

Dear Reviewer,

Thank you very much for your insightful and constructive comments on our manuscript (metabolites-4193302). We have carefully considered all your suggestions and have revised the manuscript accordingly. All changes in the revised version are highlighted in **blue** to distinguish them from first-round revisions (which were in red). Below we provide our point-by-point responses to your comments.

Previous Major 6:

The authors were suggested to benchmark enrichment against background gene sets and discuss whether identified pathways are enriched beyond what is expected from the training gene list alone. But the authors did not explicitly compare their enrichment results against the training gene background. The concern about pathways like PI3K-Akt and focal adhesion simply reflecting training data bias remains partially unaddressed. The authors were suggested to include a control analysis showing enrichment p-values when using the training positive genes as the test set, demonstrating that predicted genes show additional pathway enrichment beyond known drivers.

Response: We thank the reviewer for this important follow-up. To directly address the concern about potential bias from the training data, we performed the recommended control analysis. Specifically, we conducted the same GO and KEGG enrichment analysis (hypergeometric test, Benjamini-Hochberg FDR correction) on the set of known driver genes (positive samples) used for model training. The results are summarized in **Supplementary Table S6**.

The comparison reveals that both the predicted driver genes and the known driver genes are significantly enriched in classic cancer pathways such as PI3K-Akt signaling (predicted: $FDR = 1.04e-20$; known: $FDR = 2.68e-31$) and focal adhesion (predicted: $FDR = 1.55e-27$; known: $FDR = 3.80e-15$). Importantly, the predicted genes exhibit additional significant enrichments in processes related to anatomical structure morphogenesis, regulation of cellular processes, and cell surface receptor signaling—terms that are not significantly enriched in the known driver gene set. This indicates that the pathway enrichments observed in our predicted genes are not merely a reflection of the training data but include novel biological signals that extend beyond known driver genes. We have added a brief description of this analysis in Section 3.4 (highlighted in **blue**) and provided the full results in Supplementary Table S6. We believe this addition fully addresses the reviewer's concern.

--- Page 15, Section 3.4 Enrichment Analysis, line 549-561

Minors

Comments1. In the Line 7: 'Graph Convolutional Network (GCN)' - The acronym GCN is introduced here but not used consistently throughout. Consider using 'GCN' after first definition.

Response 1: We thank the reviewer for this careful observation. In the revised manuscript, we have checked all occurrences of "Graph Convolutional Network" after its first definition in the Abstract and consistently replaced them with the acronym "GCN" to improve readability and consistency. The changes have been highlighted in **blue** to distinguish them from first-round revisions (which were marked in red). We believe this enhances the clarity of the text.

--- Page 2-3, Introduction, lines 48, 53, 54, 91, 97, 109;

Page 5, Section 2.2.2 Graph Structure Construction, line 191;

Page 16, Discussion, line 580

Page 17, Conclusions, line 620

Comments2. In the Line 9: 'Knowledge distillation is incorporated' - This passive construction could be more direct: 'We incorporate knowledge distillation...'

Response 2: We thank the reviewer for this helpful suggestion. In the revised manuscript, we have changed the passive construction to an active one in the Methods section of the Abstract. The sentence now reads: "We incorporate knowledge distillation by using soft labels from a pre-trained teacher model to enhance feature representation." This revision has been highlighted in **blue** to distinguish it from first-round changes (which were in red). We believe this improves the directness and clarity of the text.

--- Page 1, Abstract, line 9

Comments3. In the Line 96-100: The sentence introducing Figure 1's four modules is grammatically complex. Consider breaking into two sentences for clarity.

Response 3: We thank the reviewer for this helpful suggestion. In the revised manuscript, we have split the sentence describing the four core modules into two sentences to improve readability. The revised text now reads: “As illustrated in Figure 1, the M-GNN framework consists of four core modules. These are: a multi-omics encoder, a cross-group interaction fusion layer, a feature enhancement module, and a dual prediction head.” This change has been highlighted in **blue** to distinguish it from first-round revisions (which were in red). We believe this improves the clarity of the introduction.

--- Page 3, Introduction, lines 97-99

Comments4. In the Lines 136-140: The description of Z-score normalization could benefit from a brief explanation of why this prevents data leakage, as some readers may not immediately understand the concern.

Response 4: We thank the reviewer for this helpful suggestion. In the revised manuscript, we have added a brief explanation to clarify why calculating the normalization parameters solely on the training set prevents data leakage. The added sentence (highlighted in **blue**) states: “This ensures that no information from the validation or test sets influences the normalization parameters, which would otherwise constitute data leakage and lead to overly optimistic performance estimates.” We believe this clarification makes the methodological description more accessible to readers.

--- Page 4, Section 2.1 Materials line 139-142

Comments5. In the Line 226-229, the MLP subscripts (MLP1, MLP2, MLP3) suggest separate networks, but it's unclear if these share any parameters. Please clarify them.

Response 5: We thank the reviewer for this important observation. In the revised manuscript, we have added a brief clarification to the sentence describing the MLP processing. The text now reads: “These patterns are subsequently processed by a multilayer perceptron (MLP) for each view H21, H23, H24 (with independent parameters, i.e., no weight sharing) and are ultimately spliced and fused into a comprehensive representation of the expression omics H2.” This addition makes it explicit that the three MLPs are independent and do not share parameters, which aligns with the subscript notation and the architecture illustrated in Figure 1. The change has been highlighted in **blue** to distinguish it from first-round revisions

(which were in red). We believe this clarifies the model design.

--- Page 6, Section 2.2.3 Multi-omics Encoder, line 227

Comments6. In the Line 270-271: 'hidden dimensions of 128 in the first classifier layer and 512 in the second'. This sentence could be clearer: specify whether these are sequential layers or parallel branches.

Response 6: We thank the reviewer for this helpful suggestion. In the revised manuscript, we have added a brief clarification to the description of the teacher model's classifier layers to explicitly state that they are sequential. The text now reads: "with hidden dimensions of 128 in the first classifier layer and 512 in the second classifier layer (which are sequential layers, i.e., the output of the first serves as the input to the second)." This matches the standard design of the student model and the architecture illustrated in Figure 1. The change has been highlighted in **blue** to distinguish it from first-round revisions (which were in red). We believe this makes the architectural description clearer.

--- Page 7, Section 2.2.5, Feature Enhancement Module, lines 274-275

Comments7. In the Line 284: ' λ was set to 0.5 based on validation performance' - Consider mentioning what range of values was explored during tuning.

Response 7: We thank the reviewer for this suggestion. In the revised manuscript, we have added a brief description of the range explored for λ during hyperparameter tuning. The sentence now reads: "The value of λ was set to 0.5 based on validation performance (explored within the range of 0.1 to 0.9 with a step of 0.1)." This range was chosen to cover typical values used in knowledge distillation and was evaluated using grid search on the validation set, consistent with the hyperparameter optimization described in Supplementary Material Section 3.5. The change has been highlighted in **blue** to distinguish it from first-round revisions (which were in red). We believe this addition improves the transparency and reproducibility of the method.

--- Page 7, Section 2.2.5 Feature Enhancement Module, line 289

Comments8. In the Line 313 (Equation 15): The output layer bias is denoted b' (with prime), while Equation 19 uses b'' (double prime). This notation is inconsistent

with standard practice; consider using different subscripts instead.

Response 8: We thank the reviewer for this important observation. To improve consistency and clarity, we have revised the notation for the output layer parameters in Equations (15) and (19). Specifically, we now use subscript G for the GCN classifier (i.e., W_G , b_G) and subscript MLP for the MLP predictor (i.e., W_{MLP} , b_{MLP}). This avoids the ambiguous prime/double-prime notation and follows standard practice. The revised equations and the accompanying text have been highlighted in **blue** to distinguish them from first-round revisions (which were in red). We believe this change makes the notation clearer and more systematic.

--- Page 8-9, Section 2.2.5 Architectural Principle of Dual Predictive Head, lines 318-321, 334-337, Equations (15) and (19)

Comments9. In the Line 381-382: 'M-GNN consistently outperforms existing baseline methods' - This claim should be softened to 'consistently achieves the highest or second-highest performance' to match the actual Table 1 results.

Response 9: We thank the reviewer for this important observation. We agree that the original wording was too strong and did not accurately reflect the results in Table 1, where M-GNN achieves the highest AUPR on STRINGdb and IRefIndex but the second-highest on CPDB (behind SGCD). Therefore, we have revised the sentence to state that M-GNN "consistently achieves the highest or second-highest AUPR" across the three networks. This change has been highlighted in **blue** to distinguish it from first-round revisions (which were in red). We believe this revised description is now fully consistent with the reported data.

--- Page 10, Section 3.1 Performance Evaluation of the M-GNN, lines 389-390

Comments10. In the Figure 2: The legend font size appears small in the reproduced figure. Consider increasing for better readability in print.

Response 10: We thank the reviewer for the helpful suggestion. In the revised figure (Figure 2), we have increased the font sizes of the legend, axis labels, tick labels, and title to improve readability in print. All modifications are clearly indicated in the updated figure.

--- Page 11, Section 3.1 Performance Evaluation of the M-GNN, Figure 2

We sincerely appreciate your time and expertise in reviewing our work. We hope that the revised manuscript now meets your expectations. Please do not hesitate to contact us if any further clarification is needed.

Sincerely,

Wen Zhu (on behalf of all authors)

School of Mathematics and Statistics, Hainan Normal University

Email: syzhuwen@163.com