Annotating protein functions via fusing multiple biological modalities

Wenjian Ma, Xiangpeng Bi, Huasen Jiang, Zhiqiang Wei & Shugang Zhang
Nature Communications Biology
doi.org/10.1038/s42003-024-07411-y

Gökçe Uludoğan
PhD Candidate
LifeLU Reading Group | 2 January 2025

Motivation

- Multimodal Integration for Protein Function Prediction
    - methods rely only on sequences or structures, ignoring:
        - Protein-protein interactions
        - Domain knowledge
        - Subcellular localization and other biological contexts.
- Component-based approach
    - Modality-specific feature subspaces & unavailability of unified and high-quality protein representations.
    - Difficulty in fusing diverse biological modalities effectively
- Supervised methods vs Self-supervised learning

MIF2GO: Multimodal Information Fusion for GO Annotation

Objective: To fuse diverse biological modalities using a unified framework

**Modalities:** domain, localization, pathway, interaction, homology, sequence
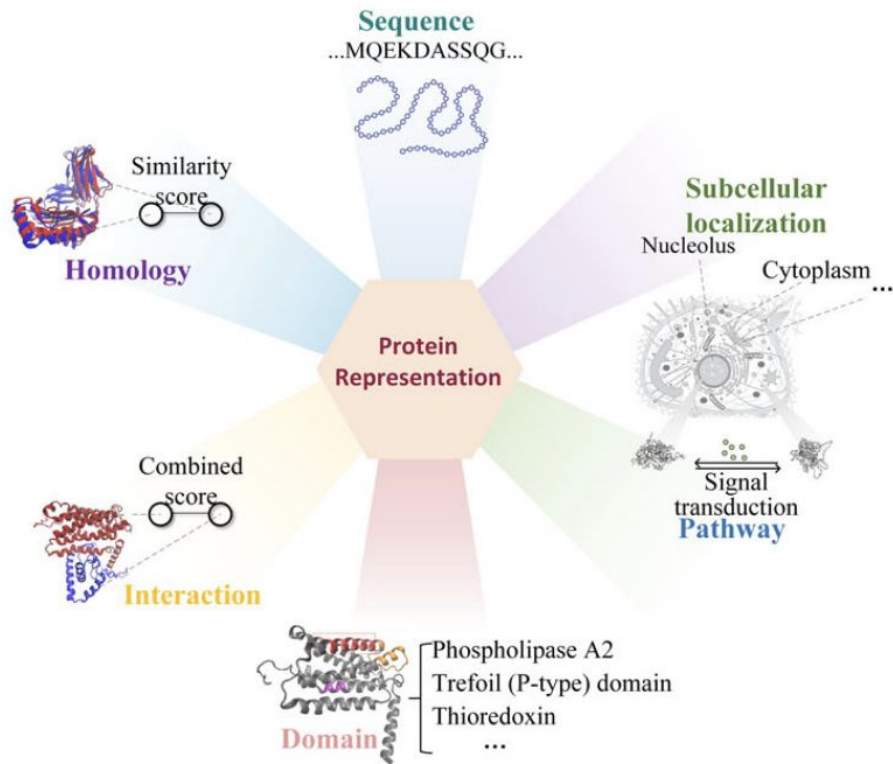
MIF2GO: Multimodal Information Fusion for GO Annotation

Siamese Contrastive Autoencoder (SCA): Encodes domain, localization, pathway, interaction, and homology relationships.
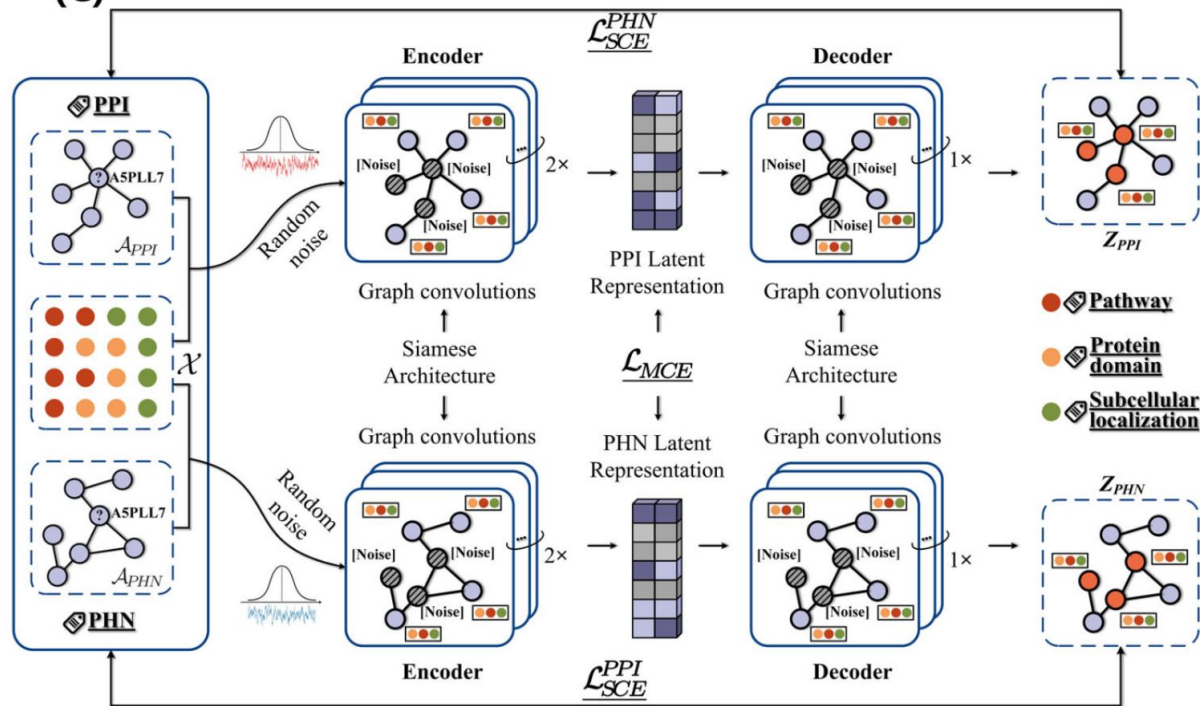
LM-HAW: Self-supervised language model for sequence features with hierarchical adaptive weighting.

Modal Hypernode Pooling (MHP): Final fusion of all modalities into unified protein representations.
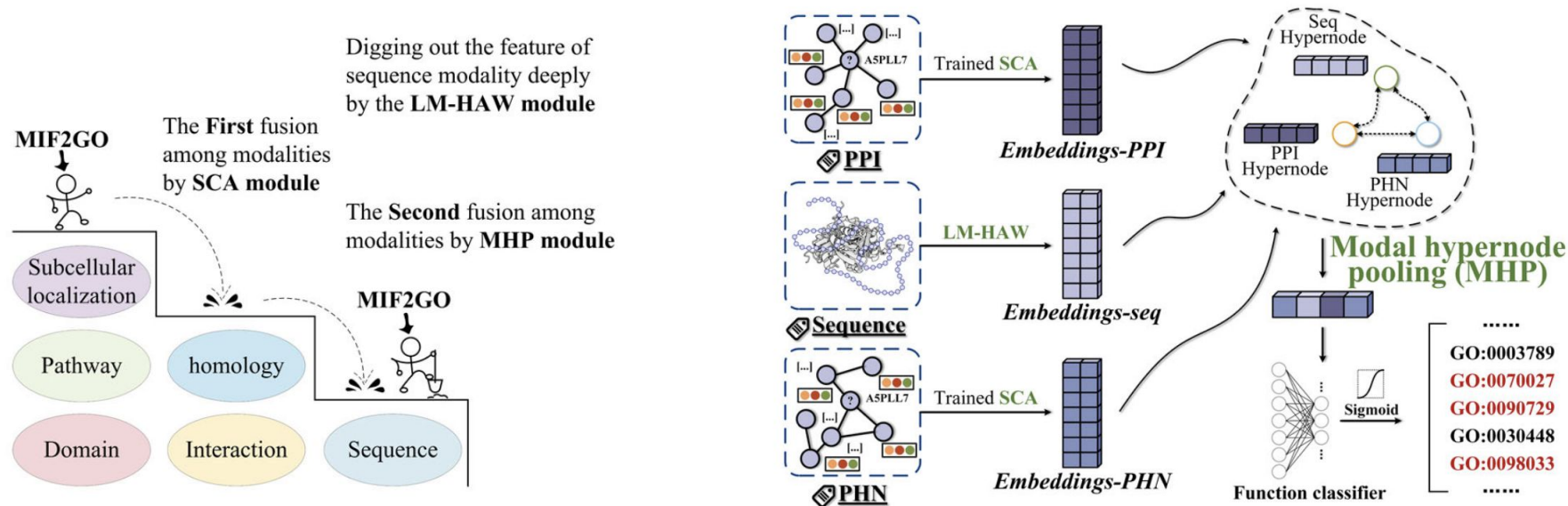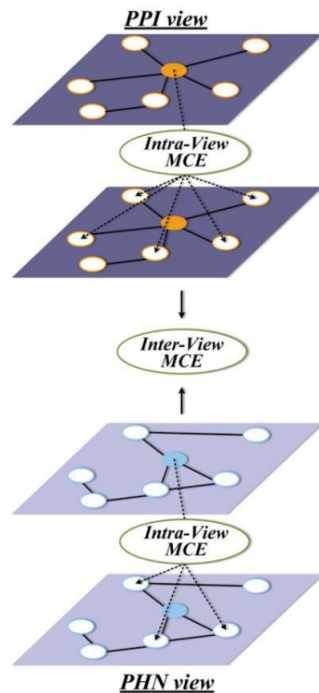
MIF2GO: Siamese Contrastive Autoencoder (SCA)

PPI: STRING **PHN:** BLAST

Siamese Contrastive autoencoder (SCA)

MIF2GO: Siamese Contrastive Autoencoder (SCA)

Encoder

$$H_{PPI} = \text{GCN}_{PPI}^{enc}\left(\widetilde{\mathcal{X}}, \mathcal{A}_{PPI}\right)$$

$$H_{PHN} = \text{GCN}_{PHN}^{enc}\left(\widetilde{\mathcal{X}}, \mathcal{A}_{PHN}\right)$$

$$\text{GCN}^{enc}(\widetilde{\mathcal{X}}, \mathcal{A}) = \widetilde{\mathcal{A}} ReLU\left(\widetilde{\mathcal{A}} \widetilde{\mathcal{X}} W_0\right) W_1$$

Decoder

$$Z_{PPI} = \text{GCN}_{PPI}^{dec}\left(H_{PPI}, \mathcal{A}_{PPI}\right)$$

$$Z_{PHN} = \text{GCN}_{PHN}^{dec}\left(H_{PHN}, \mathcal{A}_{PHN}\right)$$

Loss

$$\mathcal{L}_{SCA} = \left(\mathcal{L}_{SCE}^{PPI} + \mathcal{L}_{SCE}^{PHN}\right) + \lambda\left(\mathcal{L}_{MCE}^{PPI} + \mathcal{L}_{MCE}^{PHN}\right)$$

Scaled Cosine Error

$$\mathcal{L}_{\text{SCE}} = \frac{1}{|\widetilde{\mathcal{V}}|} \sum_{v_i \in \widetilde{\mathcal{V}}} \left(1 - \frac{x_i^T z_i}{\|x_i\| \cdot \|z_i\|}\right)^{\gamma}, \gamma \geq 1$$
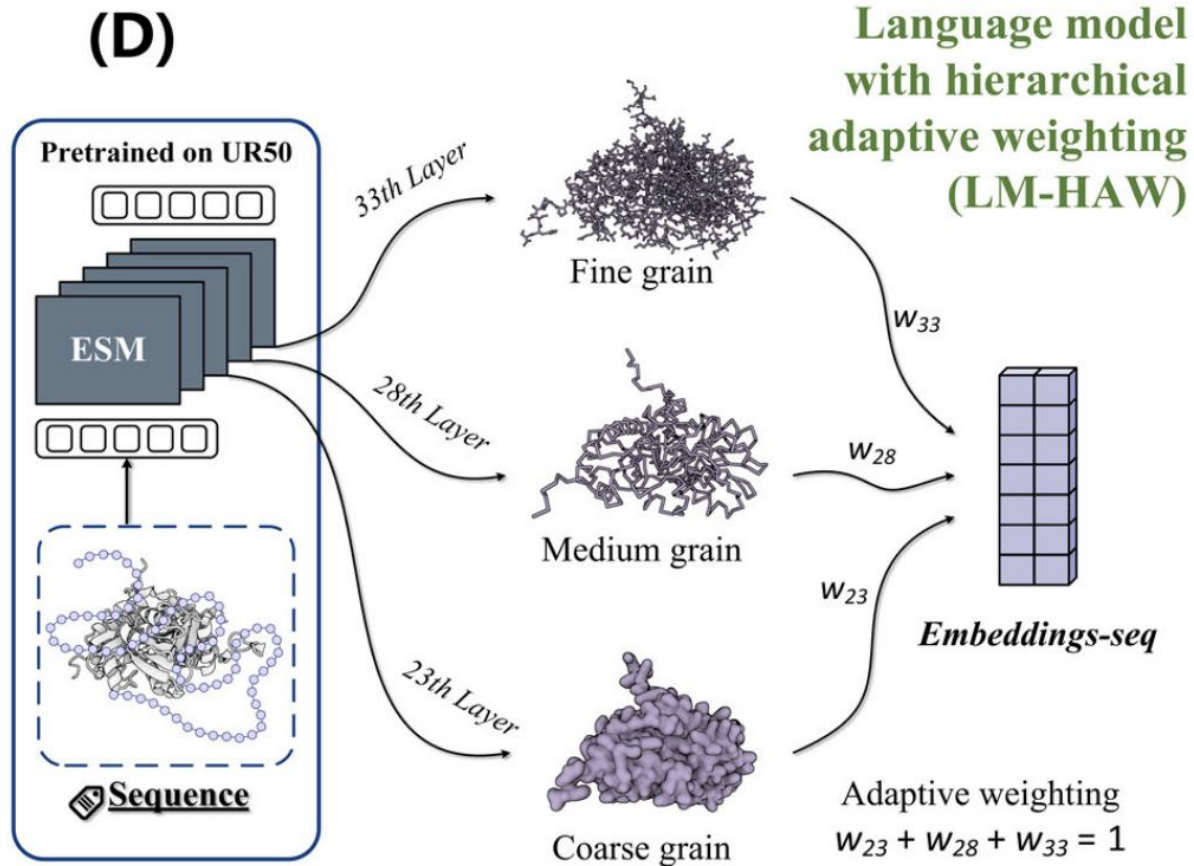
reconstruction

Multiview Contrastive Error

$$\mathcal{L}_{MCE}^{PPI} = \mathcal{L}_{Inter-view}^{PPI} + \mathcal{L}_{Intra-view}^{PPI}$$

$$= -\frac{1}{N} \sum_{i=0}^{N} \left(\sum_{j=0}^{N} \mathcal{A}_{ij}^{PHN} \times \log\left(\frac{\exp\left(h_i^{PPI} \cdot h_j^{PHN^T}/\tau\right)}{\sum_{k=0}^{N} \exp\left(h_i^{PPI} \cdot h_k^{PHN^T}/\tau\right)}\right) \right.$$

$$\left. + \sum_{j=0}^{N} \mathcal{A}_{ij}^{PPI} \times \log\left(\frac{\exp\left(h_i^{PPI} \cdot h_j^{PPI^T}/\tau\right)}{\sum_{k=0}^{N} \exp\left(h_i^{PPI} \cdot h_k^{PPI^T}/\tau\right)}\right) \right)$$
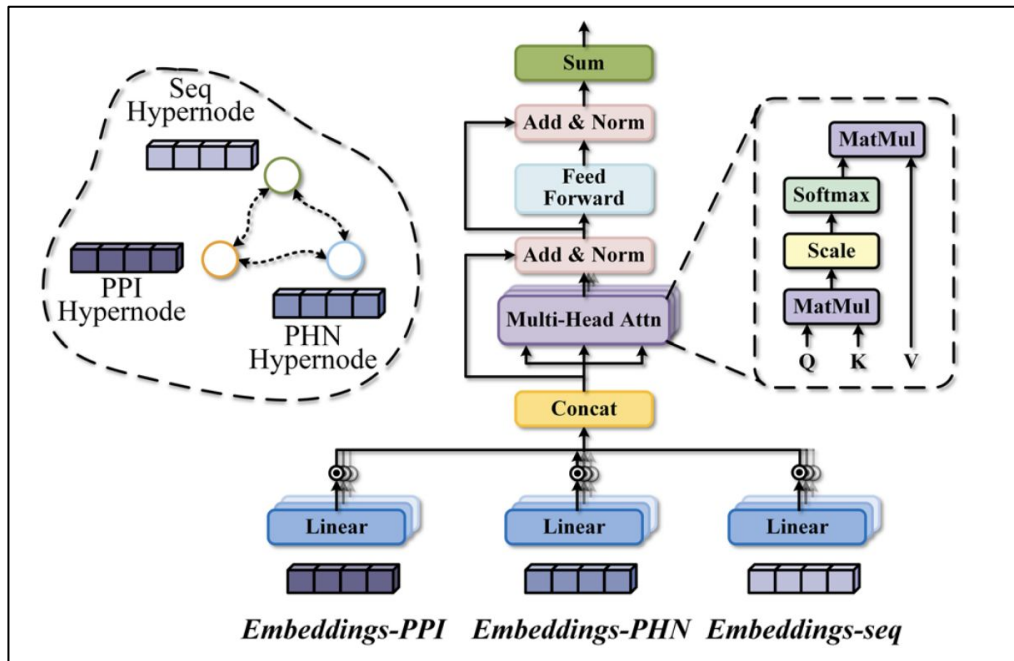
aligning interaction and homology modalities

MIF2GO: LM-HAW

$$emb^p_{seq} = \text{MLP}_{seq} \left(W_{33} \frac{1}{L} \sum_{i=0}^{L} \mathcal{X}^{33}_{seq_i} + W_{28} \frac{1}{L} \sum_{i=0}^{L} \mathcal{X}^{28}_{seq_i} \right.$$
$$\left. + W_{23} \frac{1}{L} \sum_{i=0}^{L} \mathcal{X}^{23}_{seq_i} \right)$$

(D)

Pretrained on UR50

ESM

Sequence

33th Layer

28th Layer

23th Layer

Fine grain

Medium grain

Coarse grain

Language model with hierarchical adaptive weighting (LM-HAW)

w_{33}

w_{28}

w_{23}

Embeddings-seq

Adaptive weighting
$w_{23} + w_{28} + w_{33} = 1$

MIF2GO: Modal hypernode pooling (MHP)

$$F_{(0)}^{p} = \left[\text{MLP}_{PPI}\left(emb_{PPI}^{p}\right) \parallel \text{MLP}_{PHN}\left(emb_{PHN}^{p}\right) \right.$$

$$\left. \parallel \text{MLP}_{seq}\left(emb_{seq}^{p}\right) \right]$$

$$F_{(1)}^{p} = \text{MultiHead}\left(F_{(0)}^{p}, F_{(0)}^{p}, F_{(0)}^{p}\right) + \text{Norm}\left(F_{(0)}^{p}\right)$$

$$F_{(2)}^{p} = \text{FFN}\left(F_{(1)}^{p}\right) + \text{Norm}\left(F_{(1)}^{p}\right)$$

MIF2GO outperforms existing methods for human protein function prediction

Methods	MF			BP			CC		
	M- AUPR	m- AUPR	F- max	M- AUPR	m- AUPR	F- max	M- AUPR	m- AUPR	F- max
DeepNF	0.489 (N/A)	0.636 (N/A)	0.618 (N/A)	0.261 (N/A)	0.425 (N/A)	0.447 (N/A)	0.454 (N/A)	0.626 (N/A)	0.597 (N/A)
Mashup	0.540 (N/A)	0.704 (N/A)	0.681 (N/A)	0.302 (N/A)	0.490 (N/A)	0.496 (N/A)	0.497 (N/A)	0.696 (N/A)	0.655 (N/A)
DeepFRI (GCN)	0.520 (0.003)	0.737 (0.002)	0.708 (0.002)	0.187 (0.001)	0.378 (0.005)	0.415 (0.003)	0.334 (0.004)	0.635 (0.007)	0.615 (0.004)
DeepFRI (GAT)	0.532 (0.002)	0.742 (0.001)	0.709 (0.002)	0.192 (0.002)	0.384 (0.003)	0.417 (0.003)	0.348 (0.004)	0.647 (0.003)	0.622 (0.003)
DeepFRI (SAGEConv)	0.517 (0.002)	0.740 (0.003)	0.707 (0.003)	0.188 (0.003)	0.386 (0.005)	0.417 (0.003)	0.345 (0.002)	0.649 (0.006)	0.623 (0.004)
DeepFRI (ChebConv)	0.517 (0.006)	0.737 (0.004)	0.705 (0.003)	0.189 (0.003)	0.384 (0.002)	0.417 (0.003)	0.340 (0.005)	0.643 (0.003)	0.621 (0.003)
GAT-GO (GCN)	0.530 (0.001)	0.749 (0.002)	0.703 (0.001)	0.200 (0.003)	0.393 (0.007)	0.420 (0.003)	0.357 (0.005)	0.667 (0.002)	0.626 (0.001)
GAT-GO (GAT)	0.531 (0.001)	0.751 (0.002)	0.706 (0.001)	0.204 (0.002)	0.401 (0.002)	0.425 (0.001)	0.365 (0.003)	0.670 (0.001)	0.628 (0.002)
GAT-GO (SAGEConv)	0.521 (0.004)	0.727 (0.004)	0.699 (0.006)	0.202 (0.002)	0.381 (0.003)	0.417 (0.002)	0.352 (0.002)	0.647 (0.002)	0.619 (0.002)
GAT-GO (ChebConv)	0.517 (0.009)	0.729 (0.006)	0.699 (0.010)	0.198 (0.004)	0.376 (0.005)	0.411 (0.003)	0.349 (0.005)	0.637 (0.007)	0.616 (0.003)
HEAL (GCN)	0.491 (0.004)	0.730 (0.003)	0.698 (0.002)	0.169 (0.007)	0.363 (0.008)	0.395 (0.007)	0.300 (0.005)	0.635 (0.009)	0.605 (0.005)
HEAL (GAT)	0.486 (0.007)	0.726 (0.003)	0.696 (0.004)	0.166 (0.007)	0.364 (0.003)	0.394 (0.002)	0.305 (0.006)	0.633 (0.004)	0.607 (0.002)
HEAL (SAGEConv)	0.494 (0.006)	0.729 (0.006)	0.701 (0.004)	0.166 (0.003)	0.365 (0.006)	0.394 (0.005)	0.311 (0.006)	0.632 (0.006)	0.606 (0.005)
HEAL (ChebConv)	0.496 (0.004)	0.735 (0.005)	0.700 (0.003)	0.170 (0.003)	0.369 (0.004)	0.398 (0.003)	0.313 (0.002)	0.637 (0.005)	0.609 (0.004)
MultiPredGO	0.444 (0.008)	0.603 (0.019)	0.591 (0.015)	0.242 (0.006)	0.454 (0.006)	0.465 (0.003)	0.374 (0.011)	0.698 (0.001)	0.657 (0.002)
DeepFusionGO	0.546 (0.003)	0.587 (0.013)	0.656 (0.005)	0.294 (0.006)	0.437 (0.011)	0.481 (0.005)	0.427 (0.006)	0.608 (0.011)	0.643 (0.009)
Graph2GO	0.560 (N/A)	0.761 (N/A)	0.718 (N/A)	0.284 (N/A)	0.488 (N/A)	0.490 (N/A)	0.494 (N/A)	0.751 (N/A)	0.686 (N/A)
DeepGraphGO	0.591 (0.001)	0.786 (0.002)	0.751 (0.001)	0.277 (0.002)	0.470 (0.003)	0.489 (0.001)	0.453 (0.002)	0.678 (0.006)	0.649 (0.003)
CFAGO	0.353 (0.005)	0.721 (0.002)	0.669 (0.004)	0.192 (0.002)	0.484 (0.003)	0.488 (0.002)	0.365 (0.001)	**0.771 (0.001)**	**0.705 (0.001)**
MIF2GO (Ours)	**0.624 (0.002)**	**0.804 (0.002)**	**0.758 (0.001)**	**0.317 (0.002)**	**0.523 (0.001)**	**0.514 (0.001)**	**0.528 (0.002)**	0.757 (0.003)	0.697 (0.002)

MIF2GO demonstrates good generalization and robustness across other species

Supplementary Table 1. Performance comparison of MIF2GO and other baseline methods on the fruit fly dataset.

Methods	MF			BP			CC		
	M-AUPR	m-AUPR	F-max	M-AUPR	m-AUPR	F-max	M-AUPR	m-AUPR	F-max
DeepFRI	0.611	0.768	0.731	0.250	0.420	0.440	0.413	0.697	0.667
	(0.005)	(0.004)	(0.006)	(0.003)	(0.009)	(0.005)	(0.006)	(0.007)	(0.003)
GAT-GO	0.589	0.750	0.697	0.253	0.441	0.449	0.397	0.710	0.664
	(0.009)	(0.006)	(0.006)	(0.008)	(0.003)	(0.004)	(0.006)	(0.017)	(0.011)
MultipredGO	0.444	0.603	0.591	0.242	0.454	0.465	0.374	0.698	0.657
	(0.009)	(0.019)	(0.015)	(0.006)	(0.007)	(0.004)	(0.011)	(0.001)	(0.002)
HEAL	0.597	0.761	0.715	0.223	0.400	0.419	0.381	0.684	0.645
	(0.005)	(0.006)	(0.009)	(0.005)	(0.010)	(0.008)	(0.006)	(0.007)	(0.008)
DeepFusionGO	0.546	0.587	0.656	0.294	0.437	0.481	0.427	0.608	0.643
	(0.003)	(0.013)	(0.005)	(0.006)	(0.011)	(0.005)	(0.006)	(0.011)	(0.009)
CFAGO	0.375	0.634	0.590	0.226	0.502	0.496	0.391	0.777	0.708
	(0.001)	(0.001)	(0.001)	(0.002)	(0.001)	(0.001)	(0.003)	(0.001)	(0.001)
MIF2GO	**0.682**	**0.808**	**0.764**	**0.378**	**0.546**	**0.536**	**0.576**	**0.794**	**0.735**
	(0.006)	(0.003)	(0.003)	(0.004)	(0.002)	(0.003)	(0.005)	(0.004)	(0.003)

Supplementary Table 2. Performance comparison of MIF2GO and other baseline methods on the mouse dataset.

Methods	MF			BP			CC		
	M-AUPR	m-AUPR	F-max	M-AUPR	m-AUPR	F-max	M-AUPR	m-AUPR	F-max
DeepFRI	0.533	0.733	0.709	0.174	0.362	0.399	0.307	0.621	0.602
	(0.004)	(0.003)	(0.002)	(0.001)	(0.004)	(0.002)	(0.005)	(0.007)	(0.004)
GAT-GO	0.541	0.743	0.696	0.177	0.385	0.408	0.333	0.650	0.609
	(0.007)	(0.002)	(0.002)	(0.002)	(0.003)	(0.002)	(0.006)	(0.002)	(0.003)
MultipredGO	0.311	0.606	0.594	0.128	0.400	0.430	0.274	0.621	0.603
	(0.007)	(0.009)	(0.008)	(0.001)	(0.002)	(0.003)	(0.002)	(0.003)	(0.002)
HEAL	0.501	0.726	0.694	0.148	0.341	0.376	0.289	0.618	0.591
	(0.006)	(0.003)	(0.002)	(0.006)	(0.006)	(0.004)	(0.008)	(0.005)	(0.003)
DeepFusionGO	0.560	0.695	0.703	0.241	0.398	0.455	0.390	0.599	0.616
	(0.003)	(0.005)	(0.002)	(0.001)	(0.002)	(0.002)	(0.003)	(0.003)	(0.002)
CFAGO	0.369	0.711	0.659	0.177	0.480	0.485	0.356	**0.771**	**0.703**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	**(0.001)**	**(0.001)**
MIF2GO	**0.627**	**0.794**	**0.748**	**0.284**	**0.489**	**0.493**	**0.509**	0.745	0.685
	(0.002)	(0.001)	(0.001)	(0.002)	(0.001)	(0.002)	(0.005)	(0.004)	(0.001)

MIF2GO outperforms other baseline methods on the GO terms with few protein samples

Table 2 | Comparison of different methods on the GO terms with few protein samples using F-max metric

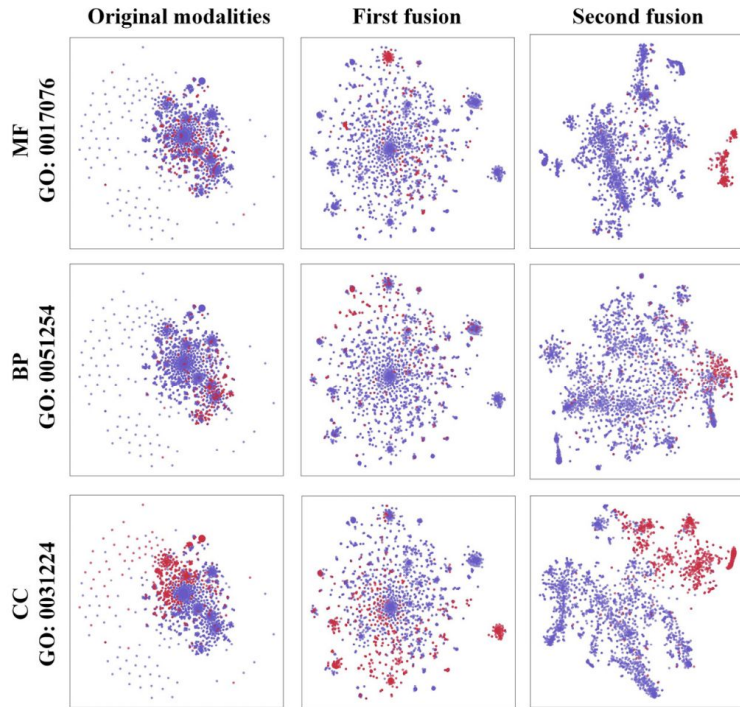| Methods	Min-30[a]			31-100			101-300		
	MF	BP	CC	MF	BP	CC	MF	BP	CC
DeepFRI	0.497 (0.004)	0.172 (0.007)	0.307 (0.004)	0.513 (0.003)	0.231 (0.009)	0.301 (0.007)	0.657 (0.008)	0.279 (0.006)	0.359 (0.003)
GAT-GO	0.489 (0.002)	0.189 (0.001)	0.341 (0.005)	0.505 (0.002)	0.240 (0.004)	0.312 (0.003)	0.641 (0.001)	0.287 (0.001)	0.355 (0.003)
MultiPredGO	0.235 (0.007)	0.108 (0.005)	0.258 (0.005)	0.345 (0.003)	0.202 (0.004)	0.287 (0.002)	0.523 (0.002)	0.284 (0.004)	0.336 (0.003)
HEAL	0.468 (0.007)	0.147 (0.003)	0.278 (0.003)	0.480 (0.010)	0.199 (0.005)	0.266 (0.008)	0.640 (0.006)	0.247 (0.004)	0.323 (0.007)
DeepFusionGO	0.526 (0.004)	0.222 (0.003)	0.393 (0.003)	0.531 (0.006)	0.273 (0.002)	0.346 (0.005)	0.629 (0.002)	0.312 (0.001)	0.372 (0.003)
CFAGO	0.463 (0.001)	0.251 (0.001)	0.471 (0.003)	0.468 (0.002)	0.310 (0.001)	0.436 (0.001)	0.588 (0.001)	0.363 (0.002)	0.476 (0.002)
MIF2GO	**0.600 (0.003)**	**0.278 (0.002)**	**0.481 (0.003)**	**0.600 (0.001)**	**0.350 (0.005)**	**0.450 (0.002)**	**0.717 (0.004)**	**0.403 (0.006)**	**0.486 (0.002)**

[a]'Min-30' refers to the GO terms whose number of protein samples is less than 30.
Bold indicates the optimal performance.

Performance comparison of MIF2GO on shallow terms and specific terms

**Table 3 | F-max performance of MIF2GO on different depth of shallow terms and specific terms**

		Depth-2	Depth-3	Depth-4	Depth-5	Depth-6
MF	Shallow terms	0.848 (0.001)	0.841 (0.003)	0.818 (0.001)	0.781 (0.004)	0.764 (0.002)
	Specific terms	0.741 (0.003)	0.694 (0.003)	0.762 (0.002)	0.775 (0.001)	0.772 (0.003)
	The number of GO term	12	27	74	165	325
BP	Shallow terms	0.635 (0.005)	0.691 (0.006)	0.657 (0.002)	0.606 (0.003)	0.576 (0.002)
	Specific terms	0.526 (0.003)	0.462 (0.005)	0.401 (0.002)	0.357 (0.004)	0.372 (0.002)
	The number of GO term	14	43	135	364	779
CC	Shallow terms	0.661 (0.001)	0.684 (0.003)	0.672 (0.003)	0.662 (0.005)	0.671 (0.004)
	Specific terms	0.358 (0.003)	0.429 (0.006)	0.478 (0.002)	0.529 (0.005)	0.560 (0.002)
	The number of GO term	34	87	165	289	403

MIF2GO learns distinguishable representations after two stages of fusion

Original modalities **First fusion** **Second fusion**

MF GO: 0017076

BP GO: 0051254

CC GO: 0031224

Supplementary Figure 2. The evaluation of feature distribution in the human dataset throughout the fusion process. For each category, a representative GO term is selected for visualization. Samples with that GO term (i.e., positive samples) are colored in red. With the progress of fusion, the positive samples tend to be more clustered and more even.
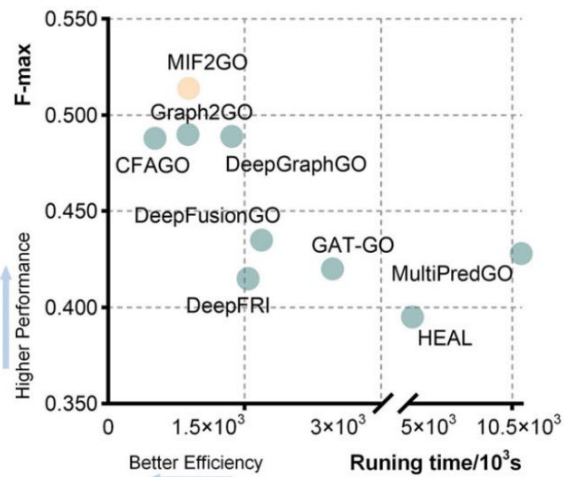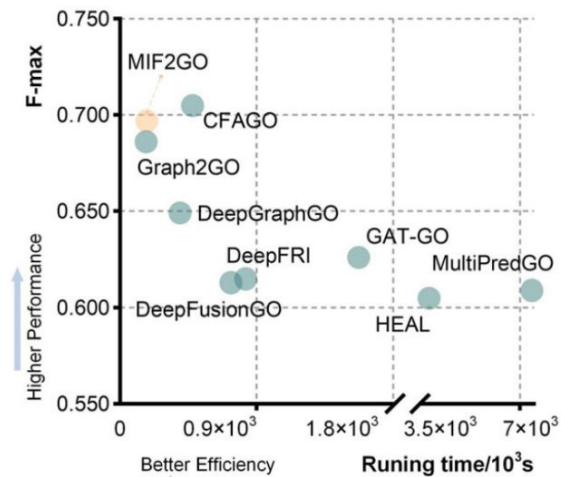
Ablation Studies
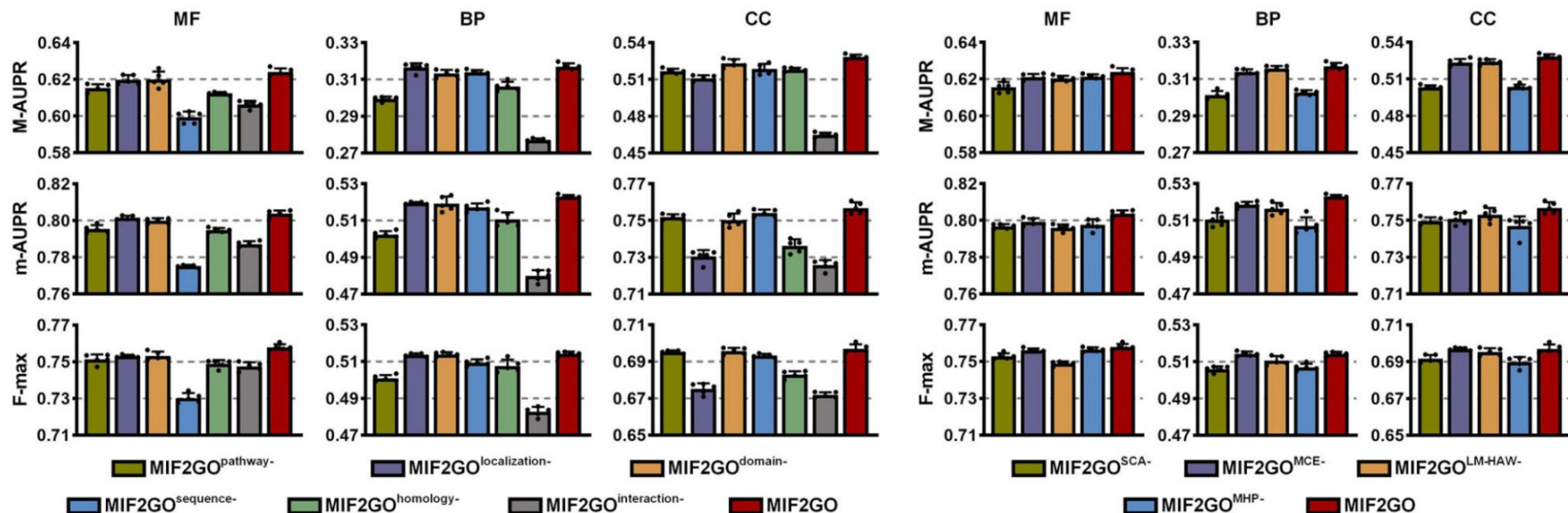
Execution Times
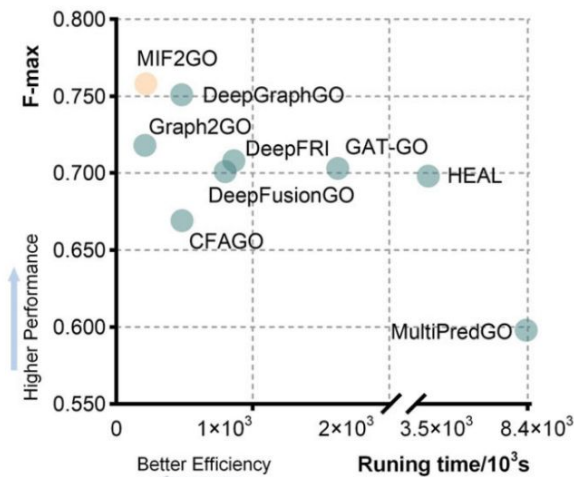
(A) MF **(B) BP** **(C) CC**

Modal missing scenarios of MIF2GO on the CAFA-subset benchmark

Methods	MF			BP			CC		
	M- AUPR	m- AUPR	F- max	M- AUPR	m- AUPR	F- max	M- AUPR	m- AUPR	F- max
DeepFRI	0.341 (0.002)	0.553 (0.019)	0.547 (0.010)	0.115 (0.004)	0.392 (0.014)	0.441 (0.011)	0.186 (0.001)	0.633 (0.015)	0.610 (0.005)
GAT-GO	0.357 (0.003)	0.579 (0.006)	0.555 (0.004)	0.122 (0.002)	0.445 (0.004)	0.470 (0.002)	0.205 (0.004)	0.649 (0.003)	0.610 (0.001)
MultiPredGO	0.152 (0.004)	0.451 (0.005)	0.471 (0.006)	0.065 (0.005)	0.387 (0.007)	0.433 (0.006)	0.089 (0.006)	0.568 (0.006)	0.569 (0.003)
HEAL	0.342 (0.006)	0.541 (0.014)	0.541 (0.011)	0.115 (0.004)	0.413 (0.006)	0.449 (0.004)	0.175 (0.001)	0.653 (0.008)	0.618 (0.002)
DeepFusionGO	0.359 (0.005)	0.578 (0.006)	0.565 (0.002)	0.145 (0.001)	0.472 (0.001)	0.490 (0.002)	0.198 (0.004)	0.663 (0.002)	0.618 (0.001)
CFAGO	0.145 (0.001)	0.509 (0.002)	0.505 (0.003)	0.105 (0.001)	0.465 (0.003)	0.478 (0.004)	0.150 (0.001)	0.697 (0.001)	0.649 (0.001)
MIF2GO (LM-HAW)	0.392 (0.004)	0.609 (0.001)	0.579 (0.001)	0.150 (0.002)	0.466 (0.010)	0.488 (0.001)	0.199 (0.004)	0.659 (0.001)	0.620 (0.002)
MIF2GO	**0.398 (0.002)**	**0.611 (0.011)**	**0.586 (0.007)**	**0.206 (0.002)**	**0.524 (0.006)**	**0.532 (0.001)**	**0.293 (0.005)**	**0.699 (0.005)**	**0.650 (0.005)**

Missing Modalities: Challenges and Solutions

1. **Sequence Information:**
   - Always available for all proteins.
   - Provides a robust backbone for MIF2GO.
2. **Protein Homology Network (PHN):**
   - Adaptable threshold for similarity scores ensures sufficient connectivity.
3. **Protein-Protein Interaction (PPI) Network:**
   - Interaction data gaps addressed by transferring relationships from homologous proteins.
4. **Other Modalities (Pathway, Subcellular Localization, Domains):**
   - Initialized as zero vectors when unavailable.
   - Biological features aggregated from neighboring nodes via GCN.

Conclusion

- **MIF2GO** is proposed to integrate **multimodal** biological information for accurate Gene Ontology (GO) term prediction.
  1. **Siamese Contrastive Autoencoder (SCA):** Self-supervised encoding of domain, localization, and pathway modalities into interaction and homology relationships.
  2. **Language Model with Hierarchical Adaptive Weighting (LM-HAW):** Captures fine-grained and coarse-grained sequence features.
  3. **Modal Hypernode Pooling (MHP):** Fuses interaction, homology, and sequence modalities into unified representations.
- **Performance:** Outperformed state-of-the-art methods across seven benchmark datasets.
- **Robustness:** Demonstrated generalizability across species and resilience to missing modalities.
- **Insights:** Revealed critical associations between modalities and protein functions, validating the importance of multimodal integration.