequences. These networks have been used for interpreting effects of genetic variants, cis-regulatory elements interactions, and PPI binding specificity. Lastly, Zheng et al. [75] developed DeepAraPPI, an integrative deep learning framework for predicting PPIs in Arabidopsis thaliana, demonstrating excellent performance and promising cross-species predictive ability.

References

1. Rajpurkar, P.; Chen, E.; Banerjee, O.; Topol, E.J. AI in health and medicine. Nat. Med. 2022, 28, 31–38.
2. Cetinic, E.; She, J. Understanding and creating art with AI: Review and outlook. ACM Trans. Multimed. Comput. Commun. Appl. (TOMM) 2022, 18, 1–22.
3. Chamola, V.; Hassija, V.; Gupta, V.; Guizani, M. A comprehensive review of the COVID-19 pandemic and the role of IoT, drones, AI, blockchain, and 5G in managing its impact. IEEE Access 2020, 8, 90225–90265.

4. Aggarwal, A.; Mittal, M.; Battineni, G. Generative adversarial network: An overview of theory and applications. Int. J. Inf. Manag. Data Insights 2021, 1, 100004.
5. Jabbar, A.; Li, X.; Omar, B. A survey on generative adversarial networks: Variants, applications, and training. ACM Comput. Surv. (CSUR) 2021, 54, 1–49.
6. Cai, Z.; Xiong, Z.; Xu, H.; Wang, P.; Li, W.; Pan, Y. Generative adversarial networks: A survey toward private and secure applications. ACM Comput. Surv. (CSUR) 2021, 54, 1–38.
7. Chen, Y.; Yang, X.H.; Wei, Z.; Heidari, A.A.; Zheng, N.; Li, Z.; Chen, H.; Hu, H.; Zhou, Q.; Guan, Q. Generative adversarial networks in medical image augmentation: A review. Comput. Biol. Med. 2022, 144, 105382.
8. Gui, J.; Sun, Z.; Wen, Y.; Tao, D.; Ye, J. A review on generative adversarial networks: Algorithms, theory, and applications. IEEE Trans. Knowl. Data Eng. 2021, 35, 3313–3332.
9. Barron, J.T.; Mildenhall, B.; Verbin, D.; Srinivasan, P.P.; Hedman, P. Mip-nerf 360: Unbounded anti-aliased neural radiance fields. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, New Orleans, LA, USA, 18–24 June 2022; pp. 5470–5479.
10. Martin-Brualla, R.; Radwan, N.; Sajjadi, M.S.; Barron, J.T.; Dosovitskiy, A.; Duckworth, D. Nerf in the wild: Neural radiance fields for unconstrained photo collections. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 7210–7219.
11. Mildenhall, B.; Srinivasan, P.P.; Tancik, M.; Barron, J.T.; Ramamoorthi, R.; Ng, R. Nerf: Representing scenes as neural radiance fields for view synthesis. Commun. ACM 2021, 65, 99–106.
12. Zhang, K.; Riegler, G.; Snavely, N.; Koltun, V. Nerf++: Analyzing and improving neural radiance fields. arXiv 2020, arXiv:2010.07492.
13. Yu, A.; Li, R.; Tancik, M.; Li, H.; Ng, R.; Kanazawa, A. Plenoctrees for real-time rendering of neural radiance fields. In Proceedings of the IEEE/CVF International Conference on Computer Vision, Montreal, QC, Canada, 10–17 October 2021; pp. 5752–5761.
14. Pumarola, A.; Corona, E.; Pons-Moll, G.; Moreno-Noguer, F. D-nerf: Neural radiance fields for dynamic scenes. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, Nashville, TN, USA, 20–25 June 2021; pp. 10318–10327.
15. OpenAI. GPT-4 Technical Report. OpenAI Technical Report. 2023. Available online: https://cdn.openai.com/papers/gpt-4.pdf (accessed on 15 May 2023).
16. Radford, A.; Wu, J.; Child, R.; Luan, D.; Amodei, D.; Sutskever, I. Language Models Are Unsupervised Multitask Learners. OpenAI Technical Report. 2019. Available online: https://d4mucfpksywv.cloudfront.net/better-language-models/language_models_are_unsupervised_multitask_learners.pdf (accessed on 15 May 2023).
17. Radford, A.; Narasimhan, K.; Salimans, T.; Sutskever, I. Improving Language Understanding by Generative Pre-Training. OpenAI Technical Report. 2018. Available online: https://s3-us-west-2.amazonaws.com/openai-assets/research-covers/language-unsupervised/language_understanding_paper.pdf (accessed on 15 May 2023).
18. Brown, T.; Mann, B.; Ryder, N.; Subbiah, M.; Kaplan, J.D.; Dhariwal, P.; Neelakantan, A.; Shyam, P.; Sastry, G.; Askell, A.; et al. Language models are few-shot learners. Adv. Neural Inf. Process. Syst. 2020, 33, 1877–1901.
19. Zhang, Y.; Shi, X.; Zhang, H.; Cao, Y.; Terzija, V. Review on deep learning applications in frequency analysis and control of modern power system. Int. J. Electr. Power Energy Syst. 2022, 136, 107744.
20. Yazici, I.; Beyca, O.F.; Delen, D. Deep-learning-based short-term electricity load forecasting: A real case application. Eng. Appl. Artif. Intell. 2022, 109, 104645.
21. Choudhary, K.; DeCost, B.; Chen, C.; Jain, A.; Tavazza, F.; Cohn, R.; Park, C.W.; Choudhary, A.; Agrawal, A.; Billinge, S.J.; et al. Recent advances and applications of deep learning methods in materials science. npj Comput. Mater. 2022, 8, 59.
22. Tang, B.; Pan, Z.; Yin, K.; Khateeb, A. Recent advances of deep learning in bioinformatics and computational biology. Front. Genet. 2019, 10, 214.
23. Yazdani, A.; Lu, L.; Raissi, M.; Karniadakis, G.E. Systems biology informed deep learning for inferring parameters and hidden dynamics. PLoS Comput. Biol. 2020, 16, e1007575.
24. Zampieri, G.; Vijayakumar, S.; Yaneske, E.; Angione, C. Machine and deep learning meet genome-scale metabolic modeling. PLoS Comput. Biol. 2019, 15, e1007084.
25. Lu, H.; Zhou, Q.; He, J.; Jiang, Z.; Peng, C.; Tong, R.; Shi, J. Recent advances in the development of protein–protein interactions modulators: Mechanisms and clinical trials. Signal Transduct. Target. Ther. 2020, 5, 213.

26. Bryant, P.; Pozzati, G.; Elofsson, A. Improved prediction of protein–protein interactions using AlphaFold2. Nat. Commun. 2022, 13, 1265.

27. Hu, L.; Wang, X.; Huang, Y.A.; Hu, P.; You, Z.H. A survey on computational models for predicting protein–protein interactions. Brief. Bioinform. 2021, 22, bbab036.

28. Richards, A.L.; Eckhardt, M.; Krogan, N.J. Mass spectrometry-based protein–protein interaction networks for the study of human diseases. Mol. Syst. Biol. 2021, 17, e8792.

29. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Tunyasuvunakool, K.; Bates, R.; Žídek, A.; Potapenko, A.; et al. Highly accurate protein structure prediction with AlphaFold. Nature 2021, 596, 583–589.

30. Chen, M.; Ju, C.J.T.; Zhou, G.; Chen, X.; Zhang, T.; Chang, K.W.; Zaniolo, C.; Wang, W. Multifaceted protein–protein interaction prediction based on Siamese residual RCNN. Bioinformatics 2019, 35, i305–i314.

31. Wang, X.; Xu, J.; Shi, W.; Liu, J. OGRU: An Optimized Gated Recurrent Unit Neural Network. J. Phys. Conf. Ser. 2019, 1325, 012089.

32. Hashemifar, S.; Neyshabur, B.; Khan, A.A.; Xu, J. Predicting protein–protein interactions through sequence-based deep learning. Bioinformatics 2018, 34, i802–i810.

33. Wu, Z.; Pan, S.; Chen, F.; Long, G.; Zhang, C.; Philip, S.Y. A comprehensive survey on graph neural networks. IEEE Trans. Neural Netw. Learn. Syst. 2020, 32, 4–24.

34. Fan, W.; Ma, Y.; Li, Q.; He, Y.; Zhao, E.; Tang, J.; Yin, D. Graph neural networks for social recommendation. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 13–17 May 2019; pp. 417–426.

35. Liu, M.; Gao, H.; Ji, S. Towards deeper graph neural networks. In Proceedings of the 26th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Virtual, 6–10 July 2020; pp. 338–348.

36. Zhou, J.; Cui, G.; Hu, S.; Zhang, Z.; Yang, C.; Liu, Z.; Wang, L.; Li, C.; Sun, M. Graph neural networks: A review of methods and applications. AI Open 2020, 1, 57–81.

37. Zhang, S.; Tong, H.; Xu, J.; Maciejewski, R. Graph convolutional networks: A comprehensive review. Comput. Soc. Netw. 2019, 6, 1–23.

38. Wu, F.; Souza, A.; Zhang, T.; Fifty, C.; Yu, T.; Weinberger, K. Simplifying graph convolutional networks. In Proceedings of the International Conference on Machine Learning, PMLR, Long Beach, CA, USA, 9–15 June 2019; pp. 6861–6871.

39. Chen, M.; Wei, Z.; Huang, Z.; Ding, B.; Li, Y. Simple and deep graph convolutional networks. In Proceedings of the International Conference on Machine Learning, PMLR, Virtual, 13–18 July 2020; pp. 1725–1735.

40. Wang, X.; He, X.; Cao, Y.; Liu, M.; Chua, T.S. Kgat: Knowledge graph attention network for recommendation. In Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining, Anchorage, AK, USA, 4–8 August 2019; pp. 950–958.

41. Song, W.; Xiao, Z.; Wang, Y.; Charlin, L.; Zhang, M.; Tang, J. Session-based social recommendation via dynamic graph attention networks. In Proceedings of the Twelfth ACM International Conference on Web Search and Data Mining, Melbourne, VIC, Australia, 11–15 February 2019; pp. 555–563.

42. Wang, X.; Ji, H.; Shi, C.; Wang, B.; Ye, Y.; Cui, P.; Yu, P.S. Heterogeneous graph attention network. In Proceedings of the World Wide Web Conference, San Francisco, CA, USA, 15–17 May 2019; pp. 2022–2032.

43. Albu, A.I.; Bocicor, M.I.; Czibula, G. MM-StackEns: A new deep multimodal stacked generalization approach for protein–protein interaction prediction. Comput. Biol. Med. 2023, 153, 106526.

44. Azadifar, S.; Ahmadi, A. A novel candidate disease gene prioritization method using deep graph convolutional networks and semi-supervised learning. BMC Bioinform. 2022, 23, 422.

45. Baranwal, M.; Magner, A.; Saldinger, J.; Turali-Emre, E.S.; Elvati, P.; Kozarekar, S.; VanEpps, J.S.; Kotov, N.A.; Violi, A.; Hero, A.O. Struct2Graph: A graph attention network for structure based predictions of protein–protein interactions. BMC Bioinform. 2022, 23, 370.

46. Dai, X.; Xu, F.; Wang, S.; Mundra, P.A.; Zheng, J. PIKE-R2P: Protein-protein interaction network-based knowledge embedding with graph neural network for single-cell RNA to protein prediction. BMC Bioinform. 2021, 22 (Suppl. 6), 139.

47. Gao, J.; Gao, J.; Ying, X.; Lu, M.; Wang, J. Higher-Order Interaction Goes Neural: A Substructure Assembling Graph Attention Network for Graph Classification. IEEE Trans. Knowl. Data Eng. 2023, 35, 1594–1608.

48. Hinnerichs, T.; Hoehndorf, R. DTI-Voodoo: Machine learning over interaction networks and ontology-based background knowledge predicts drug-target interactions. Bioinformatics 2021, 37, 4835–4843.

49. Jha, K.; Saha, S.; Singh, H. Prediction of protein–protein interaction using graph neural networks. Sci. Rep. 2022, 12, 8360.
50. Kim, S.; Bae, S.; Piao, Y.; Jo, K. Graph Convolutional Network for Drug Response Prediction Using Gene Expression Data. Mathematics 2021, 9, 772.
51. Kishan, K.C.; Li, R.; Cui, F.; Haake, A.R. Predicting Biomedical Interactions with Higher-Order Graph Convolutional Networks. IEEE/ACM Trans. Comput. Biol. Bioinform. 2022, 19, 676–687.
52. Mahbub, S.; Bayzid, M.S. EGRET: Edge aggregated graph attention networks and transfer learning improve protein–protein interaction site prediction. Brief. Bioinform. 2022, 23, bbab578.
53. Quadrini, M.; Daberdaku, S.; Ferrari, C. Hierarchical representation for PPI sites prediction. BMC Bioinform. 2022, 23, 96.
54. Reau, M.; Renaud, N.; Xue, L.C.; Bonvin, A.M.J.J. DeepRank-GNN: A graph neural network framework to learn patterns in protein–protein interfaces. Bioinformatics 2023, 39, btac759.
55. Saxena, R.; Patil, S.P.; Verma, A.K.; Jadeja, M.; Vyas, P.; Bhateja, V.; Lin, J.C.W. An Efficient Bet-GCN Approach for Link Prediction. Int. J. Interact. Multimed. Artif. Intell. 2023, 8, 38–52.
56. Schapke, J.; Tavares, A.; Recamonde-Mendoza, M. EPGAT: Gene Essentiality Prediction with Graph Attention Networks. IEEE/ACM Trans. Comput. Biol. Bioinform. 2022, 19, 1615–1626.
57. Lemieux, G.S.P.; Paquet, E.; Viktor, H.L.; Michalowski, W. Geometric Deep Learning for Protein-Protein Interaction Predictions. IEEE Access 2022, 10, 90045–90055.
58. Strokach, A.; Lu, T.Y.; Kim, P.M. ELASPIC2 (EL2): Combining Contextualized Language Models and Graph Neural Networks to Predict Effects of Mutations. J. Mol. Biol. 2021, 433.
59. Wang, Y.; Wang, L.L.; Wong, L.; Li, Y.; Wang, L.; You, Z.H. SIPGCN: A Novel Deep Learning Model for Predicting Self-Interacting Proteins from Sequence Information Using Graph Convolutional Networks. Biomedicines 2022, 10, 1543.
60. Wang, R.H.; Luo, T.; Zhang, H.L.; Du, P.F. PLA-GNN: Computational inference of protein subcellular location alterations under drug treatments with deep graph neural networks. Comput. Biol. Med. 2023, 157, 106775.
61. Williams, N.P.; Rodrigues, C.H.M.; Truong, J.; Ascher, D.B.; Holien, J.K. DockNet: High-throughput protein–protein interface contact prediction. Bioinformatics 2023, 39, btac797.
62. Yuan, Q.; Chen, J.; Zhao, H.; Zhou, Y.; Yang, Y. Structure-aware protein–protein interaction site prediction using deep graph convolutional network. Bioinformatics 2022, 38, 125–132.
63. Zaki, N.; Singh, H.; Mohamed, E.A. Identifying Protein Complexes in Protein-Protein Interaction Data Using Graph Convolutional Network. IEEE Access 2021, 9, 123717–123726.
64. Zhou, H.; Wang, W.; Jin, J.; Zheng, Z.; Zhou, B. Graph Neural Network for Protein-Protein Interaction Prediction: A Comparative Study. Molecules 2022, 27, 6135.
65. Zhou, Y.; Jiang, Y.; Yang, Y. AGAT-PPIS: A novel protein–protein interaction site predictor based on augmented graph attention network with initial residual and identity mapping. Brief. Bioinform. 2023, 24, bbad122.
66. Vithayathil Varghese, N.; Mahmoud, Q.H. A survey of multi-task deep reinforcement learning. Electronics 2020, 9, 1363.
67. Feng, D.; Haase-Schütz, C.; Rosenbaum, L.; Hertlein, H.; Glaeser, C.; Timm, F.; Wiesbeck, W.; Dietmayer, K. Deep multi-modal object detection and semantic segmentation for autonomous driving: Datasets, methods, and challenges. IEEE Trans. Intell. Transp. Syst. 2020, 22, 1341–1360.
68. Capel, H.; Feenstra, K.A.; Abeln, S. Multi-task learning to leverage partially annotated data for PPI interface prediction. Sci. Rep. 2022, 12, 10487.
69. Li, Y.; Zeng, M.; Wu, Y.; Li, Y.; Li, M. Accurate Prediction of Human Essential Proteins Using Ensemble Deep Learning. IEEE/ACM Trans. Comput. Biol. Bioinform. 2022, 19, 3263–3271.
70. Linder, J.; La Fleur, A.; Chen, Z.; Ljubetic, A.; Baker, D.; Kannan, S.; Seelig, G. Interpreting neural networks for biological sequences by learning stochastic masks. Nat. Mach. Intell. 2022, 4, 41.
71. Pan, J.; You, Z.H.; Li, L.P.; Huang, W.Z.; Guo, J.X.; Yu, C.Q.; Wang, L.P.; Zhao, Z.Y. DWPPI: A Deep Learning Approach for Predicting Protein-Protein Interactions in Plants Based on Multi-Source Information with a Large-Scale Biological Network. Front. Bioeng. Biotechnol. 2022, 10, 807522.
72. Peng, W.; Tang, Q.; Dai, W.; Chen, T. Improving cancer driver gene identification using multi-task learning on graph convolutional network. Brief. Bioinform. 2022, 23, bbab432.

73. Schulte-Sasse, R.; Budach, S.; Hnisz, D.; Marsico, A. Integration of multiomics data with graph convolutional networks to identify new cancer genes and their associated molecular mechanisms. Nat. Mach. Intell. 2021, 3, 513.
74. Dong, T.N.; Brogden, G.; Gerold, G.; Khosla, M. A multitask transfer learning framework for the prediction of virus-human protein–protein interactions. BMC Bioinform. 2021, 22, 572.
75. Zheng, J.; Yang, X.; Huang, Y.; Yang, S.; Wuchty, S.; Zhang, Z. Deep learning-assisted prediction of protein–protein interactions in Arabidopsis thaliana. Plant J. 2023, 114, 984–994.