UC San Diego
UC San Diego Previously Published Works

Title

From Genotype to Phenotype: Augmenting Deep Learning with Networks and Systems Biology.

Permalink

Authors

Gazestani, Vahid

Lewis, Nathan

Publication Date

DOI

Peer reviewed

From Genotype to Phenotype: Augmenting Deep Learning with Networks and Systems Biology

Vahid H. Gazestani[1,2,3], **Nathan E. Lewis**[1,3,4,*]

[1]Department of Pediatrics, University of California San Diego, La Jolla, CA 92093, USA

[2]Autism Center of Excellence, Department of Neuroscience, University of California San Diego, La Jolla, CA 92093, USA

[3]Novo Nordisk Foundation Center for Biosustainability, University of California San Diego, La Jolla, CA 92093, USA

[4]Department of Bioengineering, University of California San Diego, La Jolla, CA 92093, USA

Abstract

Cells, as complex systems, consist of diverse interacting biomolecules arranged in dynamic hierarchical modules. Recent advances in deep learning methods now allow one to encode this rich existing knowledge in the architecture of the learning procedure, thus providing the models with the knowledge that is absent in the training data. By encoding biological networks in the architecture, one can develop flexible deep models that propagate information through the molecular networks to successfully classify cell states. Moreover, this flexibility in the architecture can be harnessed to model the hierarchical structure of real biological systems, efficiently converting gene-level data to pathway-level information with an ultimate impact on cell phenotype. Furthermore, such models could require fewer training samples, are more generalizable across diverse biological contexts, and can make predictions that are more consistent with the current understanding on the inner-working of biological systems.

Introduction

Biological networks are widely used to combine and contextualize knowledge across diverse domains from genetics to transcriptomics, proteomics and biochemistry and from small-scale to large-scale experiments (Figure 1a) [1–4]. These enable one to map the dynamics of molecular subsystems over time and under various conditions to identify their contribution to the overall behavior of a cell [4–7]. In the era of big data, where massive high-throughput molecular data are accumulating rapidly, machine learning has emerged as a powerful toolbox for extracting systems-level information from high-dimensional data and their integration with prior knowledge.

[*]Correspondence to Nathan E. Lewis (nlewisres@ucsd.edu).

Recent advances in deep neural networks (DNNs) are revolutionizing the field[8–14]. Deep learning generalizes other machine learning approaches by enabling the design of flexible architectures that 1) streamline feature representation and classification and 2) incorporate domain knowledge into their learning procedure. DNNs are constructed by stacking multiple layers of artificial neurons (Figure 1b) [15–17]. Each layer receives and processes data from the previous layers and sends it to the next layer, while the last layer is tasked with the data classification. In this design, the intermediate hidden layers learn how to represent the data to make different class labels separable for the classification layer[15]. By adjusting the number of intermediate layers, one changes a DNN model's flexibility and its ability to capture complex structures, while avoiding strong assumptions about the data[15,18]. The other power of DNNs arises from their flexibility in the connection patterns of artificial neurons between layers. This allows one to efficiently encode domain knowledge (potentially absent in input data) into the architecture of the learning machinery. For example, convolutional architectures are designed to capture patterns that re-occur with modest variations in different locations such as edges and lines in images[15,19] or cis-regulatory motifs in DNA sequences[20] (for a comprehensive review on the technical aspects of deep learning architectures see [15,16]). As we summarize recent advances below, the towering flexibility of DNN architectures allows one to encode knowledge of biological systems into DNN models, thus enabling new applications of machine learning approaches in networks and systems biology.

Deep learning and identification of active subsystems

Genes involved in the same biological process or phenotype often form connected subnetworks[5,21]. By exploiting this property, network propagation algorithms flow information across genes through molecular networks (i.e., genes connected by interactions such as protein-protein interactions, genetic interactions, gene co-expression, etc.) to identify active subsystems and discover their other gene members. These approaches are useful for functional annotation of genes, drug target identification, and inference of genes associated with complex diseases[5,21,22]. A variety of propagation algorithms have been developed by making different assumptions about the flow of information on the molecular networks. However, a more flexible propagation framework would allow the adoption of information flow dynamics based on the properties of the underlying network (e.g., network size, interaction density or interaction type) and the connection pattern of molecules involved in the subsystem of interest (e.g., genes involved in a detailed biological process versus those participating in cancer progression).

Recent advances enable network propagation in the context of neural networks[23–27]. For example, in a simple yet effective DNN architecture, each gene is connected to itself and its interacting partners from the previous layer (Figure 1c) [23]. Therefore, at each propagation layer, genes receive information from their neighbors while retaining some of their own. This parallels the traditional network propagation techniques where signals shared among interacting neighbors are amplified, while suppressing noise and false positives in the input data. However, the DNN model has more flexibility in three ways. First, although the structure of propagation layers is dictated by the existing network level information, the coefficients (i.e., parameters) controlling the network propagation is automatically learned

from input training data and the topology of the encoding network. Second, by varying the number of stacked propagation layers, one can obtain network propagation machineries of different flexibility. The number of propagation layers is effectively a model hyperparameter and can be optimized by well-established techniques such as cross validation. Third, one can easily instruct the learning algorithm to assign different propagation coefficients based on the amount of available training data. More specifically, the model can consider different propagation coefficients for different genes or the propagation coefficients can be tied to one another based on the interaction type. For example, all gene regulatory interactions could share the same coefficient, while different than those of physical binding interactions. GraphSAGE presents an alternative framework wherein a DNN model is trained to predict a node role based on the functions of its first and second degree neighbors in the network[27]. Under this framework, one can train a GraphSAGE DNN model on well annotated networks and apply it on other networks, a task that is not feasible by traditional network propagation algorithms. Authors indeed explored this by training a GraphSAGE DNN model on some tissue specific gene networks and showed that the trained DNN model can predict the functions of genes in networks from other tissues[27].

The other advantage of formulating the network propagation in the context of neural networks is that one could readily add other types of artificial layers on top of the propagation layers to streamline the processes of network propagation with other tasks such as cluster identification or disease classification. Lin et al. explored the advantages of such design by linking propagation layers with the classification layers to identify the cell type and cell state, based on the single cell expression data[28]. In this design, the propagation layers amplify the network-level dysregulations by aggregating the relative expression data of each gene with its network neighbors, empowering the classification layers to make accurate predictions based on expression data from one experiment.

Deep learning and modeling biological systems

DNNs are a natural choice for representing the hierarchical structure of biological systems where a genotype ultimately influences biological processes and phenotypes [14,29,30]. In DNNs, each neuron receives and combines information from the connected neurons in the preceding layer in a nonlinear way. As data pass through the layers, it can undergo a gradual abstraction (Figure 1d). To explore this concept, Ma et al. constructed a DNN architecture, called DCell, that encodes the multi-scale structure of a Gene Ontology tree, where several genes cooperate to perform a detailed biological process. They then combine with one another in the following layers to deliver bigger processes with an ultimate impact on the phenotype[14]. In this design, in contrast to propagation layers that amplify the network level dysregulations, DCell reduces the dimensionality of input data from genotype to Gene Ontology terms to phenotype. Indeed, DCell successfully predicted the effect of gene mutations on cell proliferation in yeast[14].

One limitation of DNN models with dense intermediate layers is that their inner working is a black box; i.e., while the model performs well in predicting the output given the input, the meaning and function of neurons in the intermediate layers are not clear. A major benefit of biologically inspired architectures such as DCell is that neurons in the intermediate layers

are more interpretable since they are engineered to represent specific Gene Ontology terms. Therefore, the trained model potentially could provide mechanistic insights on the inner working of the biological systems. However, care should be taken in interpreting the results. Although the architecture of DNNs can be engineered to be similar to biological systems, there is no guarantee that the inner-working of the two will be consistent. It is possible that physicochemical laws governing biological systems do not follow the same rules as the mathematically derived systems in the DNNs. This is especially true when DNNs are trained based on the input (e.g., gene mutations) and output (e.g., cell phenotype) of the biological system and the model is blind to "intermediate-phenotypes" of the biological system including mRNA and protein expression or cell metabolism. Consequently, follow-up experiments are needed to establish biological relevance of mechanistic interpretations from such models. A recent study by Wang et al. took initial steps in addressing this issue by integrating data on intermediate-phenotypes in the learning procedure[29]. Briefly, Wang et al. constructed a global gene regulatory network by generating data on chromatin structure and gene expression levels and then combined it with nucleotide binding preferences of gene regulators from databases[29]. Authors next encoded the hierarchical structure of the constructed gene regulatory network in a DNN model to connect genotype data with intermediate-phenotypes including gene expression, tissue cell type composition, and enhancer activity as indicators of status of inner-working subsystems. Wang et al. demonstrated that incorporation of intermediate-phenotypes indeed boosts the performance of DNNs in predicting the disease status[29].

Recent studies demonstrate that incorporating vast amounts of biological knowledge in the learning procedure lowers the number of model parameters, up to an order of magnitude (Figure 1), while preserving or even improving the model performance, compared to DNN models without biologically inspired architectures [14,28]. Therefore, the engineering of the architecture allows development of learning procedures that require fewer training samples and are still capable of capturing complex nonlinear relationships among biological subsystems. One challenge in designing such architectures is the missing information. There are biological processes and interactions that are yet to be discovered but important for the proper functioning of biological systems. Such missing gaps might impede the performance of the DNNs or severely influence the interpretations from their inner workings. Recent studies have approached this issue in different ways. Ma et al. supplemented the current state of knowledge on the Gene Ontology annotations by computational predictions[14]. Wang et al. based the architecture on a gene regulatory network that was systematically constructed by genome-wide data[29]. In an alternate approach, Lin et al. incorporated additional artificial neurons in the design of intermediate layers to provide the model with the required flexibility to account for gene interactions that are unknown yet[28].

Deep learning model generalizability, transferability and interpretability

One primary obstacle in inference from high-throughput molecular data is the curse of dimensionality[31,32]; the number of input features far exceeds the number of available samples in most biological problems. Biological systems are intricate, composed of numerous subsystems with a hierarchical structure and complex interactions (e.g., synergy, redundancy, and competition among RNA binding proteins and miRNAs or between

enzymes). However, there is an upper limit on the number of samples one can obtain from a given condition due to technological (e.g., instrument resolution, associated cost) or fundamental biological reasons (e.g., limits on cohort size with a given condition, or the inherent correlation structures among SNPs). In these conditions, spurious correlation structures arise between features, thus hindering the accuracy of the inferences and predictions. One common solution is to use data-driven feature reduction techniques such as principal component analysis (PCA)[33]. However, these approaches reduce the dimension by aggregating true signal with the correlating confounding factors. Another approach is gene set enrichment analysis (GSEA)[34] and related approaches[35] wherein data are converted from gene to pathway level, thereby reducing the number of features. However, in addition to having low flexibility in the way that gene set activity is defined, these knowledge-driven approaches commonly ignore the hierarchical structure of biological subsystems. Consequently, the activity of subsystems at different levels of hierarchy are predicted independently based on only the gene level data.

Engineering DNN architectures would help distinguish and disentangle these false correlation structures that exist in data by identifying processes that involve distinct subsystems but show parallel responses in high-throughput data. In this context, similar to GSEA, initial layers reduce feature dimension by aggregating gene data to pathway level data. However, in addition to being more flexible, the DNN model incorporates the hierarchical structure of biological systems. Therefore, such models would learn feature abstraction rules that are better generalizable across different biological conditions while avoiding issues related to the curse of dimensionality[31,36,37]. This increase in generalizability allows the application of transfer learning techniques where in models are pre-trained on broader contexts and are next fine-tuned based on context specific datasets with fewer available samples. As a hypothetical example, consider a classification task on a human tissue (e.g., disease versus normal) based on gene expression data. As discussed above, there is a limit on the number of samples one can obtain from a specific tissue with a specific condition. However, in the broader context, there already exists a large repository of human transcriptome datasets from different tissues and under different biological conditions. One can develop an encoder-decoder DNN model that exploits this vast amount of existing data to learn how to reduce the dimension of gene level data in an unsupervised manner. If the architecture of the encoder part of the model is inspired by the connection patterns of biological subsystems, such a model would learn how to abstract gene expression to pathway level data at different abstraction levels. This encoder model, trained on all available human expression data, can be next coupled with the classification layers that can be trained by far fewer samples for the classification of tissue of interest.

DNN architectures that are designed to resemble biological systems could make predictions that are consistent with the knowledge on the hierarchical structure of biological systems. For example, Kulmanov et al. used a DNN model to predict protein functions based on their sequence and network context[38]. By encoding the structure of the Gene Ontology tree in the classification layers, Kulmanov et al. enforced the learning procedure to predict functions that are consistent with the hierarchical structure of the tree[38]. As elaborated above, the engineering the architecture of DNN models using biological knowledge increases the interpretability of the inner working of the DNN models as well.

Conclusion and future perspectives

In different biological conditions, hundreds to thousands of biological subsystems work together, in parallel, and in a hierarchy to adapt the system and tune its response. It is impossible to capture this complexity by a single dataset or a single type of data[30,31]. Flexibility of DNN architectures allow us to go beyond a general-purpose learning algorithm and encode the extensive, existing network- and systems-level knowledge that is generated by combing diverse data types. Such design would inform the model on aspects of the biological systems that are important for making accurate predictions but are not available in the input data.

Although powerful, there remain major challenges in the application of DNNs in the field. These include the need for large training datasets, missing biological knowledge on the biological systems, and interpretability of results. Throughout the text, we discussed possible solutions for each of these issues. For example, DNNs would require fewer samples if their architectures are inspired by biological systems. These architectures are also expected to have better generalizability across different biological conditions, making them amenable for transfer learning[39] or the conceptually similar technique of multi-task learning[40]. The missing biological knowledge can be also resolved by complementing biological knowledge with systematic genome-wide analyses or considering such uncertainties in the architecture of the DNN model. On the interpretability of DNN models, effective methods have been also developed to understand input characteristics that contribute to the outcome of these models[20,41]. Designing architectures that mirror those of biological systems could also enhance the interpretability of the intermediate hidden layers in DNN models.

Another challenge is in the design of high-performance DNN architectures that can extract meaningful patterns from encoded biological networks. First, current DNN architectures are designed primarily for regular graphs where each node has a fixed number of interactions (e.g., in an image, wherein each pixel has a fixed number of neighboring pixels). However, the number of interactions vary between nodes in biological networks and an ideal architecture could extract patterns from such complex structures. Encoding the Gene Ontology hierarchy where genes involved in similar processes are grouped together provides initial steps towards such designs[14]. Second, biological networks are complex, heterogenous, and context dependent. While the interactions between neighboring pixels in images represent a fixed Euclidian distance, the interactions within biological networks could have diverse interpretations, such as physical binding interactions, molecular similarity, or gene regulation. These interactions could also be from physical measurements, functional relationships, or computational predictions. Moreover, interactions in biological networks do not represent a static entity and rewire significantly under different conditions. Several endeavors are currently underway to design efficient architectures and we expect they will be pivotal in the success of deep learning methods in the field of network biology[24,41,42].

Acknowledgements

References

1. Barabási A-L, Gulbahce N, Loscalzo J: Network medicine: a network-based approach to human disease. Nature reviews genetics 2011, 12:56.

2. Vidal M, Cusick ME, Barabási A-L: Interactome networks and human disease. Cell 2011, 144:986–998. [PubMed: 21414488]

3. Wang B, Mezlini AM, Demir F, Fiume M, Tu Z, Brudno M, Haibe-Kains B, Goldenberg A: Similarity network fusion for aggregating data types on a genomic scale. Nat Methods 2014, 11:333–337. [PubMed: 24464287]

4. Kitano H: Systems biology: a brief overview. Science 2002, 295:1662–1664. [PubMed: 11872829]

5. Cowen L, Ideker T, Raphael BJ, Sharan R: Network propagation: a universal amplifier of genetic associations. Nat Rev Genet 2017, 18:551–562. [PubMed: 28607512]

6. Eduati F, Mangravite LM, Wang T, Tang H, Bare JC, Huang R, Norman T, Kellen M, Menden MP, Yang J, et al.: Prediction of human population responses to toxic compounds by a collaborative competition. Nat Biotechnol 2015, 33:933–940. [PubMed: 26258538]

7. Gazestani VH, Pramparo T, Nalabolu S, Kellman BP, Murry S, Lopez L, Pierce K, Courchesne E, Lewis NE: Transcriptional organization of autism spectrum disorder and its connection to ASD risk genes and phenotypic variation. bioRxiv 2018:435917.

8. Angermueller C, Pärnamaa T, Parts L, Stegle O: Deep learning for computational biology. Molecular systems biology 2016, 12:878. [PubMed: 27474269]

9. Wainberg M, Merico D, Delong A, Frey BJ: Deep learning in biomedicine. Nature Biotechnology 2018, 36:829.

10. Miotto R, Wang F, Wang S, Jiang X, Dudley JT: Deep learning for healthcare: review, opportunities and challenges. Briefings in bioinformatics 2017.

11. Min S, Lee B, Yoon S: Deep learning in bioinformatics. Briefings in bioinformatics 2017, 18:851–869. [PubMed: 27473064]

12. Mamoshina P, Vieira A, Putin E, Zhavoronkov A: Applications of deep learning in biomedicine. Molecular pharmaceutics 2016, 13:1445–1454. [PubMed: 27007977]

13. Ching T, Himmelstein DS, Beaulieu-Jones BK, Kalinin AA, Do BT, Way GP, Ferrero E, Agapow P-M, Zietz M, Hoffman MM: Opportunities and obstacles for deep learning in biology and medicine. Journal of The Royal Society Interface 2018, 15:20170387.

14. Ma J, Yu MK, Fong S, Ono K, Sage E, Demchak B, Sharan R, Ideker T: Using deep learning to model the hierarchical structure and function of a cell. Nat Methods 2018, 15:290–298. [PubMed: 29505029] **This paper encodes the hierarchical structure of gene ontology tree in a DNN architecture to sucessfully predict effect of gene mutations on cell proliferation.

15. LeCun Y, Bengio Y, Hinton G: Deep learning. Nature 2015, 521:436–444. [PubMed: 26017442] **This paper provides intuitive technical insights on the DNN models.

16. Goodfellow I, Bengio Y, Courville A, Bengio Y: Deep learning, vol 1: MIT press Cambridge; 2016.

17. Bengio Y: Learning deep architectures for AI. Trends in Machine Learning 2009, 2:1–127.

18. Hornik K, Stinchcombe M, White H: Universal approximation of an unknown mapping and its derivatives using multilayer feedforward networks. Neural networks 1990, 3:551–560.

19. LeCun Y, Boser BE, Denker JS, Henderson D, Howard RE, Hubbard WE, Jackel LD: Handwritten digit recognition with a back-propagation network.In Advances in neural information processing systems: 1990:396–404.

20. Alipanahi B, Delong A, Weirauch MT, Frey BJ: Predicting the sequence specificities of DNA-and RNA-binding proteins by deep learning. Nat Biotechnol 2015, 33:831–838. [PubMed: 26213851]

21. Sharan R, Ulitsky I, Shamir R: Network-based prediction of protein function. Molecular systems biology 2007, 3:88. [PubMed: 17353930]

22. Mostafavi S, Ray D, Warde-Farley D, Grouios C, Morris Q: GeneMANIA: a real-time multiple association network integration algorithm for predicting gene function. Genome biology 2008, 9:S4.

23. Kipf TN, Welling M: Semi-supervised classification with graph convolutional networks. arXiv 2016.*This paper introduces a simple, yet effective approach for encoding network structure in the architecture of DNNs.

24. Defferrard M, Bresson X, Vandergheynst P: Convolutional neural networks on graphs with fast localized spectral filtering. In Advances in Neural Information Processing Systems: 2016:3844–3852.

25. Scarselli F, Gori M, Tsoi AC, Hagenbuchner M, Monfardini G: The graph neural network model. IEEE Transactions on Neural Networks 2009, 20:61–80. [PubMed: 19068426]

26. Gori M, Monfardini G, Scarselli F: A new model for learning in graph domains. In Neural Networks, 2005. IJCNN'05. Proceedings. 2005 IEEE International Joint Conference on: IEEE: 2005:729–734.

27. Hamilton W, Ying Z, Leskovec J: Inductive representation learning on large graphs. In Advances in Neural Information Processing Systems: 2017:1024–1034.*This paper introduces GraphSage in which a DNN is trained to predict the role of a gene based on its network neighbors

28. Lin C, Jain S, Kim H, Bar-Joseph Z: Using neural networks for reducing the dimensions of single-cell RNA-Seq data. Nucleic Acids Res 2017, 45:e156.**This paper encodes the netowrk propagation layers in a DNN architecture to accurately predict a cell type and its state based on gene expression data.

29. Wang D, Liu S, Warrell J, Won H, Shi X, Navarro FCP, Clarke D, Gu M, Emani P, Yang YT, et al.: Comprehensive functional genomic resource and integrative model for the human brain. Science 2018, 362.**This paper encodes a gene regulatory netowrk in the architecture of DNN to incorporate intermediate phenotypes (e.g., gene expression, tissue composition) in the learning procedure, resulting in an increased accuracy in predicting tissue status from genetic information.

30. Csete ME, Doyle JC: Reverse engineering of biological complexity. Science 2002, 295:1664–1669. [PubMed: 11872830]

31. Brenner S: Sequences and consequences. Philosophical Transactions of the Royal Society B: Biological Sciences 2010, 365:207–212.

32. Donoho DL: High-dimensional data analysis: The curses and blessings of dimensionality. AMS math challenges lecture 2000, 1:32.

33. Meng C, Zeleznik OA, Thallinger GG, Kuster B, Gholami AM, Culhane AC: Dimension reduction techniques for the integrative analysis of multi-omics data. Briefings in bioinformatics 2016, 17:628–641. [PubMed: 26969681]

34. Subramanian A, Tamayo P, Mootha VK, Mukherjee S, Ebert BL, Gillette MA, Paulovich A, Pomeroy SL, Golub TR, Lander ES, et al.: Gene set enrichment analysis: a knowledge-based approach for interpreting genome-wide expression profiles. Proc Natl Acad Sci U S A 2005, 102:15545–15550. [PubMed: 16199517]

35. Khatri P, Sirota M, Butte AJ: Ten years of pathway analysis: current approaches and outstanding challenges. PLoS computational biology 2012, 8:e1002375. [PubMed: 22383865]

36. Poggio T, Mhaskar H, Rosasco L, Miranda B, Liao Q: Why and when can deep-but not shallow-networks avoid the curse of dimensionality: a review. International Journal of Automation Computing 2017, 14:503–519.

37. Mhaskar H, Liao Q, Poggio TA: When and why are deep networks better than shallow ones? In AAAI: 2017:2343–2349.

38. Kulmanov M, Khan MA, Hoehndorf R: DeepGO: predicting protein functions from sequence and interactions using a deep ontology-aware classifier. Bioinformatics 2017, 34:660–668.

39. Weiss K, Khoshgoftaar TM, Wang D: A survey of transfer learning. Journal of Big Data 2016, 3:9.

40. Ruder S: An overview of multi-task learning in deep neural networks. arXiv 2017.

41. Shrikumar A, Greenside P, Kundaje A: Learning important features through propagating activation differences. arXiv 2017.

42. Chang S, Han W, Tang J, Qi G-J, Aggarwal CC, Huang TS: Heterogeneous network embedding via deep architectures. In Proceedings of the 21th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining: ACM: 2015:119–128.
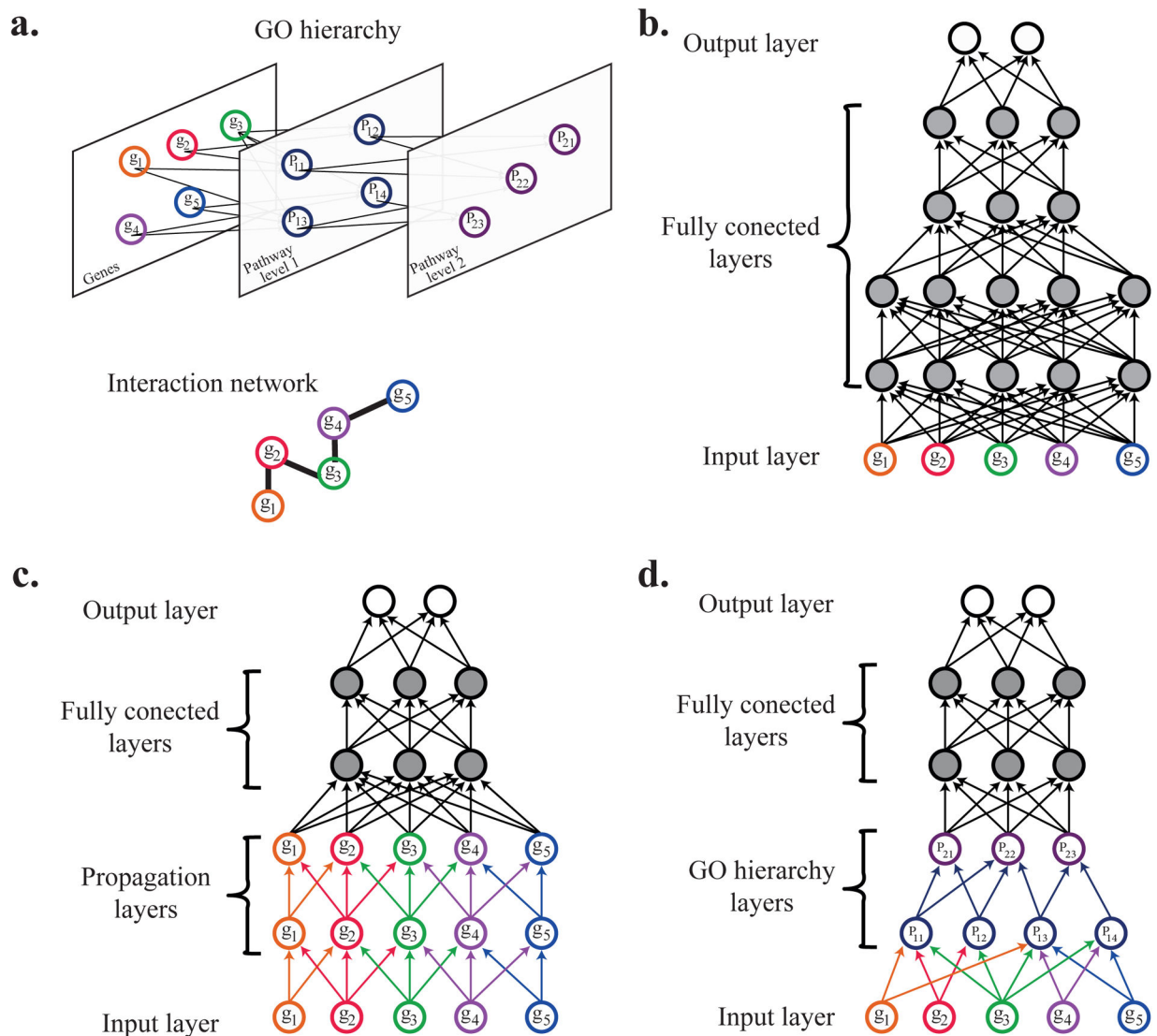
Figure 1. Encoding network- and systems-level biological knowledge in the architecture of DNN models.

(A) Gene Ontology (GO) hierarchy and gene interaction networks are two rich resources on complex structure of biological systems. (B) Schematic representation of a fully connected DNN architecture with four intermediate hidden layers. In this architecture, each neuron is connected to every neuron in the previous layer. (C) Encoding the biological network represented in panel A in the architecture of DNN models. The first two intermediate layers are propagation layers. These layers are designed to amplify network-level dysregulation and suppress potential noise. The last two intermediate layers receive information from the final propagation layer and modify them nonlinearly so that the classes are separable in the output layer. Note that the propagation layers do not change the dimension of input feature space; i.e., for each gene in the input feature, there exists one corresponding gene in the propagation layer. Therefore, although they amplify the network-level signals, they are followed by additional layers to reduce the dimension of input data (D) The hierarchical structure of biological systems can be encoded into the architecture of DNN models. The

architecture of first two intermediate layers is based on the Gene Ontology tree represented in panel A. As illustrated, data undergo gradual abstractions through these layers consistent with the current knowledge on biological systems. Note that in contrast to propagation layers, Gene Ontology based layers reduce the dimensionality of input features from the gene level to the pathway level. Moreover, genes can be involved in more than one pathway. Likewise, pathways can intersect in their gene composition. Additionally, pathways in the second Gene Ontology-based layer (i.e., P_{21}, P_{22}, and P_{23}) receive data from the preceding Gene Ontology based layer (i.e., P_{11}, P_{12}, P_{13}, and P_{14}), rather than the gene level data that is in the input layer. This is in contrast to the traditional approaches on measuring pathway level activity such as GSEA. In all four panels, g_i represents the gene level information. P_{ij} represents *Pathway j* in layer *i*.