Enhancing Protein Function Prediction Through the Fusion of Multi-Type Biological Knowledge With Protein Language Model and Graph Neural Network

Wenkang Wang, Yunyan Shuai, Yiming Li, Min Zeng, and Min Li

***Abstract*—Proteins play crucial roles in diverse biological functions. Accurately annotating their functions is essential for understanding cellular mechanisms and developing therapies for complex diseases. Computational methods have been proposed as alternatives to labor-intensive and expensive experimental approaches. Existing computational methods have demonstrated that protein evolution information and Protein-Protein Interactions (PPIs) are essential for protein function prediction. However, traditional computational approaches for generating evolution information are time-consuming. On the other hand, proteins lacking interactions are ignored in previous studies. To address these limitations, we propose a novel deep learning framework, named DeepFMB, which incorporates multi-type biological knowledge. DeepFMB leverages a pre-trained protein language model to extract evolution information. Moreover, DeepFMB generates PPI-related features and orthology-related features using graph neural networks on the constructed PPI and orthology networks. Then, these multi-type features are fused adaptively for protein function prediction. Compared to eight state-of-the-art methods, DeepFMB outperforms all of them in terms of F-max and AUPR. Additionally, with the combination of sequence similarity-based inference, our predicted model predicts protein functions more accurately. Experimental results also validate the superior performance of our methods in predicting low-frequency GO terms. Ablation studies demonstrate that the multi-type biological knowledge we use is highly relevant to protein functions.**

***Index Terms*—Protein function prediction, protein-protein interactions, orthology relations, pre-trained protein language model, graph neural network.**

I. Introduction

PROTEINS are essential biomolecules within living organisms and play indispensable roles in diverse biological processes, including cell motility and metabolic pathways [1]. Accurately annotating protein functions is crucial for understanding biological activities and disease pathology mechanisms [2]. However, traditional experimental approaches for functional annotation are costly, time-consuming, and unable to keep up with the rapidly increasing number of protein sequence generated by high-throughput sequencing technologies [3]. Consequently, over 99% of protein sequences in the UniProt [4] database lack experimentally validated functional annotations. Faced with this challenge, it is essential to develop computational approaches to accurately predict protein functions [5]. Such approaches not only enhance our understanding of protein functions but also expedite new drug discoveries and facilitate the development of disease treatment strategies.

During the early stages of protein function prediction research, the lack of a unified standard led to significant differences in how biologists defined protein functions, posing challenges to this field. To address this issue, researchers established specialized databases such as Gene Ontology (GO) [6], which provides a standardized definition of biological information by integrating descriptions of gene products across databases. GO classifies protein functions into three aspects: Cellular Component (CC), Molecular Function (MF), and Biological Process (BP). Each function is considered as a 'GO term' with a unique identification, which together form a Directed Acyclic Graph (DAG) through the relationships between different GO terms, including 'is-a', 'part-of', etc [7].

Till now, significant progress has been made in this field. Protein sequences, as the basics of proteins, are closely correlated with protein functions and have been widely used in existing computational methods [3], [8], [9], [10], [11]. Sequence-based methods, such as Diamond [12] and BlastKNN [13], predict functions by transferring functions from known proteins via the sequence similarities. With the development of deep learning (DL), several DL-based approaches, such as DeepGOPlus [10] and TALE [14], use Convolutional Neural Network (CNN) [15] and Transformer architectures [16], respectively, to extract sequence motifs for function prediction. Benefiting from the advantages of deep learning in extracting latent features, these methods surpass traditional approaches based on the sequence similarities [7].

Additionally, since proteins do not perform functions independently, the proteins that interact with each other tend to be more likely to perform similar functions [17]. Protein-Protein

Received 1 May 2024; revised 26 November 2024; accepted 31 December 2024. Date of publication 14 January 2025; date of current version 3 April 2025. This work was supported in part by the National Science Fund for Distinguished Young Scholars under Grant 62225209, in part by the Science and Technology Innovation Program of Hunan Province under Grant 2021RC4008, in part by Hunan Postgraduate Research and Innovation Project under Grant CX20240018, and in part by the High Performance Computing Center of Central South University. *(Corresponding author: Min Li.)*

The authors are with the School of Computer Science and Engineering, Central South University, Changsha 410083, China (e-mail: wangwk@csu.edu.cn; shuaiyy@csu.edu.cn; lym1998@csu.edu.cn; zengmin@csu.edu.cn; limin@mail.csu.edu.cn).

The source code can be downloaded from https://github.com/CSUBioGroup/DeepFMB.

This article has supplementary downloadable material available at https://doi.org/10.1109/TCBBIO.2025.3529301, provided by the authors.

Digital Object Identifier 10.1109/TCBBIO.2025.3529301

2998-4165 © 2025 IEEE. All rights reserved, including rights for text and data mining, and training of artificial intelligence and similar technologies. Personal use is permitted, but republication/redistribution requires IEEE permission. See https://www.ieee.org/publications/rights/index.html for more information.

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

Interactions (PPIs) are also crucial in detecting protein functions [18]. Consequently, several methods have attempted to annotate protein functions by incorporating protein sequence and PPI network information. For instance, DeepGraphGO [19] uses Graph Neural Networks (GNNs) [20] to aggregate sequence features of interacted proteins through PPI network. DeepGOA [21] employs Bidirectional Long Short-Term Memory (Bi-LSTM) [22] and Convolutional Neural Network (CNN) to extract sequence profiles, adopts fully connected layers to extract InterPro features [23], [24], [25] and protein embeddings from PPI network which are generated by Node2Vec [26], demonstrating the effective of the strategy that considering protein sequence and PPI for protein function prediction.

Although these DL-based approaches have significantly improved the accuracy of protein function prediction than traditional sequence-based computational methods, they still face some challenges. Firstly, sequence profiles that contain co-evolution information play a vital role in protein function prediction [21], which are generated by BLAST [13] or HHsuite [27]. These tools need to scan the whole background database for the target proteins, which is very time-consuming. Additionally, PPI is indeed critical for protein function prediction. Existing PPI-based approaches struggle to annotate functions for proteins that lack interactions [19]. This limitation leads to poor performance in predicting the functions of such proteins.

Recently, several protein Large Language Models (pLLMs) [28], [29] have been proposed and achieved significant improvement in many fields [30], [31]. These models are trained on large scale protein sequences, enabling them to capture co-evolution information and generate high-quality features from protein sequences. Meanwhile, orthology relations also provide the information between proteins that share similar functions [17]. As orthology relations are widely predicted or detected from protein sequences, they cover a broader range of proteins compared to PPIs [32], [33].

To address these limitations mentioned above, we propose a novel DL-based method, named DeepFMB (A deep learning-based model for protein function prediction by fusing multi-type biological knowledge). DeepFMB adopts various forms of biological knowledge, including protein sequences, PPIs and orthology relations, to accurately predict protein functions. DeepFMB integrates these various sources of biological knowledge effectively. Specifically, DeepFMB first generates co-evolution information as the evolution-related features from protein sequences using a pre-trained pLLM [28]. Then, DeepFMB constructs functional features via InterProScan [34] and aggregates features from similar proteins through PPIs and orthology relations as the PPI-related and orthology-related features, respectively. Finally, three types of features are fused to predict protein functions. Furthermore, we integrate the results of DeepFMB and a sequence similarity-based method, BlastKNN, denoted as DeepFMB+. To evaluate the performance of DeepFMB and DeepFMB+, we compare them with 11 state-of-the-art methods on the latest datasets, including 8 single methods and 3 composite methods. The results show that DeepFMB outperforms other single methods in terms of Fmax and AUPR. After incorporating the sequence similarity, DeepFMB+ achieves the best performance over all other methods in terms of Fmax and AUPR for BP, MF and CC. Further ablation study and several cases demonstrate the positive effect of fusing multi-type biological knowledge in DeepFMB for protein function prediction.

II. Methods

The framework of DeepFMB is illustrated in Fig. 1, which mainly consists of four modules: (1) Construction of evolution-related features. (2) Construction of PPI-related features and orthology-related features. (3) Fusion of multi-type biological features and protein function prediction. (4) Extension with sequence similarity-based methods.

A. Construction of Evolution-Related Features

Evolution-related features play a pivotal role in DeepFMB, following the recognition of the significance of evolution information in previous studies [9], [21]. However, different from previous work, in this study, DeepFMB utilizes a distinct approach by utilizing a pre-trained protein large language model [28] to generate evolution-related features. The pLLM has been trained on large-scale protein sequences with a self-supervised strategy, enabling it to learn the evolutionary process of proteins. An obvious advantage is that the pLLM can be easily merged into DeepFMB, as it only requires protein sequences as input. This contrasts with traditional methods that require pre-prepared protein profiles, resulting in time-consuming preparation steps.

Specifically, for each protein P, a famous pLLM, ESM-1b [28], is used to generate initial evolution-related features. We extract the output of 33rd layer of ESM-1b as protein embedding, denoted as $f^{pre} \in R^{1*1280}$. Then, the initial features are fed into three non-linear layers to get more sophisticated and representative features:

$$f^{evo} = \sigma\left(\sigma\left(\sigma\left(f^{pre} * W_1 + b_1\right) * W_2 + b_2\right) * W_3 + b_3\right) \quad (1)$$

where (W_i, b_i) represent the corresponding linear layers $(i = 1, 2, 3)$ and $\sigma(\cdot)$ is activation function LeakReLU. Finally, we can obtain the evolution-related features $f^{evo} \in R^{1*512}$.

B. Construction of PPI-related Features and Orthology-Related Features

1) Generating Functional Features: InterPro [23], [24], [25] is a comprehensive database that integrates 14 databases including SMART [35], CDD [36], Pfam [37], SUPERFAMILY [38]. This platform offers researchers a wealth of information on protein functions, structures, and sequence annotations. Through InterPro, users can access functional characteristics on specific proteins, including their domains, motifs, and families. These features provided by InterPro are crucial for gaining a deeper understanding of protein functions and have been widely used in previous studies [2], [19].

Specifically, we utilize the InterProScan tool [34] to extract interpro properties (domains/motifs/families) of proteins. Then, for each protein P, these properties are encoded as 39227-dimensional binary features, denoted as $f^{inter} \in \{0, 1\}^{1*39227}$,

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

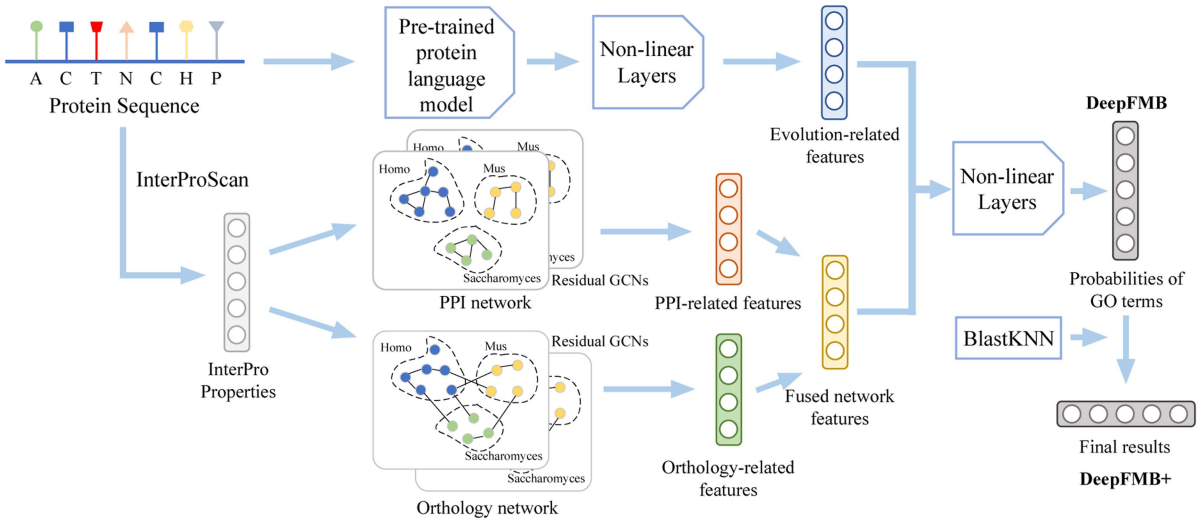

Fig. 1. The framework of DeepFMB and DeepFMB+. Firstly, evolution-related features are extracted using the pre-trained protein large language model ESM-1b. Then, InterPro properties are generated from the protein sequences by InterProScan. Subsequently, two residual GNNs are used to aggregate the neighboring features through the PPI network and the orthology network, separately. The PPI-related and orthology-related features, extracted from the two types of networks, are fused by a linear attention mechanism. Then, the fused features are combined with evolution-related feature to predict protein functions. DeepFMB+ incorporates BlastKNN results into DeepFMB, further improving the performance of protein function prediction.

where each position represents a specific property, with a value of 1 indicating that P has the corresponding property. Subsequently, the binary features are fed into an embedding layer to obtain a dense representation with low dimensions:

$$f^{func} = \tau\left(f^{inter} * W_{emd} + b_{emd}\right) \tag{2}$$

where (W_{emd}, b_{emd}) are the parameters of embedding layer and $\tau(\cdot)$ is non-linear activation function ReLU. Finally, $f^{func} \in R^{1*256}$ denotes the functional features that will be used to generate PPI-related features and orthology-related features in later steps.

2) Constructing PPI Network and Orthology Network: In this subsection, we construct two networks: a PPI network A_p and an orthology network A_o, which are the foundation for generating PPI-related features and orthology-related features, respectively.

For the construction of the PPI network, the STRING database [39], [40] is used in this study. The database integrates comprehensive data from various sources and provides extensive information on PPIs across different species. To reduce the complexes of our model and noises in the original data, we only retain the interactions with confidence higher than 300. The final PPI network is denoted as $A_p \in R^{n*n}$, where n is the number of proteins and $0 \leq A_p(i,j) \leq 1$ represents the confidence of interaction between protein P_i and P_j, which can be obtained from the STRING database.

Different from PPI network, the orthology network is constructed from the eggNOG database [32], [33]. The database provides comprehensive information on orthology relations and evolutionary histories across different species. The final orthology network can be denoted as $A_o \in \{0,1\}^{n*n}$, where n is the number of proteins and $A_o(i,j) = 1$ indicates the existence of an orthology relation between protein P_i and P_j. Consequently, we construct two different types of networks. The PPI network focuses on the relations between proteins in the same species, while the orthology network reflects the relations between proteins among different species. This ensures a broader coverage of proteins compared to relying on a single network alone.

3) Generating Network-Related Features: For each network A_p and A_o, Graph Convolutional Networks (GCNs) are adopted to aggregate information of neighbor proteins to learn PPI-related features and orthology-related features, respectively. Specifically, functional features $F^{func} = [f_1^{func}; f_2^{func}; \ldots; f_n^{func}] \in R^{n*256}$ are used as node features.

For the $(l+1)$-th layer of GCN, residual operation is added to improve the representation capability of the model, the process of aggregation is as follows:

$$F_{l+1} = \tau\left(\tilde{D}^{-\frac{1}{2}}\tilde{A}\tilde{D}^{-\frac{1}{2}}F_l W_{l+1}^{GCN} + b_{l+1}^{GCN}\right) + F_l \tag{3}$$

where F_l and F_{l+1} refer the output features of l-*th* layer and $(l+1)$-*th* layer, $\tilde{A} \in \{A_p + I, A_o + I\}$ indicates the PPI network or orthology network, $\tilde{D}$ is the corresponding degree matrix of $\tilde{A}$, and $(W_{l+1}^{GCN}, b_{l+1}^{GCN})$ are the parameters of $(l+1)$-*th* GCN layer. After two residual GCN layers, we can obtain the final network-related features. Notably, in this process, separate GCN layers are applied for the PPI network and orthology network, indicating a total of four GCN layers in DeepFMB. Consequently, PPI-related features and orthology-related features can be obtained, denoted as $f^{ppi} \in R^{1*512}$ and $f^{ort} \in R^{1*512}$, respectively.

C. Fusion of Multi-Type Biological Features and Training Model

DeepFMB integrates both PPI-related features and orthology-related features, which are derived from the functional properties obtained through InterPro. The fusion process involves combining these two types of features before incorporating the

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

evolution-related features. Subsequently, DeepFMB utilizes the fused features to predict protein functions.

In this study, we introduce a linear attention mechanism to adaptively fuse two types of network-related features:

$$f^{fuse} = w^{ppi} f^{ppi} + w^{ort} f^{ort} \tag{4}$$

$$w^{ppi}, w^{ort} = softmax\left(tp^{ppi}, tp^{ort}\right) \tag{5}$$

$$tp^{ppi} = \delta\left(f^{ppi} * W_1^{att} + b_1^{att}\right) * W_2^{att} \tag{6}$$

$$tp^{ort} = \delta\left(f^{ort} * W_1^{att} + b_1^{att}\right) * W_2^{att} \tag{7}$$

where w^{ppi} and w^{ort} are learned weighted ($w^{ppi} + w^{ort} = 1$) for fusing f^{ppi} and f^{ort}, which are generated from two shared non-linear layers ($W_1^{att}, b_1^{att}, W_2^{att}$), and $\delta(\cdot)$ is activation function Tanh.

Subsequently, the fused features f^{fuse} are further concatenated with evolution-related features f^{evo} and fed into two non-linear layers to predict protein functions:

$$f^{cob} = \left[f^{fuse}, f^{evo}\right] \tag{8}$$

$$\hat{y} = \tau\left(f^{cob} * W_1^{pred} + b_1^{pred}\right) * W_2^{pred} + b_2^{pred} \tag{9}$$

where $\hat{y} \in R^{1*M}$ is the predicted probability of GO terms. To obtain a stable performance, we store the three best models that achieves better Fmax (the definition can be seen in Section III.B) on the validation data during the training epochs, and average the results predicted by these three models as the final result.

Additionally, we use binary cross entropy as the loss function L:

$$L = -\frac{1}{NM}\sum_{i=1}^{N}\sum_{j=1}^{M} y_{ij}\log\left(\hat{y}_{ij}\right) + \left(1 - y_{ij}\right)\log\left(1 - \hat{y}_{ij}\right) \tag{10}$$

where *N* is the number of proteins, and *M* is the total number of GO terms. If GO term G_j is annotated to protein P_i, y_{ij} is 1. Otherwise, y_{ij} is 0.

D. Extension With Sequence Similarity

To further enhance the predictive performance of DeepFMB, we try to fuse the model predictions with existing homology methods. Previous studies have demonstrated the effectiveness of combining model predictions with homology-based approaches, such as DeepGOPlus [10], DeepGOZero+ [11], and ATGO+ [41]. These methods incorporate the results of homology methods, such as BlastKNN and Diamond, to improve the overall performance. Inspired by these, we combine our DeepFMB prediction results with the results of BlastKNN. The specific processes are as follows:

$$\hat{y}_{BlastKNN}\left(P_i, G_j\right) = \frac{\sum_{P_k \in N(P_i)} I\left(P_k, G_j\right) * Sim\left(P_i, P_k\right)}{\sum_{P_k \in N(P_i)} Sim\left(P_i, P_k\right)} \tag{11}$$

$$\hat{y}_{DeepFMB+}\left(P_i, G_j\right) = a\hat{y}_{DeepFMB}\left(P_i, G_j\right) + \left(1 - a\right)\hat{y}_{BlastKNN}\left(P_i, G_j\right) \tag{12}$$

TABLE I
STATISTIC INFORMATION OF DATASETS GENERATED BY TIME STAMPS

Dataset	BP	CC	MF	ALL
Training	47830	42891	31756	57254
Validation	777	717	684	1432
Test	1070	898	408	1435

where $N(P_i)$ represents the top-k proteins that are most similar to protein P_i, and $I(P_k, G_j)$ equals to 1 if protein P_k has the function G_j, otherwise 0. $Sim(P_i, P_k)$ is the similarity score between P_i and P_k, which is calculated by Blast with cut-off e-value of 0.001. Then, the results of BlastKNN are combined with our model DeepFMB, where a is set to 0.5, 0.5, 1 in BP, MF, CC, respectively. To be noted, we find that BlastKNN reduces the performance of DeepFMB in CC. So we set a to 1.

E. Implemental Details

All parameters are optimized based on the performance of DeepFMB on the validation set. In order to predict as many GO terms as possible, all the GO terms that have appeared in the training set are saved as the predicted labels. During the training and testing processes, we focus on proteins that have at least one GO term in the specific ontology term (BP, MF and CC).

Our deep learning framework is implemented by Pytorch [42], a public deep learning framework developed by Facebook. For two network modules, since the original networks are too large, we keep the 100 largest edges for each node and adopt a mini-batch strategy for training and testing. The batch size and GCN layers are set as 40 and 2, respectively.

Finally, the Adam optimizer is used to train our deep learning framework. The initial learning rate is set to 0.001, the number of epochs is set to 20, batch size is set to 40, and k in BlastKNN is set as 50000.

III. EXPERIMENTS

A. Datasets

In our experiments, we construct an updated dataset following the Critical Assessment of Functional Annotation (CAFA) challenges [3], including protein sequences and their functions from the UniProt database, and GO terms from the Gene Ontology database. For additional biological knowledge used in our model, the interactions and orthology relations are extracted from the STRING database and the eggNOG database, respectively. The details of data process are as follows:

- Protein sequences and functions: We collect protein sequences and their corresponding functions from the UniProt database (April 2022 release). Following the CAFA challenges [3], we retain 23 species with 8 evidence codes, i.e., 'EXP', 'IDA', 'IPI', 'IMP', 'IGI', 'IEP', 'TAS', and 'IC'. After excluding proteins without GO annotations, our final protein dataset comprises 60121 proteins.
- PPI network: The PPI information is obtained from the STRING [40] database (version 11.5 full links). To reduce the noise and complex of the PPI network, we filter the

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

TABLE II
PREDICTIVE PERFORMANCE OF DEEPFMB (+) AGAINST COMPETING METHODS

Methods		Fmax			AUPR		
		MF	CC	BP	MF	CC	BP
Single algorithms	Diamond	0.589	0.571	0.426	0.384	0.281	0.195
	BlastKNN	0.614	0.595	0.443	0.482	0.383	0.257
	DeepGO	0.352	0.579	0.321	0.265	0.583	0.251
	DeepGOA	0.552	0.627	0.400	0.496	0.609	0.334
	DeepGOCNN	0.366	0.563	0.323	0.302	0.570	0.251
	DeepGOZero	0.600	0.613	0.443	0.576	0.572	0.393
	DeepGraphGO	0.548	0.633	0.427	0.515	0.586	0.381
	ATGO	0.455	0.599	0.395	0.437	0.596	0.338
	DeepFMB	**0.624**	**0.652**	**0.453**	**0.604**	**0.690**	**0.409**
Composite algorithms	DeepGOPlus	0.586	0.630	0.437	0.548	0.625	0.366
	DeepGOZero+	0.623	0.633	0.463	0.618	0.592	0.412
	ATGO+	0.619	0.630	0.454	0.593	0.634	0.396
	DeepFMB +	**0.639**	**0.652**	**0.469**	**0.645**	**0.690**	**0.413**

Note: The best performance values are highlighted in bold and the next best performances are underlined.

interactions with a confidence score lower than 300. Finally, 29285116 interactions of 19 species with 196490 proteins are retained.

- Orthology network: Orthology relations are downloaded from the eggNOG [33] database version 5.0, encompassing all proteins used in our study. Finally, 1473178 orthology relations of 22 species between 54959 proteins are retained.
- GO file: The go.obo is downloaded from Gene Ontology [6], (December 4, 2022). As only the 'is-a' and 'part-of' relationships are deemed reliable for grouping protein functions in function propagation, we investigate all annotated GO terms and their ancestor nodes for each protein based on these relationships.

Similar to CAFA [3], as illustrated in Table I, we partitioned the dataset based on time stamps as follows:

- Training: released before 2020-05.
- Validation: released from 2020-05 to 2021-04.
- Test: released from 2021-05 to 2022-04.

Based on the GO terms existing in the training data, there are 19717, 2506, and 6095 GO terms in BP, CC, and MF, respectively.

B. Evaluation Metrics

For the evaluation of these models, we utilize two widely used metrics, Fmax and AUPR. Fmax is the maximum value of F1-score for all proteins and GO terms at different thresholds within the [0, 1] range. Given a threshold t, its corresponding precision and recall for protein P_i can be calculated as follows:

$$Pr_i(t) = \frac{\sum_j I(S(P_i, G_j) \geq t) * I(P_i, G_j)}{\sum_j I(S(P_i, G_j) \geq t)} \tag{13}$$

$$Rc_i(t) = \frac{\sum_j I(S(P_i, G_j) \geq t) * I(P_i, G_j)}{\sum_j I(P_i, G_j)} \tag{14}$$

where $S(P_i, G_j)$ represents the predicted score of function G_j for protein P_i. $I(S(P_i, G_j) \geq t)$ returns 1 when the predicted score is greater than or equal to the threshold t, while it returns 0 otherwise. $I(P_i, G_j)$ equals to 1 when the protein P_i has a function G_j. Otherwise, it equals to 0. Subsequently, the average precision and recall of all proteins can be obtained:

$$AvgPr(t) = \frac{1}{m(t)} * \sum_{i=1}^{m(t)} Pr_i(t) \tag{15}$$

$$AvgRc(t) = \frac{1}{n} * \sum_{i=1}^{n} Rc_i(t) \tag{16}$$

where n is the number of proteins, and $m(t)$ is the number of proteins that accurately predict at least one GO term. Then, Fmax is the best performance of F1-score among different thresholds t:

$$F\max = \max_t \left\{ \frac{2 * AvgPr(t) * AvgRc(t)}{AvgPr(t) + AvgRc(t)} \right\} \tag{17}$$

Similarly, AUPR is the area under the curve of precision-recall formed by different thresholds t. Both higher Fmax and higher AUPR reflect better performance.

C. Performance Comparison With Other Methods

To evaluate the predictive performance of DeepFMB, we compare it with eight state-of-the-art single algorithms: Diamond [12], BlastKNN [13], DeepGO [43], DeepGOA [21], DeepGOCNN [10], DeepGOZero [11], DeepGraphGO [19], ATGO [41]. All the methods are implemented with default parameters. Additionally, we also compare the extended version, DeepFMB+, with other three composite algorithms that are all combined with sequence similarity-based approaches: DeepGOPlus [10], DeepGOZero+ [11] and ATGO+ [41].

Table II presents the performance of different types of methods on the test set. It is remarkable that DeepFMB and DeepFMB+ both achieve the best performance between single algorithms and composite algorithms. Specifically, for single algorithms, in terms of Fmax, DeepFMB yields improvements ranging from 1.6% to 77.3% in MF and from 2.3% to 41.1% in BP. Similarly, DeepFMB surpasses other methods

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

TABLE III
STATISTIC INFORMATION OF GO TERMS GROUPED BY DIFFERENT FREQUENCIES

Ontology	10-30	31-100	101-300	>300
MF	120	173	141	155
CC	76	112	133	169
BP	551	945	945	963

by 4.9%∼127.9% in MF and 4.1%∼109.7% in BP in terms of AUPR. Notably, the advantages of DeepFMB in CC are more pronounced, outperforming other approaches by at least 3% and 10% in terms of Fmax and AUPR, respectively. These results demonstrate that DeepFMB can predict protein functions more accurately. Additionally, although DeepGO and DeepGOA utilize orthology relations, their performance is poorer compared to DeepFMB, proving that GNNs are more effective in capturing relationships between proteins. Furthermore, compared to DeepGraphGO that also uses GNNs to aggregate InterPro features via PPI network, DeepFMB outperforms DeepGraphGO in the three sub-ontologies, especially in MF, where Fmax and AUPR are improved by 13.9% and 17.3%, respectively. This suggests that the orthology-related features and evolution-related features contribute to our model significantly. Similar conclusions can also be obtained from Supplementary Table S1 in terms of AUC and ACC.

Meanwhile, for composite algorithms that incorporating sequence similarity-based methods, DeepFMB+ achieves the best performance in all sub-ontologies. As for Fmax, there is an improvement of 1.3%∼7.3%, 2.6%∼9.0%, and 3.0%∼3.5% in BP, MF, and CC, respectively. As for AUPR, there is also an improvement of 0.2%∼12.8%, 4.4%∼17.7%, and 8.8%∼16.6% in BP, MF, and CC, respectively.

Consequently, DeepFMB+ is further improved after considering the sequence similarities between proteins. It is worth noting that DeepFMB performs well in CC and introducing BlastKNN will reduce its performance, which may indicate that sequence similarity occupies a significant role in MF and BP, excluding CC. Above all, the results illustrate that DeepFMB and DeepFMB+ hold the ability to predict protein functions accurately.

D. Performance Comparison on GO Groups With Different Frequencies

In the loosely hierarchy of GO terms, the deeper the depth of the GO term tends to represent more specific function and often corresponds to fewer proteins with this function. Consequently, it is more challenging and meaningful to accurately predict GO terms with fewer samples. In this section, we further evaluate these methods on different GO terms grouped by frequencies (the number of known samples).

As shown in Table III, according to the number of annotations per GO term, we categorize the GO terms in the training set into 4 groups: 10-30, 31-100, 101-300 and >300. Table IV illustrates the performance of these methods on different groups in terms of M-AUPR, which is the average of AUPR on each GO term. Notably, DeepFMB outperforms all other single approaches in all sub-ontologies, with only DeepGOZero slightly better when the number of annotations in BP is greater than 300. Consistent improvements demonstrate that DeepFMB holds stable advantages on different GO terms. Similar conclusions can also be obtained from DeepFMB+. Additionally, compared with DeepGOCNN and DeepGOPlus, DeepGOZero and DeepGOZero+, ATGO and ATGO+, DeepFMB and DeepFMB+, there is a common trend that the performance of all these methods is significantly improved after incorporating sequence similarity-based methods. Additionally, the one-side Wilcoxon test is provided (see Table IV), and the results also demonstrate the effective of our methods in protein function predictions.

E. Ablation Studies

To understand the specific contributions of each component in DeepFMB, we design ablation studies addressing several key questions:

- Does orthology network facilitate DeepFMB's inference of protein function?
- Does PPI network facilitate DeepFMB's inference of protein function?
- How powerful are the evolution-related features generated from pre-trained model?
- Is it helpful to fuse multi-type biological features?

Therefore, we design the following four variant models to verify our conjectures:

- DeepFMB_only_evo: only uses evolution-related features extracted by ESM-1b.
- DeepFMB_no_evo: removes evolution-related features extracted by ESM-1b, only keeps PPI-related features and orthology-related features.
- DeepFMB_no_ort: excludes the orthology network, utilizing evolution-related features and PPI-related features.
- DeepFMB_no_PPI: excludes the PPI network, utilizing evolution-related features and orthology-related features.

The performance of these models is shown in Fig. 2. Our results show that: (i) Comparing DeepFMB with DeepFMB_no_ort, it can be found that the performance in MF decreases significantly after removing orthogonal correlation features, while the performance in BP and CC is stable. This indicates that orthology relations may have a more pronounced effect in MF than BP and CC. (ii) Comparing DeepFMB with DeepFMB_no_PPI, a significant decrease can be found in MF and CC, which indicates that the PPI information also contributes protein function prediction in MF and CC. (iii) Comparing DeepFMB with DeepFMB_no_evo, in CC and MF, there is a significant decrease in terms of AUPR, and a slight decrease in Fmax. It also proves that the evolution-related features play a key role in MF and CC. (iv) Comparing DeepFMB with DeepFMB_only_evo reveals a common observation in all conditions. After removing PPI-related and orthology-related features, the performance of our model drops drastically in terms of Fmax and AUPR in MF, CC, and BP. It demonstrates that functional features (InterPro properties), PPI network and orthology network together contribute to protein function

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

TABLE IV
PREDICTIVE PERFORMANCE ON GO TERMS WITH DIFFERENT FREQUENCIES IN TERMS OF M-AUPR

Ontology	Method	10-30		31-100		101-300		>300	
		M-AUPR(std)	p-value	M-AUPR(std)	p-value	M-AUPR(std)	p-value	M-AUPR(std)	p-value
MF	Diamond	0.24(0.25)	6e-5	0.25(0.26)	4e-7	0.26(0.26)	2e-7	0.31(0.24)	5e-22
	BlastKNN	0.29(0.25)	0.06	0.31(0.27)	0.01	0.29(0.28)	2e-5	0.36(0.26)	2e-15
	DeepGO	0	-	0.03(0.11)	2e-25	0.11(0.17)	1e-15	0.21(0.18)	1e-20
	DeepGOA	0	-	0.17(0.26)	5e-15	0.34(0.29)	6e-3	0.44(0.29)	0.02
	DeepGOCNN	0	-	0.07(0.16)	2e-22	0.13(0.19)	6e-16	0.26(0.22)	8e-18
	DeepGOZero	0.30(0.25)	0.34	0.33(0.25)	3e-3	0.35(0.28)	0.10	0.44(0.28)	4e-4
	DeepGraphGO	0.24(0.24)	0.01	0.26(0.27)	2e-7	0.32(0.28)	6e-4	0.41(0.30)	1e-10
	ATGO	0.25(0.24)	4e-6	0.26(0.24)	1e-11	0.26(0.27)	2e-11	0.33(0.28)	1e-22
	DeepFMB	**0.31(0.26)**	-	**0.35(0.27)**	-	**0.36(0.29)**	-	**0.47(0.29)**	-
	DeepGOPlus	0.30(0.25)	3e-4	0.30(0.26)	1e-7	0.34(0.29)	8e-5	0.42(0.29)	5e-12
	DeepGOZero+	0.33(0.25)	0.67	0.38(0.26)	0.32	0.38(0.28)	0.28	0.47(0.29)	1e-5
	ATGO+	0.31(0.25)	3e-3	0.32(0.25)	3e-6	0.32(0.29)	3e-6	0.43(0.30)	8e-15
	DeepFMB+	**0.33(0.26)**	-	**0.38(0.27)**	-	**0.38(0.29)**	-	**0.49(0.30)**	-
CC	Diamond	0.16(0.24)	5e-6	0.18(0.24)	9e-4	0.14(0.18)	2e-9	0.17(0.15)	7e-19
	BlastKNN	0.21(0.27)	2e-3	0.21(0.26)	0.02	0.14(0.19)	1e-10	0.21(0.18)	1e-18
	DeepGO	0	-	0.05(0.16)	8e-16	0.09(0.16)	4e-14	0.20(0.22)	9e-21
	DeepGOA	0	-	0.04(0.12)	5e-16	0.17(0.22)	1e-3	0.29(0.26)	3e-5
	DeepGOCNN	0	-	0.07(0.15)	6e-11	0.11(0.17)	4e-9	0.21(0.22)	1e-17
	DeepGOZero	0.21(0.26)	0.02	0.21(0.25)	0.07	0.19(0.23)	1e-4	0.26(0.24)	5e-13
	DeepGraphGO	0.16(0.24)	1e-3	0.20(0.25)	0.01	0.20(0.24)	0.01	0.31(0.26)	0.01
	ATGO	0.15(0.21)	5e-7	0.16(0.21)	1e-6	0.11(0.18)	1e-12	0.25(0.26)	3e-12
	DeepFMB	**0.25(0.27)**	-	**0.24(0.26)**	-	**0.21(0.24)**	-	**0.32(0.26)**	-
	DeepGOPlus	0.19(0.27)	4e-6	0.22(0.26)	2e-4	0.21(0.24)	5e-3	**0.32(0.27)**	3e-6
	DeepGOZero+	0.23(0.27)	0.02	**0.28(0.28)**	0.59	0.21(0.25)	2e-3	0.31(0.26)	7e-10
	ATGO+	0.23(0.27)	3e-3	0.24(0.26)	6e-5	0.17(0.24)	2e-7	0.31(0.28)	1e-9
	DeepFMB+	**0.25(0.29)**	-	0.24(0.28)	-	**0.21(0.25)**	-	0.32(0.27)	-
BP	Diamond	0.12(0.21)	4e-18	0.12(0.21)	3e-32	0.11(0.19)	8e-21	0.17(0.14)	4e-114
	BlastKNN	0.13(0.22)	3e-10	0.13(0.22)	5e-14	0.13(0.20)	5e-10	0.16(0.17)	1e-96
	DeepGO	0	-	0	-	0.01(0.05)	2e-137	0.11(0.14)	2e-99
	DeepGOA	0	-	0	-	0	-	0.16(0.19)	1e-51
	DeepGOCNN	0	-	0.04(0.11)	5e-79	0.07(0.14)	3e-42	0.13(0.14)	1e-54
	DeepGOZero	0.12(0.21)	3e-6	0.14(0.22)	6e-3	0.15(0.21)	4e-3	**0.22(0.22)**	0.20
	DeepGraphGO	0.10(0.19)	2e-4	0.12(0.21)	2e-7	0.13(0.21)	4e-9	0.20(0.20)	5e-7
	ATGO	0.07(0.16)	2e-53	0.08(0.17)	4e-73	0.08(0.15)	2e-65	0.16(0.19)	2e-61
	DeepFMB	**0.14(0.22)**	-	**0.15(0.23)**	-	**0.15(0.22)**	-	0.21(0.21)	-
	DeepGOPlus	0.12(0.20)	8e-38	0.140(0.22)	1e-15	0.15(0.22)	2e-12	0.20(0.20)	1e-27
	DeepGOZero+	0.14(0.22)	1e-6	0.16(0.24)	6e-3	0.17(0.23)	3e-4	0.23(0.22)	0.03
	ATGO+	0.13(0.21)	8e-32	0.14(0.22)	7e-45	0.14(0.22)	1e-56	0.21(0.22)	9e-43
	DeepFMB+	**0.16(0.23)**	-	**0.17(0.24)**	-	**0.18(0.24)**	-	**0.23(0.22)**	-

Note: The best performance values are highlighted in bold and the next best performance are underlined.

prediction. Above all, all types of biological knowledge used in DeepFMB are significant and different types of biological knowledge contribute to different ontologies.

F. Case Study

To further compare DeepFMB with other methods, we use two proteins (HS3S1_MOUSE and ERG13_SCHPO) in the test set to show the predictive performance of these methods. As can be seen in Table V, HS3S1_MOUSE is experimentally annotated by 7 GO terms. However, DeepGraphGO predicts 4 GO terms and only 2 of them are correct. DeepGO predicts 4 GO terms and only 3 of them are correct. DeepGOCNN predicts 12 GO terms and only 3 of them are predicted correctly. Diamond, BlastKNN, DeepGOA, DeepGOZero, DeepGOPlus, DeepGOZero+, and ATGO+ all predict 5 GO terms, all of

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

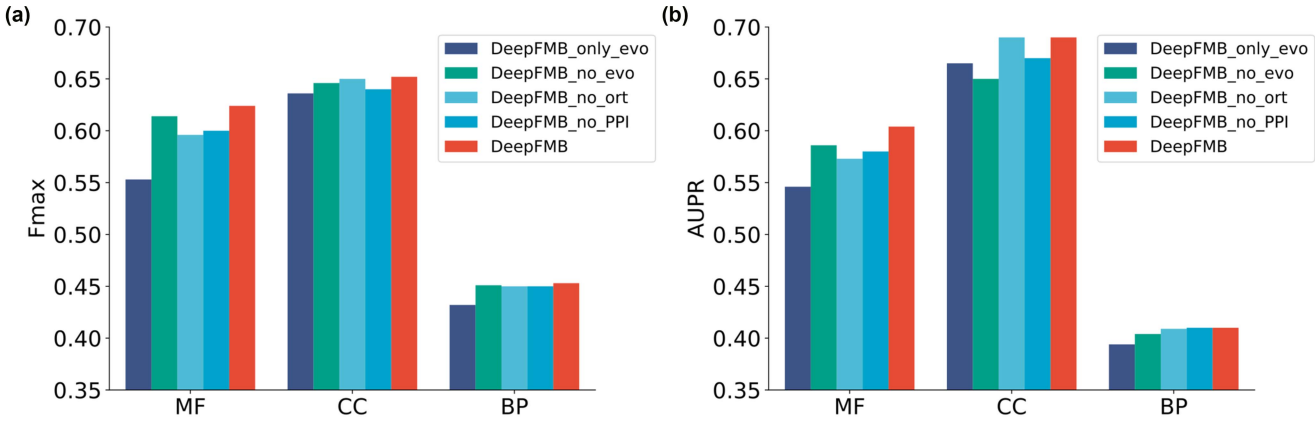

Fig. 2. The predictive performance of different variant models in terms of F-max and AUPR.

TABLE V
PREDICTIVE PERFORMANCE OF VARIOUS METHODS ON HS3S1_MOUSE

Method	Predicted GO terms	Precision	Recall	F1
Diamond	**GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	1.00	0.71	0.83
BlastKNN	**GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	1.00	0.71	0.83
DeepGO	GO:0005488, **GO:0003674, GO:0016740, GO:0003824**	0.75	0.43	0.54
DeepGOA	**GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	1.00	0.71	0.83
DeepGOCNN	GO:0016787, GO:0098772, GO:0005102, GO:0016757, **GO:0016740, GO:0003824**, GO:0016758, GO:0005488, **GO:0003674**, GO:0016788, GO:0005515	0.27	0.43	0.33
DeepGOZero	**GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	1.00	0.71	0.83
DeepGraphGO	GO:0016705, **GO:0003824, GO:0003674**, GO:00016491	0.50	0.29	0.37
ATGO	GO:0016787, **GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	0.86	0.71	0.78
DeepFMB	**GO:0016740, GO:0003824, GO:0016782, GO:0034483, GO:0003674, GO:0008146**	1.00	0.86	0.92
DeepGOPlus	**GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	1.00	0.71	0.83
DeepGOZero+	**GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	1.00	0.71	0.83
ATGO+	**GO:0016740, GO:0003824, GO:0016782, GO:0003674, GO:0008146**	1.00	0.71	0.83
DeepFMB +	**GO:0016740, GO:0003824, GO:0016782, GO:0034483, GO:0003674, GO:0008146**	1.00	0.86	0.92
Ground truth	**GO:0008467, GO:0016740, GO:0003824, GO:0016782, GO:0034483, GO:0003674, GO:0008146**	-	-	-

Note: The GO terms predicted accurately are highlighted in bold.

which are predicted correctly. While ATGO also correctly predicts these 5 GO terms, it predicts a GO term incorrectly, whereas ATGO+ can make corrections when predicting this GO term, making its prediction more accurate. It can also be obtained that Diamond and BlastKNN methods achieve comparable performance to several DL-based methods, indicating that using sequence similarity can work well for function prediction, not only for enhancement, but also for prediction alone. Our models, DeepFMB and DeepFMB+, both predict 6 correct GO terms and predict GO:0034483 which is not predicted by the rest of the methods, reflecting the effectiveness of our models. Additionally, Table VI illustrates the performance of existing methods on ERG13_SCHPO, which is annotated by 70 GO terms. It is worth noting that sequence similarity-based approaches (eg. BlastKNN and Diamond) outperform almost other deep learning-based methods, except DeepFMB and DeepFMB+. Specifically, BlastKNN and Diamond achieve F1 socre of 0.95 and 0.91, respectively. Our method, DeepFMB, gets better performance with a precision value of 1.00, a recall value of 0.99 and an F1 value of 0.99. After incorporating BlastKNN, our method, DeepFMB+, can accurately annotate ERG13_SCHPO perfectly with the best F1 score of 1.00. These cases further demonstrate the practicality of DeepFMB and DeepFMB+.

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

TABLE VI
PREDICTIVE PERFORMANCE OF VARIOUS METHODS ON ERG13_SCHPO

Methods	Precision	Recall	F1
Diamond	1.00	0.84	0.91
BlastKNN	1.00	0.91	<u>0.95</u>
DeepGO	0.61	0.54	0.57
DeepGOA	0.90	0.53	0.67
DeepGOCNN	0.62	0.49	0.55
DeepGOZero	0.79	0.49	0.60
DeepGraphGO	0.93	0.36	0.52
ATGO	0.69	0.16	0.26
DeepFMB	1.00	0.99	**0.99**
DeepGOPlus	0.93	0.76	0.84
DeepGOZero+	0.92	0.86	<u>0.89</u>
ATGO+	1.00	0.49	0.66
DeepFMB+	1.00	1.00	**1.00**

Note: The best performance values are highlighted in bold and the next best performances are underlined.

IV. DISCUSSIONS AND CONCLUSION

With the development of high-throughput technology and deep learning technologies, protein-related data are constantly being produced, including protein sequences and PPI networks, which provide more opportunities for computational methods to predict protein functions. Although several existing approaches have shown that evolution information of protein sequences is highly correlated with functions, generating evolution information by traditional methods are time-consuming. Besides, as proteins perform their functions with interactions, PPI information is also utilized by several methods to predict protein functions. However, these methods perform poorly on proteins lacking interactions. To address these limitations, we propose a noval deep learning-based model, DeepFMB, which integrates multi-type biological knowledge, including protein sequences, protein-protein interactions and orthology relations. DeepFMB utilizes pre-trained pLLMs to generate evolution-related features, which is much faster than traditional computational methods. Simultaneously, DeepFMB introduces orthology relations to compensate for the lack of interactions in some proteins.

Experimental results indicate that DeepFMB surpasses competing methods across all sub-ontologies, particularly in CC. Further enhancement is achieved by combining DeepFMB with BlastKNN, forming DeepFMB+. This combination improves performance in BP and MF, highlighting the importance of sequence similarity in these ontologies. Additionally, we also compare the performance of these methods on different GO terms categorized by their frequencies, demonstrating the capability to effectively predict GO terms that are few annotated to proteins. Ablation studies verify that our proposed framework can fuse multi-type biological knowledge effectively and different types of functions focus on different biological knowledge. Specifically, sequence information may be more crucial on CC, while orthology relations are more closely related to MF. Finally, we use a typical protein case to show the practicality of our method.

As for future work, we are probably to integrate other heterogeneous data, such as gene co-expression [2], protein structure [8], text mining [44], and other biological networks [45], to explore more possibilities to enhance protein function prediction.

ACKNOWLEDGMENT

Part of this paper is published in the IEEE International Conference on Bioinformatics and Biomedicine (BIBM) 2023 [46].

REFERENCES

[1] B. Alberts, "The cell as a collection of protein machines: Preparing the next generation of molecular biologists," *Cell*, vol. 92, no. 3, pp. 291–294, 1998.
[2] M. Li, W. Shi, F. Zhang, M. Zeng, and Y. Li, "A deep learning framework for predicting protein functions with co-occurrence of GO terms," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 20, no. 2, pp. 833–842, Mar./Apr. 2023.
[3] N. Zhou et al., "The CAFA challenge reports improved protein function prediction and new functional annotations for hundreds of genes through experimental screens," *Genome Biol.*, vol. 20, no. 1, pp. 1–23, 2019.
[4] S. Köhler et al., "The human phenotype ontology in 2021," *Nucleic Acids Res.*, vol. 49, no. D1, pp. D1207–D1217, 2021.
[5] S. De Cnudde, Y. Ramon, D. Martens, and F. Provost, "Deep learning on big, sparse, behavioral data," *Big Data*, vol. 7, no. 4, pp. 286–307, 2019.
[6] G. O. Consortium, "The gene ontology resource: 20 years and still GOing strong," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D330–D338, 2019.
[7] W. Wang, Y. Shuai, Q. Yang, F. Zhang, M. Zeng, and M. Li, "A comprehensive computational benchmark for evaluating deep learning-based protein function prediction approaches," *Brief. Bioinf.*, vol. 25, no. 2, 2024, Art. no. bbae050.
[8] B. Lai and J. Xu, "Accurate protein function prediction via graph attention networks with predicted structure information," *Brief. Bioinf.*, vol. 23, no. 1, 2022, Art. no. bbab502.
[9] F. Zhang, H. Song, M. Zeng, Y. Li, L. Kurgan, and M. Li, "DeepFunc: A deep learning framework for accurate prediction of protein functions from protein sequences and interactions," *Proteomics*, vol. 19, no. 12, 2019, Art. no. 1900019.
[10] M. Kulmanov and R. Hoehndorf, "DeepGOPlus: Improved protein function prediction from sequence," *Bioinformatics*, vol. 36, no. 2, pp. 422–429, 2020.
[11] M. Kulmanov and R. Hoehndorf, "DeepGOZero: Improving protein function prediction from sequence and zero-shot learning based on ontology axioms," *Bioinformatics*, vol. 38, no. Supplement_1, pp. i238–i245, 2022.
[12] B. Buchfink, K. Reuter, and H.-G. Drost, "Sensitive protein alignments at tree-of-life scale using DIAMOND," *Nature Methods*, vol. 18, no. 4, pp. 366–368, 2021.
[13] S. F. Altschul et al., "Gapped BLAST and PSI-BLAST: A new generation of protein database search programs," *Nucleic Acids Res.*, vol. 25, no. 17, pp. 3389–3402, 1997.
[14] Y. Cao and Y. Shen, "TALE: Transformer-based protein function Annotation with joint sequence-Label Embedding," *Bioinformatics*, vol. 37, no. 18, pp. 2825–2833, 2021.
[15] S. Albawi, T. A. Mohammed, and S. Al-Zawi, "Understanding of a convolutional neural network," in *Proc. Int. Conf. Eng. Technol.*, 2017, pp. 1–6.
[16] A. Vaswani et al., "Attention is all you need," in *Proc. Adv. Neural Inf. Process. Syst.*, 2017, pp. 6000–6010.
[17] W. Wang, X. Meng, J. Xiang, Y. Shuai, H. D. Bedru, and M. Li, "CACO: A core-attachment method with cross-species functional ortholog information to detect human protein complexes," *IEEE J. Biomed. Health Inform.*, vol. 27, no. 9, pp. 4569–4578, 2023.
[18] M. Zeng, F. Zhang, F.-X. Wu, Y. Li, J. Wang, and M. Li, "Protein–protein interaction site prediction through combining local and global features with deep neural networks," *Bioinformatics*, vol. 36, no. 4, pp. 1114–1120, 2020.
[19] R. You, S. Yao, H. Mamitsuka, and S. Zhu, "DeepGraphGO: Graph neural network for large-scale, multispecies protein function prediction," *Bioinformatics*, vol. 37, no. Supplement_1, pp. i262–i271, 2021.

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.

[20] Z. Wu, S. Pan, F. Chen, G. Long, C. Zhang, and S. Y. Philip, "A comprehensive survey on graph neural networks," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 1, pp. 4–24, Jan. 2021.
[21] F. Zhang et al., "A deep learning framework for gene ontology annotations with sequence-and network-based information," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 18, no. 6, pp. 2208–2217, Nov./Dec. 2021.
[22] P. Zhou et al., "Attention-based bidirectional long short-term memory networks for relation classification," in *Proc. 54th Annu. Meeting Assoc. Comput. Linguistics*, 2016, pp. 207–212.
[23] R. .. Apweiler et al., "The InterPro database, an integrated documentation resource for protein families, domains and functional sites," *Nucleic Acids Res.*, vol. 29, no. 1, pp. 37–40, 2001.
[24] R. D. Finn et al., "InterPro in 2017–beyond protein family and domain annotations," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D190–D199, 2017.
[25] A. Mitchell et al., "The InterPro protein families database: The classification resource after 15 years," *Nucleic Acids Res.*, vol. 43, no. D1, pp. D213–D221, 2015.
[26] A. Grover and J. Leskovec, "node2vec: Scalable feature learning for networks," in *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Mining*, 2016, pp. 855–864.
[27] M. Steinegger, M. Meier, M. Mirdita, H. Vöhringer, S. J. Haunsberger, and J. Söding, "HH-suite3 for fast remote homology detection and deep protein annotation," *BMC Bioinf.*, vol. 20, pp. 1–15, 2019.
[28] A. Rives et al., "Biological structure and function emerge from scaling unsupervised learning to 250 million protein sequences," *Proc. Nat. Acad. Sci. USA*, vol. 118, no. 15, 2021, Art. no. e2016239118.
[29] A. Elnaggar et al., "Prottrans: Toward understanding the language of life through self-supervised learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 44, no. 10, pp. 7112–7127, Oct. 2022.
[30] H. Wang, "Scientific discovery in the age of artificial intelligence," *Nature*, vol. 620, no. 7972, pp. 47–60, 2023.
[31] Z. Lin et al., "Evolutionary-scale prediction of atomic-level protein structure with a language model," *Science*, vol. 379, no. 6637, pp. 1123–1130, 2023.
[32] J. Huerta-Cepas, D. Szklarczyk, and K. Forslund, "eggNOG 4.5: A hierarchical orthology framework with improved functional annotations for eukaryotic, prokaryotic and viral sequences," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D286–D293, 2016.
[33] J. Huerta-Cepas et al., "eggNOG 5.0: A hierarchical, functionally and phylogenetically annotated orthology resource based on 5090 organisms and 2502 viruses," *Nucleic Acids Res.*, vol. 47, no. D1, pp. D309–D314, 2019.
[34] P. Jones et al., "InterProScan 5: Genome-scale protein function classification," *Bioinformatics*, vol. 30, no. 9, pp. 1236–1240, 2014.
[35] I. Letunic, T. Doerks, and P. Bork, "SMART: Recent updates, new developments and status in 2015," *Nucleic Acids Res.*, vol. 43, no. D1, pp. D257–D260, 2015.
[36] A. Marchler-Bauer, J. B. Anderson, F. Chitsaz, M. K. Derbyshire, and S. H. Bryant, "Marchler-Bauer, A CDD: Specific functional annotation with the conserved domain database," *Nucleic Acids Res.*, vol. 37, pp. D205–D210, 2008.
[37] R. D. Finn et al., "The Pfam protein families database: Towards a more sustainable future," *Nucleic Acids Res.*, vol. 44, no. D1, pp. D279–D285, 2016.
[38] M. E. Oates, "The SUPERFAMILY 1.75 database in 2014: A doubling of data," *Nucleic Acids Res.*, vol. 43, no. D1, pp. D227–D233, 2015.
[39] D. Szklarczyk et al., "The STRING database in 2011: Functional interaction networks of proteins, globally integrated and scored," *Nucleic Acids Res.*, vol. 39, no. Suppl 1, pp. D561–D568, 2011.
[40] D. Szklarczyk et al., "The STRING database in 2017: Quality-controlled protein–protein association networks, made broadly accessible," *Nucleic Acids Res.*, vol. 45, no. D1, pp. D362–D368, 2016.
[41] Y.-H. Zhu, C. Zhang, D.-J. Yu, and Y. Zhang, "Integrating unsupervised language model with triplet neural networks for protein gene ontology prediction," *PLOS Comput. Biol.*, vol. 18, no. 12, 2022, Art. no. e1010793.
[42] A. Paszke et al., "PyTorch: An imperative style, high-performance deep learning library," in *Proc. Adv. Neural Inf. Process. Syst.*, 2019, vol. 32.
[43] M. Kulmanov, M. A. Khan, and R. Hoehndorf, "DeepGO: Predicting protein functions from sequence and interactions using a deep ontology-aware classifier," *Bioinformatics*, vol. 34, no. 4, pp. 660–668, Feb. 15, 2018.
[44] R. Cao and J. Cheng, "Integrated protein function prediction by mining function associations, sequences, and protein–protein and gene–gene interaction networks," *Methods*, vol. 93, pp. 84–91, 2016.
[45] W. Peng, M. Li, L. Chen, and L. Wang, "Predicting protein functions by using unbalanced random walk algorithm on three biological networks," *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 14, no. 2, pp. 360–369, 2015.
[46] Y. Shuai, W. Wang, Y. Li, M. Zeng, and M. Li, "Protein function prediction using graph neural network with multi-type biological knowledge," in *Proc. IEEE Int. Conf. Bioinf. Biomed.*, 2023, pp. 30–35.

Wenkang Wang received the BS degree from Central South University, China, in 2020. He is currently working toward the PhD degree in computer science and engineering with Central South University, China. His currently research interests include bioinformatics, complex network analysis, machine learning, and deep learning.

Yunyan Shuai received the BS degree from Central South University, China, in 2022. She is currently working toward the MS degree in computer science and engineering with Central South University, China. Her currently research interests include bioinformatics and deep learning.

Yiming Li received the BS degree from Xiang-tan University in 2019 and the MS degree from Central South University in 2022. She is currently working toward the PhD degree with Central South University, ChangSha, China. Her currently research interests include bioinformatics, machine learning, and deep learning.

Min Zeng received the BS degree from Lanzhou University in 2013 the MS and PhD degrees from Central South University, in 2016 and 2020, respectively. He is currently an associate professor with the School of Computer Science and Engineering, Central South University. His main research interests include bioinformatics, machine learning, and deep learning.

Min Li received the PhD degree in computer science from Central South University, China, in 2008. Currently, she is a professor and dean with the School of Computer Science and Engineering, Central South University, China. Her main research interests include bioinformatics and systems biology. She has published more than 100 technical papers in referred journals such as *Nature Communications*, *Genome Research, Bioinformatics, Briefings in Bioinformatics*, *IEEE/ACM Transactions on Computational Biology and Bioinformatics*, and conference proceedings such as BIBM, GIW and ISBRA. According to Google scholar, her paper citations are more than 14000 and H-index is 63.

Authorized licensed use limited to: Nanjing Agricultural University. Downloaded on April 17,2025 at 08:08:15 UTC from IEEE Xplore. Restrictions apply.