

## Article

# Grain Protein Function Prediction Based on CNN and Residual Attention Mechanism with AlphaFold2 Structure Data

Jing Liu <sup>1</sup>, Xinping Zhang <sup>1</sup> , Kai Huang <sup>2,3</sup>, Yuqi Wei <sup>1</sup> and Xiao Guan <sup>2,3,\*</sup>

<sup>1</sup> College of Information Engineering, Shanghai Maritime University, Shanghai 201306, China; jingliu@shmtu.edu.cn (J.L.); 202230310091@stu.shmtu.edu.cn (X.Z.); 202430310291@stu.shmtu.edu.cn (Y.W.)

<sup>2</sup> School of Health Science and Engineering, University of Shanghai for Science and Technology, Shanghai 200093, China; hk@usst.edu.cn

<sup>3</sup> National Grain Industry (Urban Grain and Oil Security) Technology Innovation Center, Shanghai 200093, China

\* Correspondence: gnxo@usst.edu.cn

**Abstract:** The prediction of grain protein function is essential for the advancement of food science. Traditional experimental methods are associated with high costs and significant time requirements. Computational methods are recognized for their efficiency and reduced time demands. A new multimodal deep learning method, MMSNet, is proposed in this study, and protein data of four types of grains (japonica, indica, maize, and wheat) are analyzed. This method fuses the protein structure information predicted by AlphaFold2 and combines a multiscale one-dimensional convolutional neural network (1DCNN) with a two-dimensional convolutional neural network (2DCNN) to enable the model to capture sequence and structural information effectively. We used a residual attention mechanism to replace the traditional pooling layer, thereby improving the feature extraction capability of the network layers in 2DCNN. The experimental results indicate that secondary structure and spatial structure information contribute to improving model performance. Compared with two classical methods, MMSNet demonstrates optimal performance, which validates the effectiveness of our approach in integrating complex grain protein data and highlights its potential to open new avenues for grain protein function prediction.



Academic Editor: Hamid Mcheick

Received: 17 December 2024

Revised: 6 February 2025

Accepted: 8 February 2025

Published: 12 February 2025

**Citation:** Liu, J.; Zhang, X.; Huang, K.; Wei, Y.; Guan, X. Grain Protein Function Prediction Based on CNN and Residual Attention Mechanism with AlphaFold2 Structure Data. *Appl. Sci.* **2025**, *15*, 1890. <https://doi.org/10.3390/app15041890>

**Copyright:** © 2025 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** grain protein function prediction; deep learning; residual attention; convolutional neural network; secondary structure; alphafold2

## 1. Introduction

With the growing global population and increasing demand for healthy diets, food science and technology are facing unprecedented challenges and opportunities [1]. As an indispensable part of the human diet, grains are not only a source of primary energy and nutrition but are also utilized as essential raw materials in the food industry owing to their rich protein content [2]. The unique functional properties of grain proteins [3], such as solubility, emulsification, and gelling ability, influence the texture, taste, and nutritional value of foods. This study focuses on predicting the biological functions of grain proteins based on gene ontology (GO) terms, and the study of functional properties is beyond the scope of this study. Gene ontology describes the molecular function, biological process, or cellular component roles of proteins. The gene ontology function prediction in this study used protein amino acid sequence data and structural data. Functional properties (such as solubility, emulsification, and gelling ability) are determined by the molecular structure and chemical properties of proteins [4]. Gene ontology function prediction can

provide indirect reference data for studying functional properties. Therefore, it is necessary to understand and predict the function of grain proteins.

Traditional protein function prediction relies on time-consuming and labor-intensive biological experiments. However, with the development of high-throughput sequencing technology and bioinformatics, computational methods are now being used to effectively accelerate the annotation of protein function. Earlier methods, such as BLAST [5], transferred protein functional annotations based on sequence similarity. With the development of artificial intelligence, many machine-learning methods have been widely used to predict protein function with good results [6–9]. However, redundancy and noise problems resulting from the accumulation of large amounts of protein data have become bottlenecks that limit the performance improvement of machine-learning algorithms. Kulmanov first used a convolutional neural network (CNN) to construct a deep-learning model, DeepGO [10], and combined protein sequences and protein–protein interaction network features to predict protein functions. In the later proposed DeepGOPlus [11], Diamond (a fast sequence alignment tool) was used to obtain sequence similarity information [12], and a CNN was combined to extract sequence features to improve prediction performance. Protein function is not only determined by its primary structure (sequence) but is also closely related to its specific spatial structure. The folding and spatial conformation of proteins are decisive factors in determining their functions [13,14]. DeepFRI uses a graph convolutional network (GCN) to link sequence and structural data and achieves better performance in protein function prediction by using experimentally determined protein structure data [15]. Deep learning methods can accurately capture subtle patterns closely related to functions in massive amounts of protein data owing to their powerful pattern recognition and automatic feature extraction capabilities [16].

Previous research on protein function prediction was mostly limited to protein sequence data or relied on the acquisition of expensive experimental structural data. However, a new perspective on protein function prediction was provided through a breakthrough in protein structure prediction achieved using AlphaFold2 [17]. The accuracy of protein structure prediction has been improved to the atomic level using AlphaFold2, providing high-quality structural predictions close to the experimental levels. The availability of such high-quality structural data has greatly enriched the resources for computational approaches. TransFun extracts features from protein sequences using a pre-trained protein language model (evolutionary scale modeling, ESM) via transfer learning and combines them with protein 3D structures through equivariant graph neural networks [18]. The Struct2GO model learns the embedding of amino acid residues using graphs [19]. A graph pooling algorithm based on a self-attention mechanism was used to extract the overall graph structural features, which were then fused with sequence features obtained from protein language models to improve the accuracy of the protein function prediction and the generality of the model. Therefore, a shift has been witnessed in the frontier field of protein function prediction from single-sequence information to combined structural data and from single-sequence alignment methods to complex deep learning methods.

The multimodal MMSNet algorithm based on multi-source data fusion is proposed for grain protein function prediction in this paper, which integrates protein sequence information and high-quality protein structure information predicted by AlphaFold2. Secondary and spatial structural information of amino acids were extracted from structural data, and convolutional neural networks were employed to learn feature representations for one-dimensional sequence data and two-dimensional image data, respectively. The proposed method was compared with two representative methods, and a significant improvement in overall performance was observed. The effectiveness of the extracted secondary structure information in improving the prediction performance was demonstrated, and in-

corporating a residual attention mechanism into a 2DCNN to extract 3D structural features significantly enhanced the accuracy and stability of model predictions.

## 2. Materials and Methods

### 2.1. Datasets

Protein sequence data with GO annotations for japonica, indica, maize, and wheat were obtained from the UniProtKB-SwissProt database. All protein data in the UniProtKB-SwissProt database were carefully verified through a literature review and computer tools [20], providing high-quality grain protein data for this experiment. The corresponding grain protein structure data predicted by AlphaFold2 were obtained from the AlphaFold Protein Structure Database [21]. Simultaneously, the gene ontology's OBO format data were obtained from <https://geneontology.org/docs/download-ontology/> (accessed on 2 October 2023), and the gene ontology hierarchy was parsed.

Gene ontology annotation is widely used as a classification scheme for protein function prediction [22]. To describe the biological concepts of protein function, protein function categories were organized into three sub-ontologies: molecular function, biological process, and cellular component. Several GO terms can be used to annotate proteins. For each grain protein dataset, separate datasets were constructed for MF, BP, and CC sub-ontologies in the initial data preprocessing stage based on previous studies. Some proteins only possess functions that belong to a specific ontology. Real-path rules are used to propagate annotations in the GO hierarchy; i.e., if a protein is annotated with a specific GO term, then the ancestor terms of that term are also used for protein annotation. If a protein contains GO term annotations from three sub-ontologies, it is simultaneously included in the MF, BP, and CC sub-ontology datasets. To reduce GO term sparsity, we filtered out GO categories annotated with fewer than a certain threshold number of proteins from each dataset. Protein sequences with a length greater than 1002 or ambiguous amino acid codes (B, O, J, U, X, and Z) were discarded.

The processed dataset was randomly divided into a training set (80%) and a test set (20%). Then, 80% of the training data were used to train the model, whereas 20% were used to evaluate the trained models and select the best one. The test set was used to evaluate the performance of the final model and assess the generalized ability of the model on unseen data. The number of proteins in the four-grain protein datasets is listed in Table 1.

**Table 1.** The number of samples in each of the four-grain protein sub-ontology datasets.

Datasets	Sub-Ontology	Training Samples	Validation Samples	Test Samples	Total	Number of Classes	GO Filtering Threshold
Japonica	MF <sup>1</sup>	2157	540	675	3372	121	50
	BP <sup>2</sup>	2067	517	646	3230	254	
	CC <sup>3</sup>	2085	522	652	3259	59	
Indica	MF	260	65	82	407	52	10
	BP	68	17	22	107	28	
	CC	192	48	61	301	28	
Maize	MF	254	64	80	398	48	10
	BP	76	20	25	121	31	
	CC	152	38	48	238	27	
Wheat	MF	131	33	42	206	56	5
	BP	37	10	12	59	29	
	CC	61	16	20	97	27	

<sup>1</sup> MF represents the molecular function. <sup>2</sup> BP stands for biological process. <sup>3</sup> CC represents cellular components. The GO filtering threshold indicated that GO terms with annotated protein numbers below this value were not considered.

## 2.2. Data Representation

### 2.2.1. Protein Sequence Coding

In this study, protein sequence data were used as the basis. The selected protein sequence consisted of 20 amino acids (aa). The amino acid sequence was encoded using the  $n$ -gram ( $n = 1$ ) technique, which is a statistical language model used in natural language processing. A unigram of amino acid (AA) was generated from the protein sequence, and a lookup table was constructed with 20 amino acids corresponding to 20 specific real numbers, ranging from 1 to 20. Based on the lookup table, vocabulary was constructed for each protein sequence. This means that each amino acid in the protein sequence was mapped to a number (the index of the amino acid in the lookup table). Based on the concept of dense embedding and guided by the experimental analysis of Zuallaert [23], each numerical code in the vocabulary was mapped to a 10-dimensional dense vector through the embedding layer. The embedding layer adjusts the parameters during training to obtain the best dense vector representation. To adapt to the input data format of the deep learning model, each sequence was encoded as a vector of length 1002. If the sequence length is less than 1002, then the vector is padded with zeros. Each protein sequence is represented as a sequence feature matrix with dimensions  $1002 \times 10$ . The protein sequence was abstractly modeled using this encoding strategy to capture the local features and sequence dependencies.

### 2.2.2. Secondary Structural Features

The addition of a secondary structural feature based on sequence data can enhance the information content of input data to a certain extent. PDB files containing protein 3D structure data were converted into DSSP files containing secondary structure information using the web server of the DSSP website [24,25]. The eight-category secondary structure states [ $3_{10}$ -helix (G),  $\alpha$ -helix (H),  $\pi$ -helix (I),  $\beta$ -strand (E),  $\beta$ -bridge (B),  $\beta$ -turn (T), bend (S), and loop or irregular (L)] of the proteins were analyzed using the DSSP file. Owing to the absence of a secondary structure status for some amino acids in the DSSP file, a one-hot vector of dimension nine was used to encode the secondary structure status. The first eight dimensions correspond to eight types of secondary structures, and the last dimension indicates the absence of a secondary structure status. In the one-hot vector, the secondary structure position was set to 1, and the others were set to 0.

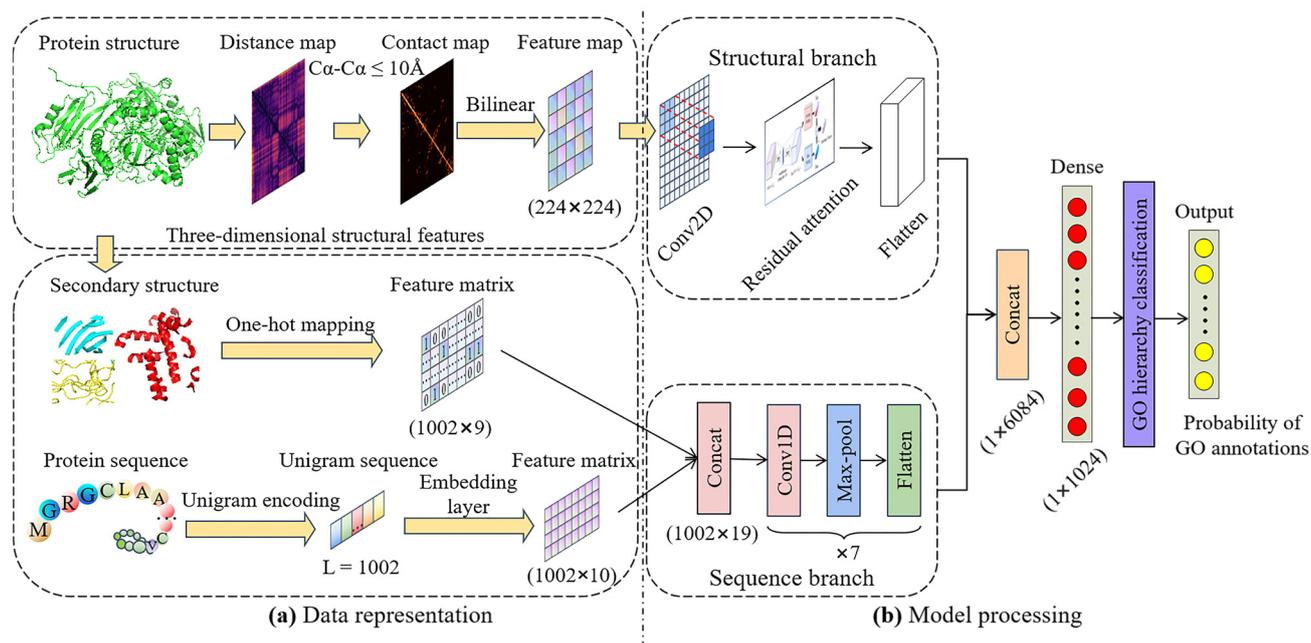
Similarly, the secondary structure information corresponding to the amino acids in the protein sequence was also constructed into a  $1002 \times 9$  feature matrix, which was spliced with the sequence feature matrix during the initial stage of the model. Therefore, using the two types of features, sequence encoding and secondary structure, a 19-dimensional feature vector was obtained for each amino acid in the protein sequence. Finally, each protein amino acid sequence was represented by a tensor size of  $1002 \times 19$  (combined sequence and secondary structure features).

### 2.2.3. Structural Contact Map Features

The  $C\alpha$  atom coordinate information for each amino acid residue was first collected from the processed DSSP file to characterize the 3D structure of the protein. Next, the Euclidean distances between each pair of residues were calculated based on the  $C\alpha$  atom coordinates, and a protein distance map was constructed to quantify spatial relationships within the protein structure. Figure 1 illustrates the spatial relationships among  $C\alpha$  atoms in protein structures. The contact map was constructed by setting the distance threshold to 10 Å; that is, if the distance between the  $C\alpha$  atoms of two residues was less than or equal to 10 Å, the two residues were considered to be in contact. The critical interactions between residues in the protein structure were highlighted by setting a reasonable contact



by performing convolutional operations on an input layer. An overview of the MMSNet approach is shown in Figure 3.

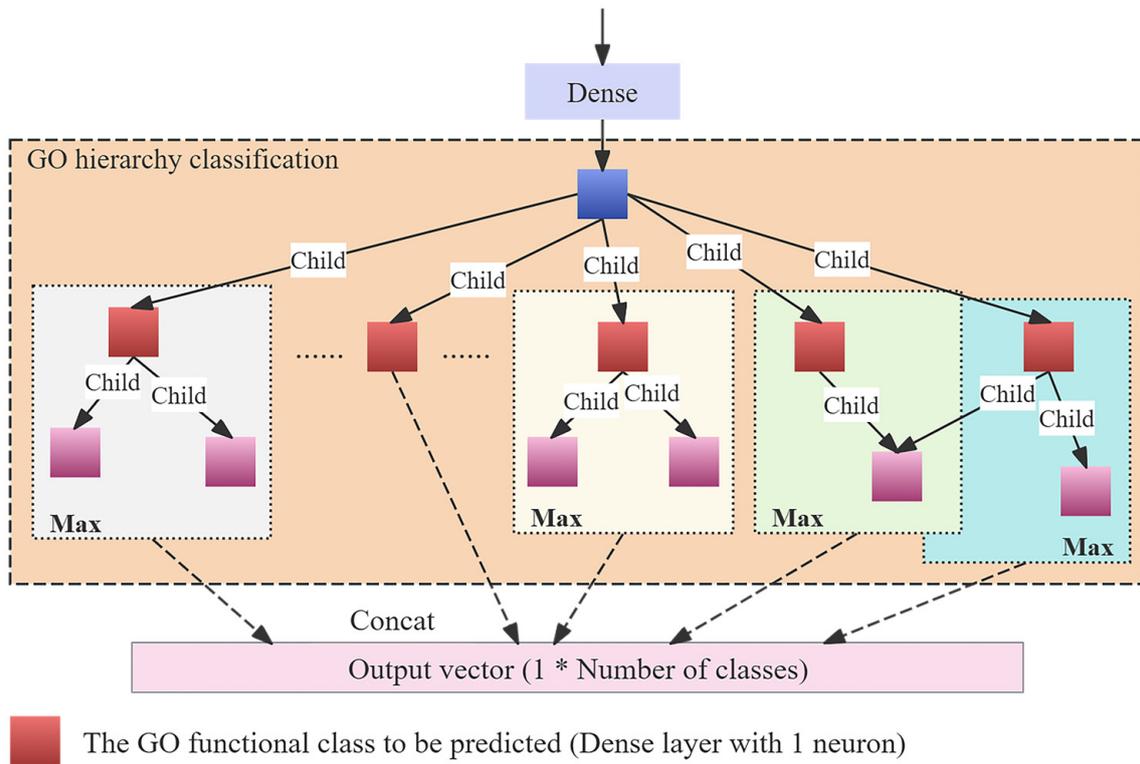


**Figure 3.** Overview of the flow of the MMSNet method. The process consists of two main parts: data processing (a) and model processing (b).

We applied a one-dimensional (1D) convolution to the sequence data. The 1DCNN can effectively capture the local patterns and features of long-sequence data. However, feature information at different levels and ranges is challenging to capture with a single-scale convolutional kernel, potentially leading to the omission of important long-distance dependent features and the introduction of unnecessary local feature redundancy. Therefore, a multiscale one-dimensional convolutional neural network was employed, consisting of seven groups of one-dimensional convolution and max pooling layers, with kernel sizes of 8, 16, 32, 64, 128, 256, and 512 for each group of convolution layers. Different convolutional kernels can capture the features of different dimensions. The features extracted by the multiscale convolutional kernel are combined to create a richer and more comprehensive representation of the sequence data.

The spatial characteristics of a protein are represented in a two-dimensional image by gradually abstracting protein structure data. We used two-dimensional convolution (Conv2D) to process the image data. Two-dimensional convolution is a technique specifically designed for processing image data that effectively captures the spatial characteristics of an image by sliding a convolution kernel across a two-dimensional plane. In protein image data processing, the capture of local patterns and features of protein structures through 2D convolution supports further analyses. Pooling layers (such as max pooling and average pooling) are traditionally used by convolutional neural networks to down-sample feature maps. However, many redundant features are generated by these pooling layers, thereby increasing the computational complexity and resulting in a decrease in overall model performance. To solve this problem, we replaced the ordinary pooling layer with a residual attention mechanism. The essential features in the feature map can be adaptively weighed by the residual attention mechanism, leading to a significant reduction in the number of redundant features and enhancement in the overall performance of the model. Through a multimodal processing method, the model can make full use of the

complementary information in different modal data. The GO hierarchical classification layer is shown in Figure 4.



**Figure 4.** GO hierarchical classification layer description. Max indicates that the category of the function to be predicted uses the maximum value in its own nodes and sub-nodes as the predicted value.

### 2.3.2. Residual Attention Mechanism

This study builds upon Zhu’s research on residual attention [27], making necessary adjustments and improvements to the original residual attention method based on the characteristics of the data to be processed with the aim of enhancing the ability of the model to extract structural features.

The core idea of the residual attention mechanism used in this study is to combine the convolution operation with the attention mechanism to enhance the representation of essential features, including convolution, global average pooling, global max pooling, and the calculation of attention scores. Let the input feature map be  $x$  with shape  $(H, W, C_{in})$  and let  $y_{raw}$  be the feature map generated by the convolutional layer with shape  $(H, W, C_{out})$ . Its expression is as follows:

$$y_{raw} = \text{Conv2D}_{1 \times 1}(x) \tag{1}$$

Conv2D is a  $1 \times 1$  convolution operation that maps the input feature map  $x$  to the output feature map  $y_{raw}$  with global average pooling and global maximum pooling performed on each channel of  $y_{raw}$ . For each channel  $k$ :

$$y_{avg,k} = \frac{1}{H \times W} \sum_{i=1}^H \sum_{j=1}^W y_{raw}(i, j, k) \tag{2}$$

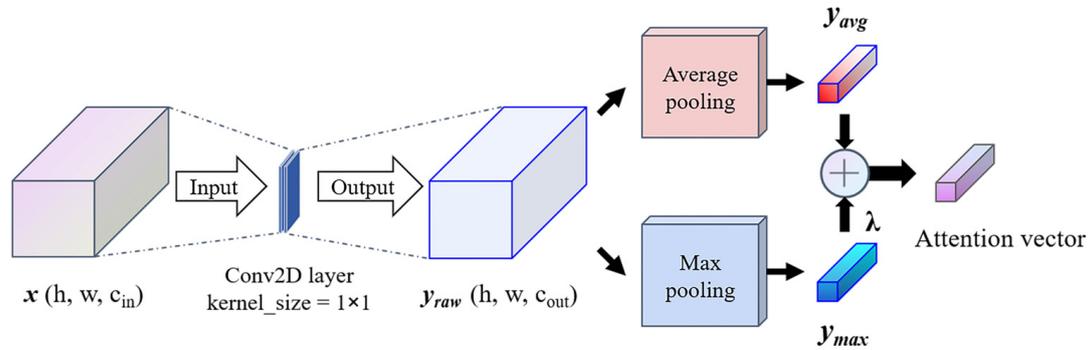
$$y_{max,k} = \max_{i=1, \dots, H; j=1, \dots, W} y_{raw}(i, j, k) \tag{3}$$

Thus, we obtain two sets of vectors,  $y_{avg} = [y_{avg,1}, y_{avg,2}, \dots, y_{avg,C_{out}}]$  and  $y_{max} = [y_{max,1}, y_{max,2}, \dots, y_{max,C_{out}}]$  of the shape  $(C_{out})$ . Finally, the final attention value

score was computed according to the following equation, with the output having a shape of ( $C_{out}$ ):

$$\text{score} = y_{avg} + \lambda \cdot y_{max} \quad (4)$$

In this formula,  $\lambda$  is defined as a weighting parameter used to adjust the relative importance of the global mean and maximum values. In our model,  $\lambda$  was set to 0.5. Figure 5 illustrates the computational relationship of the residual attention mechanism.



**Figure 5.** Computational relationship of residual attention mechanism.

### 2.3.3. Model Configuration and Training Details

The Tensorflow/Keras deep-learning library was used to implement the deep-learning framework. We used an NVIDIA A40 GPU with 48 GB of RAM to train the model to handle large amounts of data and to speed up the training process. During training, the Adam optimizer was used to optimize the model parameters, and the learning rate was set to 0.0003. A batch size of 32 was used in all experiments, and 200 epochs were set as the default. The binary cross-entropy loss function was used to measure the difference between the predicted results and true labels. The loss function is defined as follows:

$$\text{Loss} = -\frac{1}{N} \sum_{i=1}^N [y_i \log(p_i) + (1-y_i) \log(1-p_i)] \quad (5)$$

where  $N$  is the total number of samples,  $y_i$  is the true label, and  $p_i$  is the predicted label. At the end of each training round, we monitored the performance of the model on the validation set and saved the weights of the best model. To avoid overfitting, we used the early stopping criterion with patience = 10; that is, if the validation loss did not decrease for ten consecutive epochs, the training was terminated.

### 2.4. Comparison with Existing Methods

In this study, the constructed model was compared with two existing deep-learning models using three sub-ontology datasets (MF, BP, and CC) of the Japonica proteins. The DeepGOPlus prediction system combines DeepGOCNN model prediction with prediction based on sequence similarity. This study focused on a comparative analysis of the DeepGOCNN model using only the protein sequence data. DeepGOCNN consists of a series of CNN layers equipped with filters of sizes 8–128, with 512 filters per layer. DeepFRI learns protein feature representations by propagating sequence-level features of structurally neighboring residues through GCNs and uses structural data from the PDB and SWISS-MODEL databases to establish a link between protein structure and function prediction.

DeepGOCNN relies on sequence data and does not utilize structural information, which may limit its accuracy. DeepFRI integrates structural data but relies on traditional databases, especially in the case of diverse and complex proteins. In this study, both protein sequence data and high-quality structural data predicted by AlphaFold2 were used.

### 2.5. Evaluation Metrics

In this study,  $F_{\max}$ , average precision (AvgPr), average recall (AvgRc), area under the precision-recall curve (AUPR), and the Matthews correlation coefficient (MCC) were used as model evaluation metrics to measure the performance of our model and other methods for grain protein function prediction.

$F_{\max}$  is a protein-centered evaluation metric utilized in CAFA and represents the maximum value of the geometric mean of the average precision and average recall for all proteins in the test set at each threshold [28]. First, the average precision and recall were calculated using the following formula:

$$pr_i(t) = \frac{\sum_f I(f \in P_i(t) \wedge f \in T_i)}{\sum_f I(f \in P_i(t))} \quad (6)$$

$$rc_i(t) = \frac{\sum_f I(f \in P_i(t) \wedge f \in T_i)}{\sum_f I(f \in T_i)} \quad (7)$$

$$\text{AvgPr}(t) = \frac{1}{m(t)} \sum_{i=1}^{m(t)} pr_i(t) \quad (8)$$

$$\text{AvgRc}(t) = \frac{1}{n} \sum_{i=1}^n rc_i(t) \quad (9)$$

where  $f$  is the GO class,  $P_i(t)$  is the predicted annotation set of protein  $i$  at threshold  $t$ ,  $T_i$  is the true annotation set for protein  $i$ ,  $m(t)$  is the number of proteins with at least one predicted annotation at threshold  $t$ ,  $n$  is the total number of proteins.  $I$  is an identity function that returns 1 if the condition is true and 0 otherwise. We then computed  $F_{\max}$  (prediction threshold  $t \in [0, 1]$  and a step size of 0.01). If the predicted probability of a class is higher than or equal to  $t$ , it is considered as a predicted class:

$$F_{\max} = \max_t \left\{ \frac{2 \times \text{AvgPr}(t) \times \text{AvgRc}(t)}{\text{AvgPr}(t) + \text{AvgRc}(t)} \right\} \quad (10)$$

TP, TN, FP, and FN denote the true positives, true negatives, false positives, and false negatives, respectively. AUPR and MCC are two reasonable and important evaluation metrics for highly unbalanced learning problems, providing a comprehensive measurement standard [29]. The MCC calculation formula is as follows:

$$\text{MCC} = \frac{\text{TP} \times \text{TN} - \text{FP} \times \text{FN}}{\sqrt{(\text{TP} + \text{FP})(\text{TP} + \text{FN})(\text{TN} + \text{FP})(\text{TN} + \text{FN})}} \quad (11)$$

The higher the value of all the evaluation metrics, the better the algorithm performance. Additionally, a trade-off between precision and recall is typically observed, where increasing precision may reduce recall and vice versa.

## 3. Results and Discussion

### 3.1. Comparative Analysis of Algorithm Performance Combining Sequence and Structural Features

This section provides an in-depth discussion of the performance of the three algorithms based on different feature combinations in grain protein function prediction, focusing on the impact of structural features on the algorithm performance. Four types of grains were selected as research subjects, and the performances of different algorithms on the three types of protein function ontologies were analyzed. The results from the five experiments were averaged for each set of experimental data. The experimental results are presented in Table 2.

**Table 2.** Function prediction results of four grain proteins.

Grain	Sub-Ontology	Algorithm	Fmax	AvgPr	AvgRe	AUPR	MCC
Japonica	MF	MCNN-1D (S) <sup>1</sup>	0.650	0.710	0.604	0.741	0.656
		MCNN-1D (S+SS)	0.730	0.772	0.692	0.824	0.734
		MMSNet (S+SS+DS) <sup>2</sup>	<b>0.764</b>	<b>0.818</b>	<b>0.720</b>	<b>0.849</b>	<b>0.762</b>
	BP	MCNN-1D (S)	0.544	0.638	0.476	0.530	0.486
		MCNN-1D (S+SS)	0.645	0.710	0.588	0.648	0.593
		MMSNet (S+SS+DS)	<b>0.654</b>	<b>0.728</b>	<b>0.592</b>	<b>0.661</b>	<b>0.603</b>
	CC	MCNN-1D (S)	0.754	0.796	0.714	0.778	0.691
		MCNN-1D (S+SS)	0.789	0.790	<b>0.788</b>	0.814	0.724
		MMSNet (S+SS+DS)	<b>0.800</b>	<b>0.824</b>	0.776	<b>0.819</b>	<b>0.739</b>
Indica	MF	MCNN-1D (S)	0.566	0.720	0.468	0.477	0.419
		MCNN-1D (S+SS)	0.654	<b>0.808</b>	0.552	0.564	0.500
		MMSNet (S+SS+DS)	<b>0.668</b>	0.776	<b>0.594</b>	<b>0.580</b>	<b>0.522</b>
	BP	MCNN-1D (S)	0.558	<b>0.804</b>	0.430	0.627	0.457
		MCNN-1D (S+SS)	0.588	0.670	<b>0.596</b>	0.641	0.392
		MMSNet (S+SS+DS)	<b>0.604</b>	0.748	0.514	<b>0.655</b>	<b>0.482</b>
	CC	MCNN-1D (S)	0.699	0.746	0.658	0.677	0.569
		MCNN-1D (S+SS)	0.737	0.786	0.696	<b>0.731</b>	0.649
		MMSNet (S+SS+DS)	<b>0.752</b>	<b>0.794</b>	<b>0.716</b>	0.728	<b>0.667</b>
Maize	MF	MCNN-1D (S)	0.556	<b>0.690</b>	0.472	0.498	0.423
		MCNN-1D (S+SS)	0.563	0.610	<b>0.528</b>	0.509	0.439
		MMSNet (S+SS+DS)	<b>0.574</b>	0.642	0.526	<b>0.513</b>	<b>0.442</b>
	BP	MCNN-1D (S)	0.716	0.872	0.618	0.766	0.621
		MCNN-1D (S+SS)	0.784	0.918	<b>0.688</b>	0.817	0.711
		MMSNet (S+SS+DS)	<b>0.798</b>	<b>0.970</b>	0.678	<b>0.831</b>	<b>0.718</b>
	CC	MCNN-1D (S)	0.638	0.740	0.560	0.480	0.438
		MCNN-1D (S+SS)	0.714	<b>0.896</b>	0.596	0.589	0.512
		MMSNet (S+SS+DS)	<b>0.719</b>	0.870	<b>0.614</b>	<b>0.597</b>	<b>0.516</b>
Wheat	MF	MCNN-1D (S)	0.620	0.782	0.516	0.464	<b>0.402</b>
		MCNN-1D (S+SS)	0.651	0.814	0.544	0.461	0.371
		MMSNet (S+SS+DS)	<b>0.664</b>	<b>0.824</b>	<b>0.558</b>	<b>0.480</b>	0.390
	BP	MCNN-1D (S)	0.635	<b>0.720</b>	0.570	0.634	0.522
		MCNN-1D (S+SS)	0.705	0.676	0.760	0.669	0.513
		MMSNet (S+SS+DS)	<b>0.724</b>	0.680	<b>0.782</b>	<b>0.679</b>	<b>0.530</b>
	CC	MCNN-1D (S)	0.449	0.594	0.372	0.335	<b>0.314</b>
		MCNN-1D (S+SS)	0.531	0.612	0.476	0.351	0.258
		MMSNet (S+SS+DS)	<b>0.547</b>	<b>0.624</b>	<b>0.510</b>	<b>0.368</b>	0.272

S represents the protein sequence feature, SS represents the secondary structure feature, and DS represents the structural contact map feature. <sup>1</sup> MCNN-1D uses only a one-dimensional convolutional neural network, whereas <sup>2</sup> MMSNet is a model that combines sequence and structural branches. The optimal values of the experimental results are shown in bold font.

The MCNN-1D(S) algorithm is based only on sequence data. Although capturing the key features of the sequence and achieving a certain degree of prediction accuracy is possible in some cases, the overall performance is limited. On three sub-ontology datasets of four-grain proteins, the MCNN-1D(S+SS) algorithm generally achieved better performance than the MCNN-1D(S) algorithm based solely on sequence information across the five evaluation metrics, demonstrating enhanced prediction accuracy owing to secondary structure information. The MCNN-1D(S+SS) algorithm was only slightly inferior to the MCNN-1D(S) algorithm in terms of the evaluation metrics for a small number of datasets. In the CC dataset of japonica proteins, the AvgPr value of MCNN-1D(S+SS) was 0.006 lower than that of MCNN-1D(S). In the BP dataset of the indica protein, the MCNN-1D(S+SS) algorithm had a lower AvgPr value of 0.134 and MCC value of 0.065 than the MCNN-1D(S) algorithm. In the MF dataset of maize protein, the AvgPr value of MCNN-1D(S+SS) was 0.08 lower than that of MCNN-1D(S). For wheat protein, the AUPR and MCC values of MCNN-1D(S+SS) in the MF dataset were 0.003 and 0.031 lower than those of MCNN-1D(S), respectively, the AvgPr and MCC values of MCNN-1D(S+SS) in the BP dataset were

0.044 and 0.009 lower than those of MCNN-1D(S), respectively, and the MCC value of MCNN-1D(S+SS) was 0.056 lower than that of MCNN-1D(S) in the CC dataset. Overall, the effectiveness of the MCNN-1D(S+SS) algorithm was verified, demonstrating that secondary structure information is effective in predicting grain protein function.

The MMSNet algorithm adds structural branch and contact map features based on MCNN-1D(S+SS). The experimental results showed that the prediction performance of the model was the best when combined with structural contact map features. For the three ontology datasets of the four grains, the MMSNet algorithm outperformed the other two algorithms in terms of at least three evaluation metrics. In particular, MMSNet performed better on the japonica protein dataset, with an AvgRe value only 0.012 lower than that of MCNN-1D(S+SS) on the CC dataset and an AvgPr value higher than those of the other two algorithms. This may be due to the larger sample size of the Japonica protein dataset, which provided richer training information. The advantage of its data volume compared to the other three grains is that it may reduce the risk of overfitting, thereby allowing the model to generalize better to unseen data. In terms of the critical metric  $F_{\max}$ , the MMSNet algorithm outperformed the other two algorithms in all three sub-datasets of the four grains.

This demonstrated the critical role of algorithm innovation in improving the performance of grain protein function prediction models. At the same time, the excellent performance of the MMSNet algorithm powerfully demonstrates the potential of combining multidimensional biological data with advanced computing technology in the field of grain protein function prediction. It should be noted that the wheat protein dataset, with the smallest sample size, resulted in more unsatisfactory outcomes for the model across the three sub-datasets. This phenomenon reflects the limitations of data scarcity in model training.

### *3.2. Performance Comparison Between Residual Attention Mechanism and Traditional Pooling Layer*

In the structural branch, a severe performance bottleneck was encountered. Reliance on ordinary pooling operations to process structural contact map data, which contain rich structural details, was found to be insufficient for effective extraction and retention of critical local features. Moreover, essential details are prone to be lost, which negatively affects the overall recognition ability and generalization performance of the model. As a result, more advanced and intelligent feature integration strategies have been explored.

A residual attention mechanism was introduced to ensure that the most critical information for the final prediction was retained, even when feature reduction was performed, thereby preventing performance degradation caused by information loss. The effects of residual attention in the model were compared with those of traditional pooling layers on the three sub-datasets of japonica protein. Each experiment was repeated five times, and the results are shown in Table 3. The experimental results show that residual attention outperformed the other pooling layers in terms of AUPR and MCC metrics for the three sub-ontologies. In terms of  $F_{\max}$  and AvgPr metrics, the mean effect of the residual attention outperformed that of the other pooling layers. Only for the  $F_{\max}$  and AvgPr metrics of the CC sub-ontology dataset, the standard deviation of the residual attention was 0.002 and 0.003 higher than the global average pooling, respectively. In the case of the AvgRe metric, the standard deviation of the residual attention in the MF sub-ontology dataset was 0.002 higher than that of the global maximum pooling, the standard deviation of the residual attention in the BP sub-ontology dataset was 0.007 higher than that of the global average pooling, and the mean effect of the residual attention in the CC sub-ontology dataset was 0.002 lower than that of the global average pooling. When the residual at-

tention mechanism was employed, a data error of less than 1% was observed in cases of performance degradation compared with other pooling layer configurations.

**Table 3.** Experimental results of the model component comparison.

Sub-Ontology	Algorithm	Fmax	AvgPr	AvgRe	AUPR	MCC
MF	MMSNet (Max Pooling)	0.660 ± 0.033	0.690 ± 0.045	0.632 ± 0.028	0.700 ± 0.055	0.619 ± 0.043
	MMSNet (Average Pooling)	0.618 ± 0.044	0.652 ± 0.057	0.584 ± 0.036	0.645 ± 0.069	0.575 ± 0.062
	MMSNet (GlobalMax)	0.732 ± 0.025	0.776 ± 0.041	0.696 ± <b>0.026</b>	0.831 ± 0.021	0.738 ± 0.025
	MMSNet (GlobalAverage)	0.754 ± 0.031	0.798 ± 0.046	0.710 ± 0.032	0.845 ± 0.024	0.754 ± 0.028
	MMSNet (ResidualAttention)	<b>0.764 ± 0.014</b>	<b>0.818 ± 0.025</b>	<b>0.720 ± 0.028</b>	<b>0.849 ± 0.015</b>	<b>0.762 ± 0.014</b>
BP	MMSNet (Max Pooling)	0.536 ± 0.014	0.600 ± 0.052	0.488 ± 0.016	0.485 ± 0.017	0.450 ± 0.015
	MMSNet (Average Pooling)	0.518 ± 0.015	0.566 ± 0.026	0.476 ± 0.021	0.456 ± 0.012	0.420 ± 0.013
	MMSNet (GlobalMax)	0.628 ± 0.019	0.696 ± 0.024	0.574 ± 0.016	0.634 ± 0.019	0.579 ± 0.025
	MMSNet (GlobalAverage)	0.638 ± 0.015	0.698 ± 0.031	0.584 ± <b>0.008</b>	0.643 ± 0.016	0.583 ± 0.021
	MMSNet (ResidualAttention)	<b>0.654 ± 0.011</b>	<b>0.728 ± 0.020</b>	<b>0.592 ± 0.015</b>	<b>0.661 ± 0.009</b>	<b>0.603 ± 0.009</b>
CC	MMSNet (Max Pooling)	0.746 ± 0.012	0.794 ± 0.021	0.708 ± 0.032	0.734 ± 0.018	0.668 ± 0.015
	MMSNet (Average Pooling)	0.732 ± 0.020	0.764 ± 0.029	0.700 ± 0.021	0.712 ± 0.030	0.650 ± 0.024
	MMSNet (GlobalMax)	0.790 ± 0.009	0.816 ± 0.022	0.764 ± 0.008	0.816 ± 0.008	0.729 ± 0.012
	MMSNet (GlobalAverage)	0.794 ± <b>0.005</b>	0.814 ± <b>0.014</b>	<b>0.778 ± 0.010</b>	0.818 ± 0.007	0.734 ± 0.008
	MMSNet (ResidualAttention)	<b>0.800 ± 0.007</b>	<b>0.824 ± 0.017</b>	0.776 ± <b>0.005</b>	<b>0.819 ± 0.005</b>	<b>0.739 ± 0.007</b>

The maximum mean and minimum standard deviations for each set of experimental results are in bold.

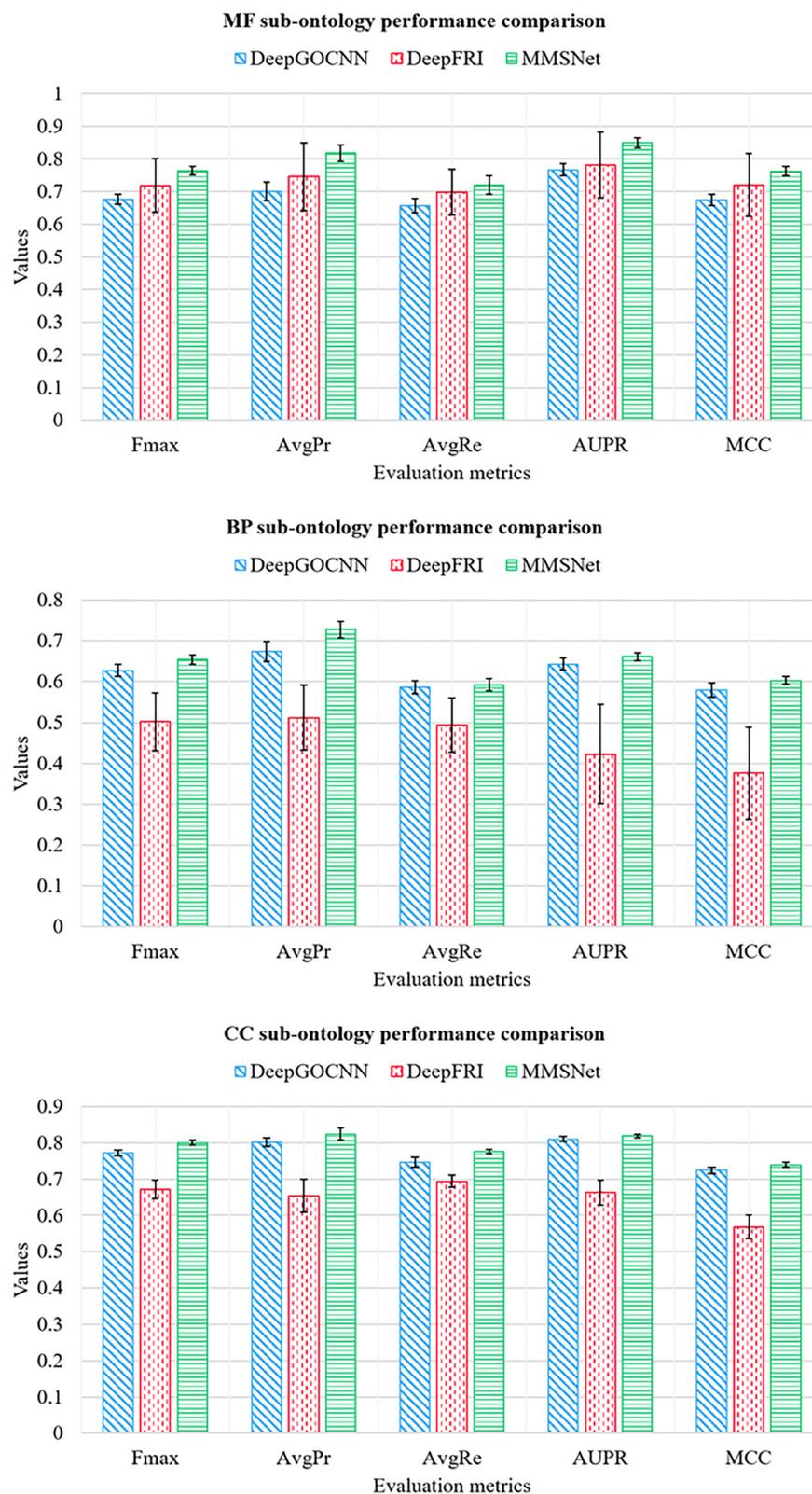
This indicates that, even in the most unfavorable situations, a performance level close to that of the optimal pooling strategy can still be maintained by the residual attention mechanism, further confirming its excellence in reducing prediction bias and improving model robustness. This finding emphasizes the importance of adopting flexible and intelligent feature-processing strategies for specific data characteristics when designing deep-learning architectures.

### 3.3. Comparative Analysis of Performance with Other Model Methods

This study provides a detailed performance comparison of three different model methods on japonica protein datasets. DeepGOCNN uses only the protein sequence data. Sequence and structural data were combined in DeepFRI using graph convolutional networks (GCN). In this study, DeepFRI and MMSNet used the same structural data. Each experiment was repeated five times, and Figure 6 shows the performance comparisons of MMSNet and the other two algorithms on the three sub-ontology datasets.

In the MF subontology dataset, DeepFRI outperformed DeepGOCNN for all the metrics. This may be because the three-dimensional structural information of proteins utilized by DeepFRI plays a crucial role in MF sub-ontology, as molecular functions are often closely related to specific structural domains or protein conformations. In the BP and CC sub-ontologies, DeepFRI performed worse than DeepGOCNN, which may be because the structural features extracted by the GCN contain more noisy data, which significantly reduces the performance of the model in predicting biological processes and cellular component functions.

Among the three sub-ontologies, MMSNet outperformed DeepFRI and DeepGOCNN in terms of the five evaluation metrics. Although both MMSNet and DeepFRI use structural information, they aim to fully utilize high-precision protein structures and multimodal integration techniques. Case studies of MMSNet have highlighted the importance of integrating multisource data to improve the accuracy and stability of predictions.



**Figure 6.** Performance comparison of MMSNet and the other two algorithms on the ontology dataset.

### 3.4. Analysis of Model Prediction Results

To further study and analyze the prediction effect of our model, the MMSNet algorithm was employed to predict the function of proteins that were not included in the training dataset. The predicted protein data were selected from the experimental prediction results of three sub-ontology datasets of proteins (MF, BP, and CC). The comparison results for the

protein functions selected in this section are shown in Table 4. We focused on analyzing the functions of some japonica proteins based on the latest research papers and model prediction results. We want to emphasize that we were not only limited to the study of model performance, but also paid more attention to the practical significance of the model.

**Table 4.** Example of protein function prediction results for four types of grains.

Grain	Sub-Ontology	Protein	Real Function	Predicted Function
Japonica	MF	P29250 (LOX2_ORYSJ)	GO:0046872	GO:0046872
			GO:0043169	GO:0043169
			GO:0043167	GO:0043167
	MF	P16081 (NIA1_ORYSJ)	GO:0005488	GO:0005488
			GO:0016491	GO:0016491
			GO:0003824	GO:0003824
	MF	P16081 (NIA1_ORYSJ)	GO:0003824	GO:0003824
			GO:0005488	GO:0005488
			GO:1901363	GO:1901363
	MF	Q0JKI9 (ARFB_ORYSJ)	GO:0097159	GO:0097159
			GO:007159	GO:007159
			GO:0046914	GO:0003676
MF	Q0JKI9 (ARFB_ORYSJ)	GO:0043168	GO:0003723	
		GO:0003677	GO:0016787	
		GO:0003676	GO:0003676	
MF	Q0JKI9 (ARFB_ORYSJ)	GO:0097159	GO:0097159	
		GO:1901363	GO:1901363	
		GO:0005488	GO:0005488	
BP	Q10RB4 (BGAL5_ORYSJ)	GO:0008152	GO:0008152	
		GO:0044238	GO:0044238	
		GO:0005975	GO:0005975	
BP	Q8W0A1 (BGAL2_ORYSJ)	GO:0005975	GO:0005975	
		GO:0044238	GO:0044238	
		GO:0008152	GO:0008152	
BP	Q8W0A1 (BGAL2_ORYSJ)	GO:0009058	GO:0009058	
		GO:0016051	GO:0016051	
		GO:0009507	GO:0009507	
CC	Q84YK8 (LOXC2_ORYSJ)	GO:0009536	GO:0009536	
		GO:0043231	GO:0043231	
		GO:0043227	GO:0043227	
CC	Q84YK8 (LOXC2_ORYSJ)	GO:0043229	GO:0043229	
		GO:0043226	GO:0043226	
		GO:0110165	GO:0110165	
CC	Q6YSJ5 (AGO16_ORYSJ)	GO:0110165	GO:0110165	
		GO:0005737	GO:0005737	
		GO:0110165	GO:0110165	
MF	A2Y9M4 (SSY1_ORYSI)	GO:0005886	GO:0005886	
		GO:0016020	GO:0016020	
		GO:0110165	GO:0110165	
MF	P0C461 (RR12_ORYSI)	GO:0003824	GO:0003824	
		GO:0016740	GO:0003824	
		GO:0003735	GO:0003735	
MF	P0C461 (RR12_ORYSI)	GO:0005198	GO:0005198	
		GO:0005488	GO:0005488	
		GO:0009987	GO:0009987	
BP	Q01IX6 (DAO_ORYSI)	GO:0008152	GO:0008152	
		GO:0044237	GO:0008152	
		GO:0044238	GO:0044238	
BP	A2YNH4 (6PGL2_ORYSI)	GO:0071704	GO:0071704	
		GO:0008152	GO:0008152	
		GO:0008152	GO:0008152	
CC	B8BK18 (MCM2_ORYSI)	GO:0032991	GO:0032991	
		GO:0110165	GO:0110165	
		GO:0005737	GO:0005737	
CC	B8ATT7 (VLN4_ORYSI)	GO:0110165	GO:0110165	
		GO:0005737	GO:0005737	
		GO:0110165	GO:0110165	

Table 4. Cont.

Grain	Sub-Ontology	Protein	Real Function	Predicted Function
Maize	MF	P29390 (FRI2_MAIZE)	GO:0003824 GO:0016491	GO:0003824
		P06677 (ZEA9_MAIZE)	GO:0045735	GO:0045735 GO:0003824
	BP	P33488 (ABP4_MAIZE)	GO:0009987	GO:0009987 GO:0044237 GO:0008152
		B6SU46 (AAMT2_MAIZE)	GO:0006950 GO:0050896	GO:0006950 GO:0050896
	CC	Q9LKX9 (RBR1_MAIZE)	GO:0032993 GO:0032991	GO:0032993 GO:0032991
		Q41764 (ADF3_MAIZE)	GO:0110165 GO:0005737	GO:0110165
Wheat	MF	Q8L803 (RK9_WHEAT)	GO:0003735 GO:0005198	GO:0003735 GO:0005198 GO:0003824
		Q5I7K3 (RS29_WHEAT)	GO:0003735 GO:0005198	GO:0003735 GO:0005198
	BP	B6DZC8 (1FEH3_WHEAT)	GO:0008152 GO:0044238 GO:0071704 GO:0005975	GO:0008152 GO:0044238 GO:0071704 GO:0071704
		O04706 (GAO1B_WHEAT)	GO:0009416 GO:0009314 GO:0009628 GO:0050896	GO:0009416 GO:0009314 GO:0009628 GO:0050896
	CC	Q41560 (HS16B_WHEAT)	GO:0005737 GO:0110165	GO:0005737 GO:0110165
		Q01481 (WIR1B_WHEAT)	GO:0110165 GO:0016020	GO:0110165

For some protein samples, all the preset labels were correctly predicted by the model, demonstrating excellent prediction accuracy. For example, in the MF sub-dataset, the P29250 (LOX2\_ORYSJ) protein was annotated as GO:0046872, GO:0043169, GO:0043167, GO:0005488, GO:0016491, and GO:0003824. The functions predicted by the model were consistent with actual annotations. The P29250 protein was also annotated with the GO:0046872 function in the UniProt database, as inferred from electronic annotation. Furthermore, Guo indicated that the P29250 protein might play a key role in biological processes related to metal ion binding (GO:0046872) [30]. The Q0JKI9 (ARFB\_ORYSJ) proteins in the test set were annotated using GO:0003677, GO:0003676, GO:0097159, GO:1901363, and GO:0005488. The function of the Q0JKI9 protein was accurately predicted by the model and aligned with its annotated function. Prathap inferred that the Q0JKI9 protein has a binding function (GO:0005488) by analyzing the ARF2 gene corresponding to the Q0JKI9 protein [31]. Auxin response factors (ARFs) generally exhibit DNA-binding activity; therefore, it can be inferred that the Q0JKI9 protein may be involved in DNA binding (GO:0003677). The UniProt database also inferred that the Q0JKI9 protein had a GO:0003677 function through automatic annotation. The actual annotated function of the Q6YSJ5 (AGO16\_ORYSJ) protein in the CC sub-dataset was correctly predicted by the model. The above shows that, in some specific cases, the key features of the data can be accurately captured, and accurate judgments can be made by the model.

However, for other protein samples, the prediction results reflected the uncertainties and limitations of the functional predictions posed by the model. Models can occasionally predict functions that are not included in a label set. For example, the test protein P16081

(NIA1\_ORYSJ) in the MF ontology dataset was manually annotated with the GO:0003824, GO:0005488, GO:1901363, GO:0097159, GO:0046914, and GO:0043168 functions. The functions GO:0003824, GO:0005488, GO:1901363, and GO:0097159 were successfully predicted by the model, whereas the annotated functions GO:0046914 and GO:0043168 were not. Additionally, the model predicted the functions of GO:0003676, GO:0003723, and GO:0016787. We mainly focused on and analyzed the function of GO:0016787 (hydrolase activity). A recent study by Haider revealed transcriptome changes in rice roots in response to phosphorus starvation [32]. The study mentioned that Os08g0468100, which encodes the P16081 protein, was downregulated under phosphorus-deficient conditions and may be related to phosphoric ester hydrolase activity. In the molecular function sub-ontology hierarchy, phosphoric ester hydrolase activity was classified as a more specific functional category under hydrolase activity (GO:0016787), suggesting that the P16081 protein may be related to hydrolase activity (GO:0016787). This demonstrates the feasibility of the proposed model for predicting unknown functions.

After manual annotation, Q8W0A1 (BGAL2\_ORYSJ) in the BP dataset had the GO:0005975, GO:0044238, and GO:0008152 functions. In addition to the annotated functions, the two functions GO:0016051 and GO:0009058 were predicted unexpectedly, which may be due to the correlation between the additional predicted and annotated functions. The gene expression of japonica germplasms under heat stress conditions was analyzed by Li [33], who revealed that the expression of genes associated with the function of GO:0009058 was downregulated. This downregulation suggests that proteins involved in the GO:0009058 function may be present in japonica and that the GO:0009058 function is a subset of the GO:0008152 function. Therefore, in this study, a certain degree of possibility was suggested for predicting Q8W0A1 protein function.

The Q84YK8 (LOXC2\_ORYSJ) protein in the CC dataset is annotated with GO:0009507, GO:0009536, GO:0043231, GO:0043227, GO:0043229, GO:0043226, and GO:0110165. The model additionally discovered the GO:0016020 function during prediction. In the CC ontology hierarchy, the membranes (GO:0016020) were categorized as functional subsets of cellular anatomical entities (GO:0110165). Membranes are a more specific and focused category of cellular components than cellular anatomical entities. A recent study on the genetic analysis of rice seedlings reported that the LOC\_Os08g39850 gene, corresponding to the Q84YK8 protein, encodes a chloroplast precursor lipoyxygenase [34]. As lipoyxygenases can act on the phospholipids of cell membranes, we speculate that the Q84YK8 protein may be related to membrane function (GO:0016020). Although additional predicted labels are not explicitly included in the original label set, an intrinsic biological relationship exists between them and known labels. A method for further data analysis and hypothesis generation was developed by mining latent knowledge. These three proteins—P16081 from the MF dataset, Q8W0A1 from the BP dataset, and Q84YK8 from the CC dataset—represent instances where extra functions were predicted by the model. The additional predicted functions hold significant value and implications for advancing our understanding of protein functionalities.

However, the model failed to successfully predict the actual labels for some samples. For example, the Q10RB4 (BGAL5\_ORYSJ) protein in the BP ontology dataset was manually annotated with functions GO:0005975, GO:0044238, and GO:0008152. However, the model did not predict the function of GO:0005975, possibly because of the similarities between GO:0005975 and GO:0044238. Within the BP gene ontology hierarchy, the carbohydrate metabolism process (GO:0005975) was classified as the primary metabolic process (GO:0044238), and potential biases may exist in the model when predicting specific functional categories. Similarly, the Q7XP59 (GLR31\_ORYSJ) protein in the CC ontology dataset had a GO:0005886 function in the Swiss-Prot database, and GO:0016020 and GO:0110165

functions were manually annotated. However, the GO:0005886 function was not observed during the prediction process. Owing to the inherent hierarchical relationship between the two functions, more general or abstract functional categories are likely to be recognized by the model during predictions because they appear more frequently in the training data.

We also predicted and observed the protein functions in the test sets of indica, maize, and wheat. For the indica proteins A2YNH4 (6PGL2\_ORYSI) and B8ATT7 (VLN4\_ORYSI), the maize proteins B6SU46 (AAMT2\_MAIZE) and Q9LKX9 (RBR1\_MAIZE), and the wheat proteins Q5I7K3 (RS29\_WHEAT), O04706 (GAO1B\_WHEAT), and Q41560 (HS16B\_WHEAT), the model successfully and completely predicted their functions. For the P0C461 (RR12\_ORYSI) and B8BK18 (MCM2\_ORYSI) proteins of indica, the P06677 (ZEA9\_MAIZE) and P33488 (ABP4\_MAIZE) proteins of maize, and the Q8L803 (RK9\_WHEAT) protein of wheat, the model predicted functions that had not been previously annotated. For the indica proteins A2Y9M4 (SSY1\_ORYSI) and Q01IX6 (DAO\_ORYSI), the maize proteins P29390 (FRI2\_MAIZE) and Q41764 (ADF3\_MAIZE), and the wheat proteins B6DZC8 (1FEH3\_WHEAT) and Q01481 (WIR1B\_WHEAT), the model did not fully predict their annotated functions. This prediction bias reflects the challenges faced by models when dealing with bioinformatic data with hierarchical structures and semantic associations. In summary, it was demonstrated that MMSNet generally performed well in predicting grain protein functions.

#### 4. Conclusions

In this study, a deep learning-based multimodal learning technique was developed and applied to the prediction of grain protein functions. A significant improvement in prediction performance was shown by the MMSNet method, which integrates 1D convolutional neural networks (1DCNN), 2D convolutional neural networks (2DCNN), and residual attention mechanisms, and incorporates the structural information (especially secondary structure) predicted by AlphaFold2 into the model. Four-grain protein datasets for japonica, indica, maize, and wheat were preprocessed to analyze and evaluate the prediction performance of the MMSNet model. The superiority of the MMSNet model in capturing complex patterns in protein sequences and structures was emphasized by experimental comparison with existing methods, such as DeepGOCNN and DeepFRI. In particular, the extraction of relevant features was enhanced by the residual attention mechanism through a flexible and intelligent feature integration strategy. It has been experimentally proven that the model performance was significantly improved by this mechanism, which is not achievable with traditional pooling layers. This innovation highlights the potential of computer vision-related technologies in the field of protein function prediction has been highlighted by this innovation.

Prediction of grain protein function can provide fundamental insights into food science research. For instance, predictions of enzymatic activity or molecular interactions might inspire targeted research into functional properties, such as texture enhancement or nutritional optimization, in specific food applications. This will help bridge the gap between biological understanding and practical applications in food science. The prediction function can be experimentally verified through biological methods to validate the practical applicability of our model and uncover new insights into protein function prediction. Future research should focus on exploring new methods for protein structure feature extraction, analyzing structural data from a more detailed perspective, and improving the robustness and accuracy of grain protein function prediction. In summary, new possibilities in the field of protein function prediction were identified using our multimodal, multiscale approach.

**Author Contributions:** Conceptualization, J.L.; methodology, J.L. and X.Z.; software, X.Z.; validation, X.Z.; formal analysis, J.L., K.H. and X.G.; investigation, X.Z.; resources, J.L., K.H. and X.G.; data curation, X.Z. and Y.W.; writing—original draft preparation, X.Z.; writing—review and editing, J.L., K.H. and Y.W.; visualization, X.Z.; supervision, J.L., K.H. and X.G.; project administration, J.L.; funding acquisition, J.L. and X.G. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the Program of Shanghai Municipal Education Commission, grant number Z-2024-312-099, under the project “Artificial Intelligence Promoting Research Paradigm Reform and Empowering Disciplinary Advancement Program”.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The original data presented in the study are openly available in Zenodo at <https://doi.org/10.5281/zenodo.13932372> (accessed on 15 October 2024). The analysis code for research methods is openly available on GitHub at [https://github.com/Azxp123456/Code\\_Data](https://github.com/Azxp123456/Code_Data) (accessed on 15 October 2024).

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Oluwole, O.; Ibidapo, O.; Arowosola, T.; Raji, F.; Zandonadi, R.P.; Alasqah, I.; Raposo, A. Sustainable transformation agenda for enhanced global food and nutrition security: A narrative review. *Front. Nutr.* **2023**, *10*, 1226538. [[CrossRef](#)]
2. Langyan, S.; Khan, F.N.; Kumar, A. Advancement in nutritional value, processing methods, and potential applications of pseudocereals in dietary food: A review. *Food Bioprocess Technol.* **2023**, *17*, 571–590. [[CrossRef](#)]
3. Kajzer, M.; Diowksz, A. The clean label concept: Novel approaches in gluten-free breadmaking. *Appl. Sci.* **2021**, *11*, 6129. [[CrossRef](#)]
4. Poutanen, K.S.; Karlund, A.O.; Gomez-Gallego, C.; Johansson, D.P.; Scheers, N.M.; Marklinder, I.M.; Landberg, R. Grains—A major source of sustainable protein for health. *Nutr. Rev.* **2022**, *80*, 1648–1663. [[CrossRef](#)] [[PubMed](#)]
5. Altschul, S.F.; Madden, T.L.; Schäffer, A.A.; Zhang, J.; Zhang, Z.; Miller, W.; Lipman, D.J. Gapped blast and psi-blast: A new generation of protein database search programs. *Nucleic Acids Res.* **1997**, *25*, 3389–3402. [[CrossRef](#)]
6. Liu, J.; Tang, X.; Cui, S.; Guan, X. Predicting the function of rice proteins through multi-instance multi-label learning based on multiple features fusion. *Brief. Bioinf* **2022**, *23*, bbac095. [[CrossRef](#)]
7. Yousef, M.; Jung, S.; Showe, L.C.; Showe, M.K. Learning from positive examples when the negative class is undetermined—microna gene identification. *Algorithms Mol. Biol.* **2008**, *3*, 2. [[CrossRef](#)] [[PubMed](#)]
8. Nam, J.-W.; Shin, K.-R.; Han, J.; Lee, Y.; Kim, V.N.; Zhang, B.-T. Human microna prediction through a probabilistic co-learning model of sequence and structure. *Nucleic Acids Res.* **2005**, *33*, 3570–3581. [[CrossRef](#)] [[PubMed](#)]
9. Chen, X.-W.; Liu, M. Prediction of protein–protein interactions using random decision forest framework. *Bioinformatics* **2005**, *21*, 4394–4400. [[CrossRef](#)] [[PubMed](#)]
10. Kulmanov, M.; Khan, M.A.; Hoehndorf, R. Deepgo: Predicting protein functions from sequence and interactions using a deep ontology-aware classifier. *Bioinformatics* **2018**, *34*, 660–668. [[CrossRef](#)]
11. Kulmanov, M.; Hoehndorf, R. Deepgoplus: Improved protein function prediction from sequence. *Bioinformatics* **2020**, *36*, 422–429. [[CrossRef](#)] [[PubMed](#)]
12. Buchfink, B.; Xie, C.; Huson, D.H. Fast and sensitive protein alignment using diamond. *Nat. Methods* **2015**, *12*, 59–60. [[CrossRef](#)] [[PubMed](#)]
13. Jisna, V.A.; Jayaraj, P.B. Protein structure prediction: Conventional and deep learning perspectives. *Protein J.* **2021**, *40*, 522–544. [[CrossRef](#)] [[PubMed](#)]
14. Sercinoğlu, O.; Ozbek, P. Sequence-structure-function relationships in class i mhc: A local frustration perspective. *PLoS ONE* **2020**, *15*, e0232849. [[CrossRef](#)]
15. Gligorijević, V.; Renfrew, P.D.; Kosciolk, T.; Leman, J.K.; Berenberg, D.; Vatanen, T.; Bonneau, R. Structure-based protein function prediction using graph convolutional networks. *Nat. Commun.* **2021**, *12*, 3168. [[CrossRef](#)]
16. Liu, J.; Tang, X.; Guan, X. Grain protein function prediction based on self-attention mechanism and bidirectional lstm. *Brief. Bioinf* **2023**, *24*, bbac493. [[CrossRef](#)] [[PubMed](#)]
17. Jumper, J.; Evans, R.; Pritzel, A.; Green, T.; Figurnov, M.; Ronneberger, O.; Hassabis, D. Highly accurate protein structure prediction with alphafold. *Nature* **2021**, *596*, 583–589. [[CrossRef](#)] [[PubMed](#)]

18. Boadu, F.; Cao, H.; Cheng, J. Combining protein sequences and structures with transformers and equivariant graph neural networks to predict protein function. *Bioinformatics* **2023**, *39*, i318–i325. [[CrossRef](#)]
19. Jiao, P.; Wang, B.; Wang, X.; Liu, B.; Wang, Y.; Li, J. Struct2go: Protein function prediction based on graph pooling algorithm and alphafold2 structure information. *Bioinformatics* **2023**, *39*, btad637. [[CrossRef](#)] [[PubMed](#)]
20. Boutet, E.; Lieberherr, D.; Tognolli, M.; Schneider, M.; Bansal, P.; Bridge, A.J.; Xenarios, I. Uniprotkb/swiss-prot, the manually annotated section of the uniprot knowledgebase: How to use the entry view. *Methods Mol. Biol.* **2016**, *1374*, 23–54. [[PubMed](#)]
21. Varadi, M.; Anyango, S.; Deshpande, M.; Nair, S.; Natassia, C.; Yordanova, G.; Velankar, S. Alphafold protein structure database: Massively expanding the structural coverage of protein-sequence space with high-accuracy models. *Nucleic Acids Res.* **2022**, *50*, D439–D444. [[CrossRef](#)] [[PubMed](#)]
22. Ashburner, M.; Ball, C.A.; Blake, J.A.; Botstein, D.; Butler, H.; Cherry, J.M.; Sherlock, G. Gene ontology: Tool for the unification of biology. *Nat. Genet.* **2000**, *25*, 25–29. [[CrossRef](#)] [[PubMed](#)]
23. Zuallaert, J.; Pan, X.; Saeys, Y.; Wang, X.; Neve, W.D. Investigating the biological relevance in trained embedding representations of protein sequences. In Proceedings of the Workshop on Computational Biology at ICML2019, Long Beach, CA, USA, 9–15 June 2019.
24. Joosten, R.P.; Beek, T.A.H.t.; Krieger, E.; Hekkelman, M.L.; Hooft, R.W.W.; Schneider, R.; Vriend, G. A series of pdb related databases for everyday needs. *Nucleic Acids Res.* **2011**, *39*, D411–D419. [[CrossRef](#)]
25. Kabsch, W.; Sander, C. Dictionary of protein secondary structure: Pattern recognition of hydrogen-bonded and geometrical features. *Biopolymers* **1983**, *22*, 2577–2637. [[CrossRef](#)] [[PubMed](#)]
26. Krizhevsky, B.A.; Sutskever, I.; Hinton, G.E. Imagenet classification with deep convolutional neural networks. *Commun. ACM* **2017**, *60*, 84–90. [[CrossRef](#)]
27. Zhu, K.; Wu, J. Residual attention: A simple but effective method for multi-label recognition. In Proceedings of the 2021 IEEE/CVF International Conference on Computer Vision (ICCV), Montreal, BC, Canada, 11–17 October 2021; pp. 184–193.
28. Radivojac, P.; Clark, W.T.; Oron, T.R.; Schnoes, A.M.; Wittkop, T.; Sokolov, A.; Friedberg, I. A large-scale evaluation of computational protein function prediction. *Nat. Methods* **2013**, *10*, 221–227. [[CrossRef](#)]
29. Zeng, M.; Zou, B.; Wei, F.; Liu, X.; Wang, L. Effective prediction of three common diseases by combining smote with totem links technique for imbalanced medical data. In Proceedings of the 2016 IEEE International Conference of Online Analysis and Computing Science (ICOACS), Chongqing, China, 28–29 May 2016; pp. 225–228.
30. Guo, N.; Tang, S.; Wang, J.; Hu, S.; Tang, S.; Wei, X.; Hu, P. Transcriptome and proteome analysis revealed that hormone and reactive oxygen species synergetically regulate dormancy of introgression line in rice (*Oryza sativa* L.). *Int. J. Mol. Sci.* **2023**, *24*, 6088. [[CrossRef](#)]
31. Prathap, V.; Kumar, S.; Tyagi, A. Comparative proteome analysis of phosphorus-responsive genotypes reveals the proteins differentially expressed under phosphorous starvation stress in rice. *Int. J. Biol. Macromol.* **2023**, *234*, 123760. [[CrossRef](#)] [[PubMed](#)]
32. Haider, I.; Yunmeng, Z.; White, F.; Li, C.; Incitti, R.; Alam, I.; Bouwmeester, H.J. Transcriptome analysis of the phosphate starvation response sheds light on strigolactone biosynthesis in rice. *Plant J.* **2023**, *114*, 355–370. [[CrossRef](#)]
33. Li, P.; Jiang, J.; Zhang, G.; Miao, S.; Lu, J.; Qian, Y.; Xu, J. Integrating gwas and transcriptomics to identify candidate genes conferring heat tolerance in rice. *Front. Plant Sci.* **2023**, *13*, 1102938. [[CrossRef](#)]
34. Wei, Y.; Li, X.; Li, D.; Su, X.; Huang, Y.; Li, Q.; Yang, X. Mapping and candidate gene analysis of the low-temperature-sensitive albino gene *osltsa8* in rice seedlings. *Curr. Issues Mol. Biol.* **2024**, *46*, 6508–6521. [[CrossRef](#)] [[PubMed](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.