








A Deep Learning Multi-omics Framework to Combine Microbiome and Metabolome Profiles for Disease Classification

Andrea Licciardi^(✉), Antonino Fiannaca, Massimo La Rosa,
Maurizio Alfonso Urso, and Laura La Paglia

ICAR-CNR, National Research Council of Italy, Via Ugo La Malfa 153,
90146 Palermo, Italy

andrea.licciardi@icar.cnr.it

Abstract. Microbiome and metabolome contain information about host disease. Therefore, a multi-omics analysis of these data types can provide key constraints for disease classification. However, due to multi-omics data's complex and high-dimensional nature, classical statistical methods struggle to capture the shared information between microbiome and metabolome. Deep learning represents a power framework to address this issue. We design a deep learning model for the integrated analysis of microbiome and metabolome that leverages the complementary information between the two datasets to perform a medical diagnosis of a given disease as a supervised classification task. We test our approach on six different matched microbiome/metabolome datasets, related to diverse pathologies. A comparative performance analysis shows that our proposed model called microBiome-metaBolome Network (BiBoNet) performs better than classical machine learning methods. In addition, we show that BiBoNet achieves better results than deep learning models based on individual or combined data. We highlight the importance of multi-omics integration through deep learning for improved medical diagnosis using microbiome and metabolome.

Keywords: Deep learning · Multi-omics · Microbiome · Metabolome · Data integration

1 Introduction

Bioinformatics significantly contributes to translational research, which allows the integration of different layers of biological information in a data-driven approach [1]. Moreover, the development of sequencing technology has recently led to the production of “multi-omics” data, opening new possibilities for integrated system-level methods. This innovative multi-omics approach, combined with bioinformatics and machine learning analysis methods, has the advantage of yielding a better understanding and a clearer picture of the system under study. Moreover, it allows us to holistically emphasize the links between the different

© The Author(s) 2024

M. Wand et al. (Eds.): ICANN 2024, LNCS 15023, pp. 3–14, 2024.

https://doi.org/10.1007/978-3-031-72353-7_1

types of molecules and their roles in complex biological processes [2–4]. Indeed, each omic data type provides a different knowledge of the biological system under investigation. For instance, genomics can lead to evaluate gene-environment interactions or functions of genomics variants; it can be employed in biomarker discovery and in the comprehension of diagnosing diseases such as cancer and other chronic diseases, monitoring their progression, predicting recurrence, and supporting the identification of therapeutic treatment [5]. Metabolomics is the study of the chemical processes involving the small molecule substrates, intermediates, and products of cell metabolism, called metabolites. This omics science aims at identifying and quantifying all metabolites in a given organism or biological sample [6]. Metabolomics can lead to the characterization of different response patterns in humans. It is used for pharmacokinetics, pharmacodynamics, and precision medicine applications [7]. However, single omic can give a partial view of molecular events governing the disease state [8]. Integrating diverse omics data helps clarify the disease’s underlying pathogenic alterations, which can be confirmed by further molecular study.

Recently, metabolomics has been increasingly integrated with metagenomics data, aiming to investigate co-variation patterns between metabolites and microbiota [9]. The human gut microbiome is a sophisticated biological system that helps the host produce vitamins, break down macromolecules, and strengthen the host immune system, among other essential tasks. High-throughput sequencing methods are becoming more and more accessible, and this has allowed for critical new understandings of the composition and possible use of the microbiome [10–12]. For example, metagenomics research on the gut microbiome has revealed that the metabolic potentials of lean and obese mice differ and that community structures alter in response to dietary modifications [13, 14].

In this context, machine learning models, especially neural networks (NN), provide efficient solutions to integrate different kinds of omics data and, this way, capture and take advantage of different information content [15]. In this work, we present a composite neural network model that is able to process two types of omic data, i.e. metabolome and microbiome, in order to exploit both information sources and provide better performance in a supervised classification task. In particular, we benchmark our model with six different matched metabolome/microbiome datasets, related to several pathologies. The classification is between disease patients and healthy control. Our approach is compared with classical machine learning (ML) algorithms, such as random forest and support vector machine. Experimental results demonstrate that, in most cases, our approach reaches the best results in terms of several classification scores.

2 Related Works

One of the main advantages of deep learning (DL) is its ability to learn complex and non-linear functions to map high-dimensional and heterogeneous data (such as multi-omics data) into a desired output. Indeed, DL provides an attractive framework for analysing multi-omics data [15] and references therein. In

the context of the present work (supervised classification task), NN has been used to integrate multi-omics data for the prediction of drug response [16], drug synergy scores [17, 18] and classification of breast cancer subtypes [19]. Convolutional neural networks (CNN) fed with multi-omics data have provided successful examples of classifying molecular subtypes of breast cancer [20] or predicting gene regulation mechanisms [21]. More recently, multi-omics graph convolutional networks and graph attention models have been implemented for patient classification [22] and disease diagnosis [23].

All these studies have shown that successful integration of multi-omics data through DL provides increased performance over classical ML methods or single-dataset analysis. Despite these significant advances, DL-based multi-omics studies mostly focused on gene expression data combined with additional omics (e.g. gene copy number, mutation, DNA methylation, mRNA or protein expression). Recently, notable efforts have focused on studying the relationship between the gut microbiome and metabolome with DL [24–26]. Less attention has been paid to integrating metabolome and microbiome through DL-based strategies for the classification of subjects based on a specific disease [27], which still relies on conventional ML methods (e.g. random forest [28]). This work aims to fill this gap.

3 Materials and Method

3.1 Data

This work relies on metabolome and microbiome datasets collected by previous studies and compiled by [29] (<https://github.com/borenstein-lab/microbiome-metabolome-curated-data/wiki>). This repository includes all publicly available datasets in which metabolites and microbes concentrations from human fecal samples are aligned, that is they belong to the same patient. We selected only those with more than 100 subjects from the fourteen studies in the repository. The final collection of datasets comprises six studies which we will be referring to as iHMP [30], KIM [31], PRISM [28], SINHA [32], WANG [33] and YACHIDA [34] in the rest of this paper. Table 1 reports the main characteristics of each dataset. We extracted patients’ information from the provided metadata, where each subject is attributed to one group between healthy control and study-specific pathology/condition. This information defines the target labels of our classification task. We left the original labels unchanged in all studies except for the YACHIDA dataset. For the latter, we only considered healthy and colorectal cancer (CRC) subjects by grouping patients with different stages of CRC (from stage 0 to stage 4) into a single class, as our main objective is to classify subjects based on a specific pathology rather than sub-stages of the same disease.

Data Preprocessing. Both for microbes and metabolites, we only kept features present in more than 10% of the subjects and discarded the rest. We also discarded columns with missing values in the iHMP datasets. For microbiome

Table 1. Description of the datasets used in this study. Abbreviations - IBD: inflammatory bowel disease, CD: Crohn’s disease, UC: ulcerative colitis, CRC: colorectal cancer, ESRD: end-stage renal disease, BP: before preprocessing, AP: after preprocessing.

Dataset	Ref.	Disease	# samples	classes (samples)	# metab. (BP)	# micro. (BP)	# metab. (AP)	# micro. (AP)
iHMP	[30]	IBD	105	Control (26) CD (49) UC (30)	81867	42872	5820	8893
KIM	[31]	CRC	240	Control (102) Adenoma (102) CRC (36)	462	499	462	170
PRISM	[28]	IBD	155	Control (34) CD (68) UC (53)	8848	200	8847	163
SINHA	[32]	CRC	131	Control (89) CRC (42)	530	86	530	86
WANG	[33]	ESRD	287	Control (67) ESRD (220)	277	56962	277	22114
YACHIDA	[34]	CRC	277	Control (127) CRC (150)	450	57702	298	17692

data we used species relative abundance (RA) where available, otherwise, we used genera RA (SINHA and KIM datasets) and took the log of the concentrations. For some datasets, metabolites abundance data were already normalized (WANG dataset) and/or centred around zero (KIM and SINHA datasets), otherwise we directly used counts data (PRISM, YACHIDA and iHMP datasets), keeping all reported metabolomics features (identified and unidentified metabolites). In those cases, following [24], we performed centre log-ratio transformation for counts data, adding a pseudo-count of one for metabolites with reported zero counts. The number of resulting features for metabolomics and microbiome data before and after preprocessing is reported in Table 1.

3.2 Model Description

In this work, we present BiBoNet (microBiome-metaBolome Network), a multi-layer perceptron (MLP) that performs a supervised classification task (binary or three-class) to distinguish between healthy (i.e., control) subjects and subjects affected by a given disease using multi-omics data, which are metabolome and microbiome data. A key aspect of any multi-omics workflow is how to perform the data integration (fusion). We follow the definitions used in [15] and differentiate between early and late fusion. In the former, omics data are concatenated first and then fed to a DL model (e.g. [17,21]). In the latter, the integration is performed by concatenating intermediate features from separate subnetworks (e.g., [16,18–20]). The resulting vector is fed to a subsequent model that performs the assigned task (in our case, classification). BiBoNet is sketched in Fig. 1. Its architecture is designed to perform late fusion by combining the hidden feature representation of two independently pre-trained neural networks, each one fed with a different omic (metabolic profiles and microbes abundance data, Fig. 1a

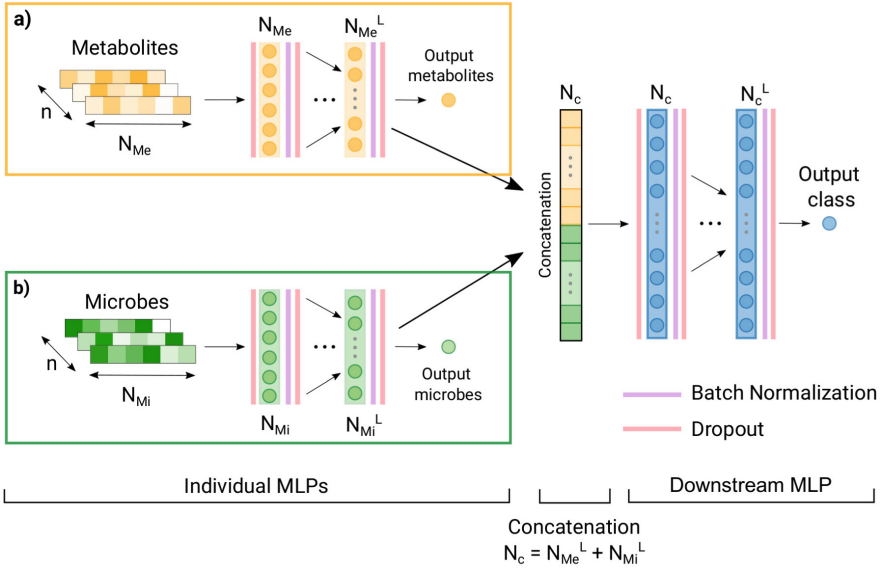


Fig. 1. Sketch of BiBoNet’s late fusion architecture. See the text for details about design and training strategy. Dimensions are reported for clarity. n is the total number of samples. N_{Me} is the number of metabolite input features, N_{Mi} is the number of microbes input features. N_C is the dimension of the concatenated hidden features vector. The superscript L refers to the dimensions of the last layer of each MLP.

and b). The two subnetworks are integrated by combining (concatenating) the output of their last corresponding hidden layer. A downstream MLP is then attached to the concatenated feature vector to effectively learn how to combine the information from both types of omics data. This type of architecture is expected to perform better than the individual networks in isolation, when trained on the same classification task (e.g., [19]).

In this work, we consider MLP as the architecture of each network. It consists of a fixed input dropout layer (with a dropout rate of 10%) and a variable number of fully connected layers with the same number of neurons, followed by a batch normalization layer, ReLU activation function and a dropout layer (Fig. 1). The output layer maps the current number of hidden features into one or three output neurons (one for each class) with a sigmoid or softmax activation function, respectively, according to the dataset-specific classification task.

To build BiBoNet, we first trained the two separated MLPs independently (Fig. 1a and b). Each model’s architecture is optimized through a grid search procedure (see next section) to get a set of best hyperparameters and weights. We then removed the output layers of the MLPs and included them in our late fusion model, each as a different subnetwork. After concatenating the output of the last hidden layer from the two subnetworks, another grid search is performed to determine the optimal architecture of the downstream MLP. BiBoNet’s training

is performed with the only caveat that the two subnetworks’ weights are kept frozen during training. In this way, the resulting model is effectively trained to learn the best parameters’ weights for maximizing the combined information from the two datasets.

4 Results

4.1 Experimental Setup

The performance of BiBoNet is compared to different baselines. For each dataset, we deployed three different model types under two settings: first, we fit each omic data (microbiome and metabolome) independently, and second, we concatenated the input features and fit the same model type. We tested Random Forest (RF), Support Vector Machine (SVM) and MLP. The results of our benchmarks are obtained through a grid search procedure on a five-fold stratified cross-validation scheme. We precompute each fold for metabolome and microbiome data separately for each dataset. The iHMP dataset is a longitudinal study, with multiple samples for the same subject. We use the samples collected at the first visit for each subject as the main data over which the training/validation split for each fold is performed. We then augment the training set of each fold with the remaining samples, ensuring that subjects present in the validation set are not used to augment the training set. Within each fold, the training data are subsequently standardized, and the validation data are rescaled using the mean and standard deviation calculated for the training set. We finally perform class rebalancing of the training data by oversampling the minority class(es) using the SMOTE algorithm [35].

A grid search is performed for each test according to the hyperparameter space reported in Table 2. MLPs are trained with a fixed number of epochs (300) and early stopping with a patience of 50 epochs on the validation loss. We use cross-entropy and binary cross-entropy losses for multiclass and binary tasks, respectively, and the Adam optimizer with default parameters except for learning rate. For each grid search iteration, we compute accuracy (ACC), Matthews correlation coefficient (MCC), F1 score (F1, F1 macro for multiclass tasks) and Area Under the Receiver Operating Characteristic Curve (ROC AUC, ‘one vs rest’ for multiclass tasks) for each fold validation set. We select the best hyperparameters set for each dataset and model type as the one that maximizes the average accuracy over all folds. Ties are broken in favour of the result with the lowest standard deviation (SD) in accuracy over the folds.

All codes are written in Python (v3.10.9). We used `scikit-learn` (v1.4.1) for data and input preparation and for building RF and SVM models. MLPs are custom-implemented using `PyTorch` (v2.2.1).

4.2 Experimental Test

The performance of each model is reported in Fig. 2. The box-and-whisker plots (one for each tested model) depict the distribution of each scoring metric over

Table 2. Hyperparameters space for each model type used in the grid search procedure. For RF and SVM, the hyperparameter names and values correspond to the arguments of the `RandomForestClassifier` and `SVC` objects in `scikit-learn`.

Model	Hyperparameter	Values
Random Forest (RF)	<code>n_estimators</code>	[10, 50, 100, 200, 400, 800, 1000]
	<code>max_depth</code>	[5, 10, 20, 30, 40, 50, 60, 70, 80, 90, <i>None</i>]
	<code>min_samples_leaf</code>	[0.001, 0.005, 0.01, 0.05, 0.1]
Support Vector Machine (SVM)	<code>C</code>	[0.0001, 0.001, 0.01, 0.1, 1, 10, 100]
	<code>gamma</code>	[0.001, 0.01, 0.1, 1, 10, 100, <i>scale</i>]
	<code>kernel_type</code>	[<i>linear</i> , <i>poly</i> , <i>rbf</i> , <i>sigmoid</i>]
Multi-layer Perceptron (MLP)	<code>n_layers</code>	[1, 2, 3]
	<code>n_neurons</code>	[16, 32, 64, 128, 256]
	<code>dropout rate</code>	[0.25, 0.5, 0.75]
	<code>learning rate</code>	[0.001, 0.0001, 0.0005]
	<code>batch size</code>	[4, 8, 16, 32]

all benchmarked datasets. Thirty samples are used in each box and whisker plot, each corresponding to the performance on the validation set of a particular fold. Results are grouped by model type and differentiated between metabolites and microbes alone, a combination of the two (through early fusion, EF) or late fusion (BiBoNet).

First, we note that BiBoNet (red box plots in Fig. 2) performed better than every other tested model, scoring the highest in every considered metric (first, second, and third quartiles are always higher than the rest). In particular, it is the only model that resulted in median scores above 0.8 (ACC), 0.6 (MCC), and 0.8 (ROC AUC). Although F1 score and ROC AUC medians are higher for BiBoNet, the difference with MLP-based models is not substantial. On the contrary, ACC and MCC indicate a sharp increase in performance for BiBoNet. We note that these results are aggregated over strongly imbalanced (both positively and negative) datasets (Table 1). In those cases, MCC is thought to provide a more reliable metric compared to F1 score and ROC AUC, at least for binary classification [36, 37].

When looking at the performance of models based on a single data type, we observe that metabolites-based models (blue box plots in Fig. 2) always perform better than their microbes counterparts (orange box plots in Fig. 2). This is true regardless of the type of model used. In addition, MLP-based models tend to perform better than RF and SVM in any setting. Concatenating the input features (i.e. early fusion, green box plots in Fig. 2) generally provides little or no increase in performance compared to the results obtained on metabolites data alone for RF and SVM models. On the contrary, although the gain in performance is still small, MLP models generally show an increased level of the third quartile compared to the best-performing single-dataset model of the same type. Nevertheless, in all the three baseline model types considered, the combination of input features seems to cause an increase in the overall spread of the results (i.e. variance, Fig. 2a). In some cases, concatenating the input features even degrades the performance (RF and SVM in Fig. 2d). These observations suggest that this

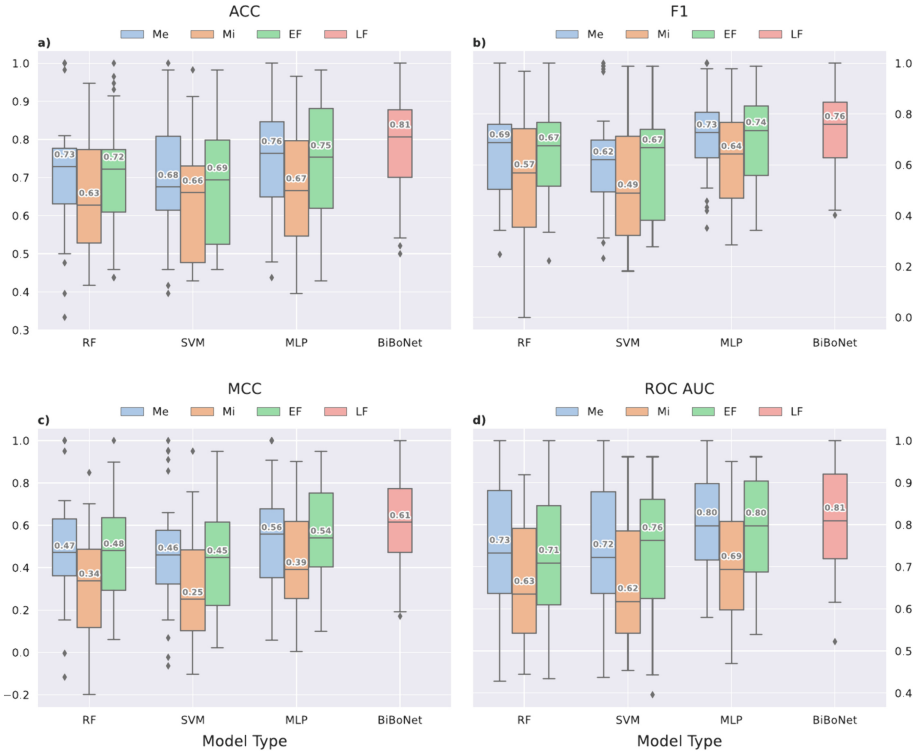


Fig. 2. Aggregated results for each fold over all datasets for four scoring metrics. Results are grouped by model type. Colours indicate results for metabolites alone (Me, blue), microbes alone (Mi, orange), early fusion (EF, green) and late fusion (LF, red). Numerical values in each box-and-whisker plot correspond to the distribution's median (second quartile). (Color figure online)

strategy does not allow for an effective mix of information between the two datasets, with RF and SVM performing worse than MLP models. On the other hand, our late fusion approach reconciles differences between the two datasets more effectively, resulting in higher scoring values than any other tested model and reduced variance compared to early fusion models.

We look more closely at MLP-based models because they have proved to perform better than RF and SVM baselines and constitute the backbone for BiBoNet. Table 3 compiles the aggregated scores (mean and standard deviation) on the five folds, per dataset, for MLP and BiBoNet. Our model performs better than single-omic and early fusion MLP models in four out of six datasets (PRISM, YACHIDA, iHMP, and KIM) with a corresponding increase in accuracy of 5.6%, 3.9%, 15.2% and 6.2% respectively, compared to the best single-MLP model. For the WANG dataset, BiBoNet achieved the same results as the single-data model based on metabolites. SINHA is the only dataset for which we register a decrease in accuracy (-1.8%) when using BiBoNet. Overall, these

Table 3. Five-fold stratified cross-validation performance on each dataset. Each reported value is the mean of that score calculated over the validation sets of the cross-validation procedure. The standard deviation is reported in brackets. Subscript Me refers to metabolites, Mi to microbes and EF (early fusion) to input feature concatenation.

Algorithm		MLP _{Me}	MLP _{Mi}	MLP _{EF}	BiBoNet
Dataset	Score				
iHMP	ACC	0.638 (0.057)	0.562 (0.047)	0.629 (0.122)	0.743 (0.065)
	F1	0.609 (0.077)	0.495 (0.048)	0.608 (0.112)	0.715 (0.087)
IBD 3 classes	MCC	0.442 (0.092)	0.323 (0.079)	0.432 (0.186)	0.611 (0.097)
	ROC AUC	0.792 (0.054)	0.684 (0.084)	0.767 (0.081)	0.795 (0.074)
KIM	ACC	0.517 (0.060)	0.471 (0.049)	0.508 (0.047)	0.550 (0.043)
	F1	0.433 (0.051)	0.406 (0.052)	0.442 (0.064)	0.447 (0.049)
CRC 3 classes	MCC	0.192 (0.098)	0.155 (0.088)	0.206 (0.084)	0.255 (0.091)
	ROC AUC	0.622 (0.036)	0.553 (0.043)	0.581 (0.033)	0.619 (0.053)
PRISM	ACC	0.774 (0.079)	0.755 (0.060)	0.774 (0.079)	0.819 (0.095)
	F1	0.773 (0.086)	0.737 (0.065)	0.780 (0.079)	0.820 (0.086)
IBD 3 classes	MCC	0.661 (0.122)	0.625 (0.096)	0.658 (0.119)	0.725 (0.146)
	ROC AUC	0.893 (0.045)	0.861 (0.052)	0.893 (0.037)	0.921 (0.039)
SINHA	ACC	0.840 (0.016)	0.718 (0.087)	0.855 (0.038)	0.840 (0.029)
	F1	0.708 (0.037)	0.496 (0.151)	0.707 (0.132)	0.693 (0.066)
CRC binary	MCC	0.631 (0.033)	0.314 (0.207)	0.674 (0.108)	0.620 (0.085)
	ROC AUC	0.780 (0.025)	0.645 (0.095)	0.792 (0.090)	0.768 (0.039)
WANG	ACC	0.993 (0.014)	0.934 (0.033)	0.951 (0.029)	0.993 (0.014)
	F1	0.995 (0.009)	0.958 (0.020)	0.969 (0.019)	0.995 (0.009)
ESRD binary	MCC	0.981 (0.038)	0.810 (0.099)	0.862 (0.085)	0.981 (0.038)
	ROC AUC	0.991 (0.019)	0.875 (0.070)	0.912 (0.054)	0.991 (0.019)
YACHIDA	ACC	0.733 (0.047)	0.679 (0.038)	0.707 (0.049)	0.762 (0.044)
	F1	0.756 (0.045)	0.704 (0.046)	0.735 (0.024)	0.782 (0.046)
CRC binary	MCC	0.469 (0.096)	0.356 (0.074)	0.419 (0.105)	0.520 (0.088)
	ROC AUC	0.731 (0.048)	0.676 (0.036)	0.704 (0.060)	0.759 (0.043)

results justify the adoption of our framework and highlight the benefits of using BiBoNet over conventional models.

5 Conclusion and Future Work

We presented BiBoNet, a DL multi-omics model that integrates gut microbiome and metabolome data. We showed that the two data types can be effectively combined using BiBoNet to classify patients under different diseases. Our proposed model leverages the complementary information from both data types and shows improved performances compared to single-data analysis and simple input feature concatenation. Our results highlight the potential of BiBoNet for

applications in precision medicine and system biology. Future work will focus on gaining insights into the microbiome/metabolome biological interactions and improving the current BiBoNet architecture. An explainable AI (XAI, [38]) analysis of BiBoNet will allow us to determine, for example, which particular component of the multi-omics input data drives model decisions and, within that component, which elements (metabolites/microbes) are more important for the final model's predictions. More sophisticated DL architectures (e.g., graph networks) combined with particular learning mechanisms (e.g., attention) can also be implemented to facilitate the model's explainability further and boost performance [23, 39]. This work provides a promising starting point for these future developments.

Acknowledgements. This work is funded by CNR project FOE-2021 DBA. AD005.225.

References

1. Athieniti, E., Spyrou, G.M.: A guide to multi-omics data collection and integration for translational medicine. *Comput. Struct. Biotechnol. J.* **21**, 134–149 (2023)
2. Rappoport, N., Shamir, R.: Multi-omic and multi-view clustering algorithms: review and cancer benchmark. *Nucleic Acids Res.* **46**(20), 10546–10562 (2018)
3. Wu, C., Zhou, F., Ren, J., Li, X., Jiang, Y., Ma, S.: A selective review of multi-level omics data integration using variable selection. *High-throughput* **8**(1), 4 (2019)
4. Subramanian, I., Verma, S., Kumar, S., Jere, A., Anamika, K.: Multi-omics data integration, interpretation, and its application. *Bioinform. Biol. Insights* **14**, 1177932219899051 (2020)
5. Schully, S.D., Khoury, M.J.: What is translational genomics? An expanded research agenda for improving individual and population health. *Appl. Transl. Genom.* **3**(4), 82–83 (2014)
6. Idle, J.R., Gonzalez, F.J.: Metabolomics. *Cell Metab.* **6**(5), 348–351 (2007)
7. Schmidt, J.C., Dougherty, B.V., Beger, R.D., Jones, D.P., Schmidt, M.A., Mattes, W.B.: Metabolomics as a truly translational tool for precision medicine. *Int. J. Toxicol.* **40**(5), 413–426 (2021)
8. Chen, C., et al.: Applications of multi-omics analysis in human diseases. *MedComm* **4**(4), e315 (2023)
9. Zhang, X., Li, L., Butcher, J., Stintzi, A., Figeys, D.: Advancing functional and translational microbiome research using meta-omics approaches. *Microbiome* **7**, 1–12 (2019)
10. Gill, S.R., et al.: Metagenomic analysis of the human distal gut microbiome. *Science* **312**(5778), 1355–1359 (2006)
11. The Human Microbiome Project Consortium: Structure, function and diversity of the healthy human microbiome. *Nature* **486**(7402), 207–214 (2012)
12. Yatsunenko, T., et al.: Human gut microbiome viewed across age and geography. *Nature* **486**(7402), 222–227 (2012)
13. Turnbaugh, P.J., Ridaura, V.K., Faith, J.J., Rey, F.E., Knight, R., Gordon, J.I.: The effect of diet on the human gut microbiome: a metagenomic analysis in humanized gnotobiotic mice. *Sci. Transl. Med.* **1**(6), 6ra14–6ra14 (2009)

14. Turnbaugh, P.J., Ley, R.E., Mahowald, M.A., Magrini, V., Mardis, E.R., Gordon, J.I.: An obesity-associated gut microbiome with increased capacity for energy harvest. *Nature* **444**(7122), 1027–1031 (2006)
15. Leng, D., et al.: A benchmark study of deep learning-based multi-omics data fusion methods for cancer. *Genome Biol.* **23**(1), 171 (2022). <https://doi.org/10.1186/s13059-022-02739-2>
16. Sharifi-Noghabi, H., Zolotareva, O., Collins, C.C., Ester, M.: MOLI: multi-omics late integration with deep neural networks for drug response prediction. *Bioinformatics* **35**(14), i501–i509 (2019). <https://doi.org/10.1093/bioinformatics/btz318>
17. Preuer, K., Lewis, R.P., Hochreiter, S., Bender, A., Bulusu, K.C., Klambauer, G.: DeepSynergy: predicting anti-cancer drug synergy with deep learning. *Bioinformatics* **34**, 1538–1546 (2018). <https://doi.org/10.1093/bioinformatics/btx806>
18. Kuru, H.I., Tastan, O., Cicek, A.E.: MatchMaker: a deep learning framework for drug synergy prediction. *bioRxiv* (2020). <https://doi.org/10.1101/2020.05.24.113241>
19. Lin, Y., Zhang, W., Cao, H., Li, G., Du, W.: Classifying breast cancer subtypes using deep neural networks based on multi-omics data. *Genes* **11**, 1–18 (2020). <https://doi.org/10.3390/genes11080888>
20. Islam, M.M., Huang, S., Ajwad, R., Chi, C., Wang, Y., Hu, P.: An integrative deep learning framework for classifying molecular subtypes of breast cancer. *Comput. Struct. Biotechnol. J.* **18**, 2185–2199 (2020). <https://doi.org/10.1016/j.csbj.2020.08.005>
21. Fu, Y., Xu, J., Tang, Z.: A gene prioritization method based on a swine multi-omics knowledgebase and a deep learning model. *Commun. Biol.* **3** (2020). <https://doi.org/10.1038/s42003-020-01233-4>
22. Wang, T., et al.: MOGONET integrates multi-omics data using graph convolutional networks allowing patient classification and biomarker identification. *Nat. Commun.* **12** (2021). <https://doi.org/10.1038/s41467-021-23774-w>
23. Xing, X., et al.: An interpretable multi-level enhanced graph attention network for disease diagnosis with gene expression data. In: 2021 IEEE International Conference on Bioinformatics and Biomedicine (BIBM), pp. 556–561 (2021). <https://doi.org/10.1109/BIBM52615.2021.9669621>
24. Reiman, D., Layden, B.T., Dai, Y.: MimeNet: exploring microbiome-metabolome relationships using neural networks. *PLoS Comput. Biol.* **17** (2021). <https://doi.org/10.1371/journal.pcbi.1009021>
25. Le, V., Quinn, T.P., Tran, T., Venkatesh, S.: Deep in the bowel: highly interpretable neural encoder-decoder networks predict gut metabolites from gut microbiome. *BMC Genom.* **21** (2020). <https://doi.org/10.1186/s12864-020-6652-7>
26. Wang, T., et al.: Predicting metabolomic profiles from microbial composition through neural ordinary differential equations. *Nat. Mach. Intell.* **5**, 284–293 (2023). <https://doi.org/10.1038/s42256-023-00627-3>
27. Khajeh, T., Reiman, D., Morley, R., Dai, Y.: Integrating microbiome and metabolome data for host disease prediction via deep neural networks. In: 2021 IEEE EMBS International Conference on Biomedical and Health Informatics (BHI), pp. 1–4 (2021). <https://doi.org/10.1109/BHI50953.2021.9508601>
28. Franzosa, E.A., Sirota-Madi, A., Avila-Pacheco, J.: Gut microbiome structure and metabolic activity in inflammatory bowel disease. *Nat. Microbiol.* **4**, 293–305 (2019). <https://doi.org/10.1038/s41564-018-0306-4>
29. Muller, E., Algavi, Y.M., Borenstein, E.: The gut microbiome-metabolome dataset collection: a curated resource for integrative meta-analysis. *Npj Biofilms Microbiomes* **8** (2022)

30. Lloyd-Price, J., Arze, C., Ananthakrishnan, A.N., et al.: Multi-omics of the gut microbial ecosystem in inflammatory bowel diseases. *Nature* **569**, 655–662 (2019). <https://doi.org/10.1038/s41586-019-1237-9>
31. Kim, M., et al.: Fecal metabolomic signatures in colorectal adenoma patients are associated with gut microbiota and early events of colorectal cancer pathogenesis. *mBio* **11** (2020). <https://doi.org/10.1128/mBio.03186-19>
32. Sinha, R., et al.: Fecal microbiota, fecal metabolome, and colorectal cancer inter-relations. *PLoS ONE* **11** (2016). <https://doi.org/10.1371/journal.pone.0152126>
33. Wang, X., Yang, S., Li, S., et al.: Aberrant gut microbiota alters host metabolome and impacts renal failure in humans and rodents. *Gut* **69**, 2131–2142 (2020). <https://doi.org/10.1136/gutjnl-2019-319766>
34. Yachida, S., Mizutani, S., Shiroma, H., et al.: Metagenomic and metabolomic analyses reveal distinct stage-specific phenotypes of the gut microbiota in colorectal cancer. *Nat. Med.* **25**, 968–976 (2019). <https://doi.org/10.1038/s41591-019-0458-7>
35. Chawla, N.V., Bowyer, K.W., Hall, L.O., Kegelmeyer, W.P.: SMOTE: synthetic minority over-sampling technique. *J. Artif. Intell. Res.* **16**, 321–357 (2002). <https://doi.org/10.1613/jair.953>
36. Chicco, D., Jurman, G.: The advantages of the Matthews correlation coefficient (MCC) over F1 score and accuracy in binary classification evaluation. *BMC Genom.* **21**, 6 (2020). <https://doi.org/10.1186/s12864-019-6413-7>
37. Chicco, D., Jurman, G.: The Matthews correlation coefficient (MCC) should replace the ROC AUC as the standard metric for assessing binary classification. *BioData Mining* **16**, 4 (2023). <https://doi.org/10.1186/s13040-023-00322-4>
38. Saranya, A., Subhashini, R.: A systematic review of explainable artificial intelligence models and applications: recent developments and future trends. *Decis. Anal. J.* **7**, 100230 (2023). <https://doi.org/10.1016/j.dajour.2023.100230>
39. Kayikci, S., Khoshgoftaar, T.M.: Breast cancer prediction using gated attentive multimodal deep learning. *J. Big Data* **10** (2023). <https://doi.org/10.1186/s40537-023-00749-w>

Open Access This chapter is licensed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license and indicate if changes were made.

The images or other third party material in this chapter are included in the chapter's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the chapter's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder.

