scientific reports

OPEN

Drug discovery and mechanism prediction with explainable graph neural networks

Conghao Wang, Gaurav Asok Kumar & Jagath C. Rajapakse

Apprehension of drug action mechanism is paramount for drug response prediction and precision medicine. The unprecedented development of machine learning and deep learning algorithms has expedited the drug response prediction research. However, existing methods mainly focus on forward encoding of drugs, which is to obtain an accurate prediction of the response levels, but omitted to decipher the reaction mechanism between drug molecules and genes. We propose the eXplainable Graph-based Drug response Prediction (XGDP) approach that achieves a precise drug response prediction and reveals the comprehensive mechanism of action between drugs and their targets. XGDP represents drugs with molecular graphs, which naturally preserve the structural information of molecules and a Graph Neural Network module is applied to learn the latent features of molecules. Gene expression data from cancer cell lines are incorporated and processed by a Convolutional Neural Network module. A couple of deep learning attribution algorithms are leveraged to interpret interactions between drug molecular features and genes. We demonstrate that XGDP not only enhances the prediction accuracy compared to pioneering works but is also capable of capturing the salient functional groups of drugs and interactions with significant genes of cancer cells.

Aiming at facilitating precision medicine in complex disease such as cancer, computational approaches have been increasingly proposed to delve into the reactions between drugs and cancer cells[1]. Recently, numerous machine learning[2,3] and deep learning[4,5] methods have been successfully applied to predict drug response levels precisely. However, most of them target at phenotypic screening[6] and do not come along with a reasonable interpretability, rendering drug reaction mechanism obscure. To expedite precision medicine, it is crucial to elucidate the mechanism of action of drugs and thereby promote novel drug discovery.

A proper representation of a drug molecule is pivotal to any drug response prediction methods. According to recent reviews of molecular representations of drugs[7], there are mainly three categories of representation: linear notations, molecular fingerprints (FPs), and graph notations. Linear notations encode the molecule with a vector of string. Two frequently used instances of linear notations are the IUPAC International Chemical Identifier (InChI)[8], and the Simplified Molecular-Input Line-Entry System (SMILES)[9]. SMILES strings are more widely used since it encodes the chemical structure into a string of ASCII characters. CaDRReS[10] applied Matrix Factorization to learn the latent features of drugs with the cell line gene expression data and drug sensitivity matrix, and compared the similarity scores derived from learned features and SMILES notations. tCNNs[11] and CDRScan[12] adopted Convolutional Neural Networks (CNN) to learn a latent representation of drugs' SMILES vector. CNN is a powerful deep learning approach to handle grid-like data in the domain of texts and images, which can be used to encode the linear notations of drugs as well. However, the SMILES notation does not possess the property of locality like texts and images since the physically adjacent atoms in the sequence of SMILES string can be far away from each other in the real molecular environment, and therefore dimisses the structural information of molecules.

Molecular fingerprints, such as Molecular Access System (MACCS)[13] and Chemically Advanced Template Search[14], identify the key structures of a molecule and represent them with a binary vector where each bit denotes the structure's existence. A drawback of this kind of representation is that only the pre-defined structure can be recognized, which might hamper the discovery of novel structures. To circumvent this problem, circular fingerprints such as Extended Connectivity FPs (ECFPs) based on Morgan algorithm[15] has been proposed to iteratively search the substructures of molecules rather than pre-define them. The information of these crucial structures is preserved in this kind of representation, whereas the positional information is lost, and we can hardly track where these sub-structures occur in the molecule. DeepDSC[16] combines Morgan fingerprints of drugs into the latent features of cancer cell lines learned by an auto-encoder. S2DV[17] applied word2vec[18] to

College of Computing and Data Science, Nanyang Technological University, Singapore 639798, Singapore. email: ASJagath@ntu.edu.sg

tokenize ECFP features or SIMLES as drugs' representations. Ma et al. used Atom Pairs (AP), MACCS and circular fingerprints as the descriptor of drugs, and performed the quantitative structure activity relationship (QSAR) study with a Deep Neural Network[19].

Graph notations have recently been brought under the spotlight in the domain of drug representation. Previously, compromising on computational complexity of molecular structures and the confined power of graph learning, aforementioned methods are preferred to denote a molecule even at the cost of loss of information. However, with the advent of Graph Neural Networks (GNN) in the deep learning domain in recent years, it is now feasible to store and analyze the information from molecules in graphs[20].

Numerous variants of GNN models have been applied in the pharmaceutical domain[21,22] and demonstrated to learn the latent representation of the molecular graphs trading off the descriptive power against complexity. A Graph Convolution Network (GCN) model[23] was proposed to predict the chemical properties of molecules and discover porous materials. The typical message passing pattern of GNN intrinsically weakens the influence of distal nodes, which might contradict the real case in the molecule, where atoms from a long topological distance can still interact such as intramolecular hydrogen bonds. An Attentive FP model proposed by[24] leveraged the graph attention mechanism to learn the impact of a node to another. This model addressed the above issue by updating the nodes with a trade-off between the topological distance and the possibly intangible linkage with the attention mechanism. GraphDRP[25] enhanced the tCNN[11] prediction precision by substituting the drug-CNN module with GNN to better encapsulate the drug features. DeepCDR[26], TGSA[27] and DualGCN[28] further explored integrating multi-omics profiles for a better representation of cancer cell lines. Besides modeling drugs with GNN, SWNet[29] introduced a self-attention mechanism to bring drug similarity into the consideration when learning cell features. An algebraic graph-assisted bidirectional transformer (AGBT) model[30] was developed to encode the 3D structure of molecules into algebraic graphs. And Molecular Topographic Map (MTM) was generated from atom features by using Generative Topographic Mapping (GTM)[31] to represent drugs in graphs[32].

In this study, we propose a framework named eXplainable Graph-based Drug response Prediction (XGDP) for predicting anti-cancer drug responses and discovering the mechanism of action. The architecture of XGDP, as shown in Fig. 1, is composed of 3 modules. The GNN module learns the latent features of drugs denoted by molecular graphs. We propose to use a set of novel features adapted from ECFPs as the node features and incorporate chemical bond types as the edge features in our graph convolutional layers. And the CNN module learns the latent features of cancer cell lines from its gene expression profiles. Then, a cross-attention module is utilized to integrate latend features from drugs and cell lines, and thereafter predict the drug responses. The experimental results indicate that, with novel node and edge features, our model outperformed the previous drug response prediction methods[11,25]. Moreover, we leverage deep learning attribution approaches such as GNNExplainer[33] and Integrated Gradients[34] to interpret our model. It is demonstrated that our developed model is capable of identifying the active substructures of drugs and the significant genes in cancer cells, and thus revealing the mechanism of action of drugs.

Methods

Datasets

We propose a deep learning-based approach to predict the drug responses of cancer with molecular graphs of drugs and gene expression data from cancer cell lines. The dataset was acquired from Genomics of Drug

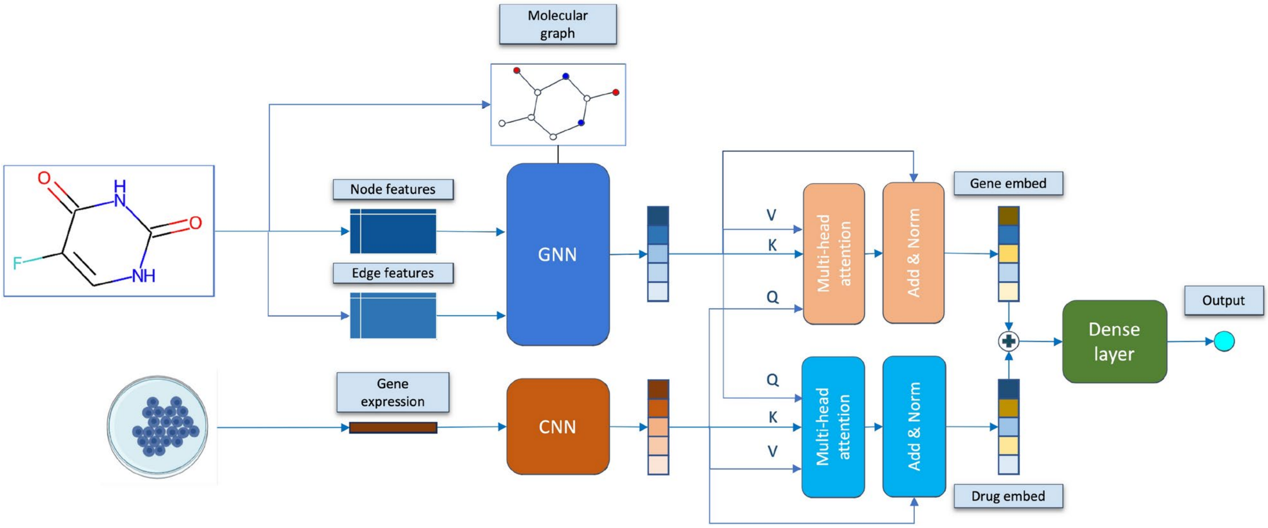

Fig. 1. The architecture of the proposed model XGDP for drug response and mechanism prediction. Molecular graph, node features and edge features are extracted from the drug molecule, and GNN is used for learning the latent features of drugs. CNN is applied to compress the gene expression features from cancer cell lines. Then two multi-head cross-attention layers are leveraged to combine drug and cell features, and the drug response is predicted with the integrated features.

Sensitivity in Cancer (GDSC) database[35], including response levels in IC50 formats, drug names, and cell line names. Gene expression data of cell lines are obtained from Cancer Cell Line Encyclopedia (CCLE)[36]. Drugs' names are retrieved in PubChem database[37] to obtain their SMILES vectors. Then the SMILES vectors are converted into molecular graphs with RDKit library[38].

We combine the GDSC and CCLE datasets by selecting cell lines whose drug responses and gene expression profiles are both recorded. In total, there are 223 drugs and 700 cell lines. After removing missing screening of drug responses, 133,212 pairs of data points are left for experiments. Each cell line is depicted by a transcriptomic profile of 13,142 genes. In order to reduce the dimensionality of the input features to avoid potential over-fitting in model training, we refer to the connectivity map proposed in LINCS L1000 research[39], and preserve only the expression values of the 956 landmark genes, since it is testified that the expression pattern of other genes can be precisely inferred by the landmark genes.

Drug representation

Previous research have demonstrated that representing drugs with molecular graphs provides better predictive power than compressed representations such as SMILES[11,25], since the structural information of a molecule can be naturally preserved in a graph. Specifically, by considering the atoms in a molecule as nodes and the chemical bonds between atoms as edges, an undirected unweighted graph is constructed to represent the drug molecule. From the molecular graphs, node features proposed by DeepChem[40] such as atom symbol, atom degree, etc., can be extracted.

In this chapter, we further enhance the predictive power of a drug's graph representation by incorporating properer node and edge features. In the previous work[25], there are five types of node features, i.e., atom symbol, atom degree, the total number of Hydrogen, implicit value of atom, and whether the atom is aromatic. Nevertheless, these features are intuitively restricted to depict an atom in a molecule. Inspired by the Morgan Algorithm and Extended-Connectivity Fingerprints (ECFP)[15], we present a circular algorithm to compute the feature of an atom, considering both the atom itself and its surrounding environment.

```
INITIALIZATION:
$X_i^0 := h(F_i)$ Identifier $X_i$ of atom $i$ in molecule $M$ is initialized by hashing its chemical properties $F_i$
$r := 0$ Initialize interested radius from 0

UPDATING:
for $r \leq 3$ do
    for $i$ in $M$ do
        $X_i^r = \|_{j \in N_i}(b_j, X_j^r)$ Concatenate bond type $b_j$ between atom $i$ and $j$, and the identifier $X_j^r$ of atom $j$. $j$ belongs to the neighbour atoms $N_i$ of $i$
        $X_i^r = h(X_i^r)$ Transfer the identifier into an integer with hashing function
    end for
    $X_i := X_i \| X_i^r$
    $r := r + 1$
end for

REDUCTION:
for $i$ in $M$ do
    for $X_i^r$ in $X_i$ do
        Convert the identifier $X_i^r$ into binary vectors
        $B_i^r = binary(X_i^r)$
        for $b$ in $range(length(B_i^r), 64)$ do
            Convert the identifier $X_i^r$ into binary vectors
            $B_i^r(b) := 0$
        end for
        $B_i := B_i \| B_i^r$
    end for
end for

OUTPUT:
The final circular feature for atom $i$ is a 256-bit binary vector $B_i$.
```

Algorithm 1. Circular atomic feature computation

In Circular Atomic Feature Computation Algorithm 1, F_i refers to the chemical properties of atom *i* to be encoded, which involves the seven Daylight atomic invariants as the initial chemical properties, including number of immediate neighbors who are non-hydrogen atoms, the valence minus the number of hydrogens (meaning total bond order ignoring bonds to hydrogens), the atomic number, the atomic mass, the atomic charge, the number of attached hydrogens, and aromaticity. X_i^r denotes the identifier of atom *i* after collecting features from its *r*-hop neighbour atoms. h is the hashing function used for feature compression and *binary* is the function to convert hashed integers back to binary features. Operator $\|$ refers to the concatenation operation.

Figure 2 provides an example of the feature extraction procedure of atom 2 in the Butyramide molecule considering interested radius of 1. In particular, this algorithm involves three stages:

1. Initial Stage: Each atom in the molecule is assigned with a unique integer identifier which is generated by hashing a set of chemical properties.
2. Updating Stage: After initialization, each atom's identifier will be updated iteratively. Starting by radius $r = 1$:

    (a) An array will be formed by collecting the radius and the core atom's current identifier.
    (b) Next, the neighboring information of the atom that is r hops away from the core atom will be incorporated into the array. Ranked by the bond order (single, double, triple, and aromatic), the bond order and the current identifier of the interested atom are appended to the array.
    (c) Then the same hash function used in the initial stage is applied again to convert the array into a new integer identifier.
    (d) The above procedure is repeated for each atom in the molecule.
    (e) The radius will be updated as $r := r + 1$ and another iteration to update identifiers for all atoms will be started, unless r has already met the user's interested radius.

3. Reduction Stage: Eventually, for each atom, all the identifiers ever generated in the updating stage are converted into a 64-bit binary vector by calculating the modulus of the decimal integer and concatenated to form the final atom feature vector of length $64 \times (\text{radius} + 1)$.In the typical ECFP algorithm, the updating stage is aimed at discovering the unique substructures in the molecule, which will be consequently integrated into the fingerprint serving as a molecular-level representation. Thus, after the updating iterations, there will be a duplicate structure removal stage to eliminate the identical features which encapsulates the same surrounding environment of atoms. However, on the contrary to a molecular-level representation, in our case we require an atom-level feature where the structural duplication amongst atoms is not hampering feature reduction. Besides, the original algorithm only considers the last generated identifiers upon reducing them into the fingerprint, which is effective in producing the unique fingerprint of the molecule, whereas we preserve all the ever-generated identifiers to produce the atom-level features. This is because the last identifier is always computed considering a relatively large radius. The surrounding substructure might thus be identical for different atoms. Therefore, if merely considering the last identifier, certain atom-level features may be duplicated, and the corresponding atoms will be indistinguishable. Under such circumstances, identifiers generated at all radius levels are appended to form the atom-level feature.

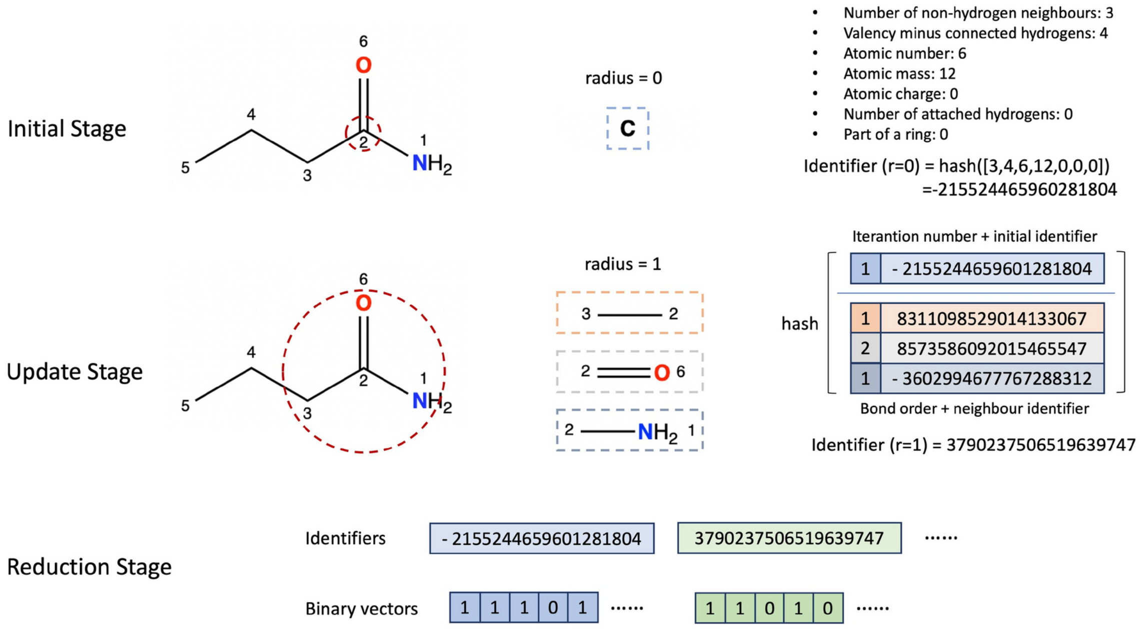

Fig. 2. Illustration of feature extraction procedure of atom 2 in the Butyramide molecule. In the initial stage, chemical features including number of non-hydrogen neibours, valency, atomic number, etc., are extracted and hashed to compute the initial identifier of each atom. In the update stage, starting from radius of 1, the bond orders and identifiers of the surrounding atoms (atom 3, 6 and 1) are combined and concatenated with the iteration number and the initial identifier of the focused atom (atom 2). The hash function is applied again on the concatenated feature to form the new identifier. Finally, in the reduction stage, each of the identifiers of atom 2 generated in the update stage are converted to a 64-bit binary vector and concatenated to form the final atom feature. In our implementation, we combined the binary vectors of radius 0, 1, 2 and 3, which forms a the final feature vector of length $4 \times 64 = 256$.

Computational framework

We utilize Graph Neural Networks (GNN) to learn the latent features of drugs' molecular graphs and Convolutional Neural Networks (CNN) to learn the representation of the gene expression data, and combine them together to predict the response level as shown in Fig. 1. Instead of concatenating latent features of drug and cell line as tCNN[11] and GraphDRP[25], we propose to leverage multi-head attention mechanism introduced by Transformer[41] to integrate the drug and cell line features effectively.

Each head H_i in the multi-head attention module can be formulated as

$$H_i = \mathit{Attention}(QW_i^Q, KW_i^K, VW_i^V) \tag{1}$$

where Q, K and V stand for the query, key and value used in an attention layer. To obtain the drug embed influenced by gene expressions, we use drug features encoded by GNN as Q and cell line features encoded by CNN as K and V. On the contrary, to learn the gene embed, we use cell line features as Q and drug features as K and V. Eventually, the integrative features are combined and fed into a predictor composed of a dense layer for drug response prediction.

After developing the model, we adopt Integrated Gradients[34] and GNNExplainer[33] to explore the saliency of inputs, i.e., atoms and bonds of the drug molecule and transcriptomic features of the cell line, which reveals the reaction mechanism of cancer cell lines and drugs.

Graph neural networks (GNN)

After constructing the molecular graphs and extracting atom-level features for the drugs, we develop GNN models to further learn the latent representation of the drugs. In this work, we take advantage of four types of GNN models and compare their performance in drug response and mechanism prediction: Graph Convolutional Networks (GCN)[42], Graph Attention Networks (GAT)[43], Relational Graph Convolutional Networks (RGCN)[44], and Relational Graph Attention Networks (RGAT)[45]. Similarly, the idea for such GNN models is to aggregate the information from a node itself and its neighborhood.

If we define the atom set in a drug's molecule as *V* and the bond set as *E*, the molecular graph of this drug can be given by $G = (V, E)$. Then we use an adjacent binary matrix $A \in \mathbb{R}^{N \times N}$ to represent the edge connection between nodes where *N* is the number of atoms, $a_{i,j} = 1$ denotes a connection between node *i* and *j*, and $a_{i,j} = 0$ denotes no connection. Additionally, a feature matrix $X \in \mathbb{R}^{N \times M}$ is used for representing the node features of atoms where *M* is the dimension of feature vector that has been extracted by the algorithm aforementioned.

The GCN layer is defined by

$$h_i = W \sum_{j \epsilon N_i \bigcup \{i\}} \frac{\widetilde{a}_{ij}}{\sqrt{\widetilde{d}_i \widetilde{d}_j}} x_i \tag{2}$$

where $\widetilde{A}$ is the adjacent matrix adding a self loop, $\widetilde{D}$ is the diagonal degree matrix with $\widetilde{d}_{i,i} = \sum_j \widetilde{a}_{i,j}$. x_i denotes the node feature vector, *W* is the weight matrix, and N_i is the neighbor node set of node *i*.

For the GAT layer, the attention coefficient of node *i* and *j* is defined as $e_{i,j} = a(Wx_i, Wx_j)$ according to[43], and is only computed when node *j* is in the neighborhood of node *i*. Then the GAT layer can be given by

$$h_i = \alpha_{i,i} W x_i + \sum_{j \epsilon N_i} \alpha_{i,j} W x_j \tag{3}$$

where x_i is the node feature vector, *W* is the weight matrix, N_i is the neighbor node set of node *i*, and $\alpha_{i,j}$ is the normalized attention coefficients with softmax function.

Notably, one drawback when adopting the typical GCN and GAT layers on molecular graphs is that they both dismiss the edge properties of the graph whereas the varying chemical bond types in a molecule could also impose a crucial impact on the drug's functional mechanism. In order to tackle this problem, we encode the chemical bond type (single, double, triple, and aromatic) into the edge features, which can be used for updating edges in the message passing procedure in GNN models. Edge features are directly supported by GAT layer and GATv2 layer which is designed to fix the static attention problem of original GAT layer[46].

To further investigate the effectiveness of edge features, we look into RGCN and RGAT models, which consider edge types as relations and differentiate the message passing patterns according to various relation types. In molecular graphs, edges represent chemical bonds that naturally possess disparate characteristics and should be treated accordingly. Therefore, we attempt to leverage RGCN and RGAT models to represent a molecule more precisely. Considering there are *R* relations in total, the RGCN layer can be defined as

$$h_i = W^{(root)} x_i + \sum_{r \epsilon R} \sum_{j \epsilon N_i^{(r)}} \frac{1}{|N_i^{(r)}|} W^{(r)} x_j \tag{4}$$

where $N_i^{(r)}$ denotes the neighbor node set of node *i* under relation *r*. Unlike general GCN layer, RGCN layer learns different weights specific to relation types. $W^{(r)}$ represents the weights corresponding to relation *r*, and $W^{(root)}$ corresponds to a special self-connected relation that is not included in *R*.

Similar to the way RGCN creates relation-specific transformations to update node representations, RGAT also proposes relation-specific attention weights for message aggregation. If we compute the attention coefficient of node *i* and *j* under relation *r* as $e_{i,j}^{(r)} = a(Wx_i^{(r)}, Wx_j^{(r)})$, RGAT layer can be formulated as

$$h_i = \sum_{r \epsilon R} \sum_{j \epsilon N_i^{(r)}} \alpha_{i,j}^{(r)} x_j^{(r)} \tag{5}$$

where $N_i^{(r)}$ denotes the neighbor node set of node *i* under relation *r* and $\alpha_{i,j}^{(r)}$ is the normalized attention coefficients with softmax function. Notably, the softmax function can be applied either over only attention coefficients under single relation type or all attention coefficients regardless of relation types, which result in within-relation GAT (WIRGAT) and across-relation GAT (ARGAT), respectively. In our experiments, ARGAT is found to outperform WIRGAT and is thereby used in the subsequent analysis.

Hyperparameters such as the radius in atom feature extraction, number of neural network layers, hidden sizes and dropout rates are searched to develop the best model. We first implemented GCN and GAT-based XGDP with the features extracted with radius of 0, 1, 2 and 3, and found radius of 3 obtained the best and most stable perfomace on the validation set. Grid search is then conducted to find the optimal parameters for number of layers of both GNN and CNN in [1, 2, 3, 4, 5], hidden sizes in [128, 256, 512] and dropout rates in [0, 0.1, 0.2, 0.3, 0.4, 0.5] . Finally, the number of layers is set to 2 for the GNN module and 3 for the CNN module. The hidden size is set to 128. And the dropout rate is configured as 0.5.

Model interpretability

To explore our proposed model's interpretability, we leverage on GNNExplainer[33] to identify the functional groups of molecular graphs and Integrated Gradients[34] implemented by Captum[47] to track the attributes of the genes in cancer cell line data.

Integrated gradients

Integrated Gradients is a gradient-based attribution method proposed by Subdararajan et al.[34]. Integrated Gradients is designed to satisfy two fundamental axioms, i.e., sensitivity and implementation invariance, and thus generate more reasonable explanations of neural network models than previous approaches such as Gradient * Input[48], Layer-wise Relevance propagation (LRP)[49], and DeepLIFT[50].

A prerequisite for a reasonable attribution using Integrated Gradients is to identify a baseline input. Take image networks as an example, the baseline inputs can be pixels equal to zero, constituting a black image. In our case, however, baseline cannot be simply set to zeros, since genes are seldomly expressed as zeros and picking zeros as baseline will render the explanation biased to certain genes with relatively high expression values. Our intention is to compute the average normal expression level as background for each gene, and study the effect when a gene is differentially expressed. Hypothesising that genes are normally expressed in most cell lines, we propose to identify suspiciously abnormal values with an interquartile range (IQR) filter. For each gene, we calculate IQR of its expression values on all the cell lines. Then expression values that are more than 2.22 times the IQR away from the median of the data are considered as outliers, which is roughly equivalent to remove the data points that have a z-score larger than 3 in the normal distribution. Thereafter we remove the outliers and compute the average of preserved expression values as the baseline of each gene.

After deciding the baseline input, Integrated Gradients computes the integral of the gradients along the path from the baseline input to the actual input. If we denote the actual input as *x* and the baseline input as $x\prime$, the integrated gradients can be defined by

$$Attribution_i\,(x) = (x_i - x'_i) \times \int_{\alpha=0}^{1} \frac{\partial F(x' + \alpha \times (x - x'))}{\partial x_i} d\alpha \tag{6}$$

where *i* refers to the interested dimension of inputs, and α is the interpolated value from $x\prime$ to *x*.

GNNExplainer

Gradient-based methods (e.g., Integrated Gradients) is suitable to explain models built on grid-like data in the text or image domain. However, GNN models built in the graph domain are developed to capture the structural information of graphs, and interpreting such models requires to explore how messages are passed through the graph structures[51]. GNNExplainer is one of the explanation methods aiming at analyzing models built in the graph domain. Comparing with gradient-based methods, GNNExplainer has been testified to be capable of capturing reasonable substructures of graphs such as functional groups of molecules, in the task of molecular property prediction[33]. Therefore, in this work, we adopt GNNExplainer to interpret the graph convolutional layers and identify the active functional groups of molecular graphs.

The theory of GNNExplainer is to identify the most salient subgraph and subset of node features for the model's prediction. It can be formulated in an optimization problem:

$$\max_{G_S} MI(Y,(G_S,X_S)) = H(Y) - H(Y|G = G_S, X = X_S) \tag{7}$$

where the mutual information *MI* reflects the change of model's output when using a subgraph G_S and subset of node features X_S. The prediction of the model can be given by $Y = \Phi(G, X)$, where Φ represents the function of the model, and *G* and *X* denote the input graph and node feature matrix. Although it is infeasible to retrieve the optimal subgraphs and feature subsets to solve the above problem directly, GNNExplainer has proposed their optimization framework to identify high-quality explanations in an empirical way.

Drug response prediction

In this section, we present the results of drug sensitivity prediction and the saliency maps of inputs. We experiment with various GNN models with and without involving edge features, and compare their performance with four baseline models, i.e., tCNN[11], GraphDRP[25], DeepCDR[26] and TGSA[27]. Particularly, gene expression data are used as cell line features in place of CNV data used in the original research of tCNN and GraphDRP. DeepCDR and TGSA focused on incorporating multi-omics profiles, whereas our work intends to investigate better profiling of drug features. For a fair comparison, we used the gene expression only version of DeepCDR and TGSA to explore if our proposed method properly represents the drugs and leads to better prediction. In addition, we decode the developed models to explore the salient functional groups of drug molecules and biomarkers that are potentially responsible for the biochemical activities.

Our models were implemented with PyTorch[52] and PyTorch Geometric[53] libraries. The performance of our experiments are evaluated by Root Mean Square Error (RMSE), Pearson Correlation Coefficient (PCC) and Coefficient of Determination (R^2). We performed a 3-fold cross-validation on our dataset. The mean and the standard deviation of the evaluation metrics obtained on the validation set are reported in the following sections.

Rediscovery of known drug and cell line responses

To test XGDP with the task of rediscovering response levels of known drugs and cell lines, we randomly shuffle all the pairs of drug and cell line and divide the dataset as described above. This strategy ensures one combination of drug and cell line can present only once in training, validation or testing set, but each drug or cell line can emerge simultaneously in all sets. The rediscovery task is designed to evaluate if the model is able to learn the reaction pattern of a drug from its response data with several cell lines, and predict the response levels between the drug and other unknown cell lines.

Table 1 presents the performances of the proposed method with different GNN layer and compares them with the baseline models. In the table, GAT_E and GATv2_E refer to GAT and GATv2 convolution with incorporation of bond types as edge features. It is shown that XGDP with GAT achieves the highest PCC and R^2 values, and all XGDP variants and the tCNN model achieve the lowest RMSE. DeepCDR and TGSA with only expression data obtain the worst RMSE, which is sensible since their research focus lie on incorporation of multi-omics profiles for drug response prediction. Compared with GraphDRP, our method extends the node features with the circular atomic descriptor as illustrated in Algorithm 1, and introduces multi-head attention to integrate the hidden features of drug and cell line rather than simple concatenation. It is evident in Table 1 that our refinement in the architecture leads to a better performance, especially on models with GAT convolution. Compared with tCNN, GAT- and GAT_E-based XGDP outperform tCNN on both PCC and R^2. Moreover,

Method	**Conv type**	**RMSE (↓)**	**PCC (↑)**	$\mathbf{R^2}$ **(↑)**
tCNN[11]	CNN	0.026 ± 0.000	0.920 ± 0.001	0.846 ± 0.001
GraphDRP[25]	GCN	0.027 ± 0.000	0.917 ± 0.001	0.840 ± 0.003
	GAT	0.042 ± 0.002	0.828 ± 0.011	0.609 ± 0.034
DeepCDR (exp)[26]	GCN	1.496 ± 0.018	0.841 ± 0.003	0.532 ± 0.057
TGSA (exp)[27]	GraphSAGE	1.072 ± 0.014	0.919 ± 0.002	0.845 ± 0.004
XGDP	GCN	0.026 ± 0.000	0.918 ± 0.001	0.843 ± 0.002
	GAT	0.026 ± 0.000	**0.923 ± 0.000**	**0.851 ± 0.001**
	GAT_E	0.026 ± 0.000	0.922 ± 0.001	0.849 ± 0.001
	GATv2_E	0.026 ± 0.000	0.921 ± 0.001	0.846 ± 0.001
	RGCN	0.026 ± 0.000	0.920 ± 0.001	0.845 ± 0.001
	RGAT	0.026 ± 0.000	0.920 ± 0.001	0.846 ± 0.002

Table 1. Performance of proposed and baseline models in the task of rediscovering known drug and cell line responses. All the models, except GraghDRP-GAT, achieve similar RMSE (~0.26). Best PCC and R^2 (marked in bold) is achieved by XGDP-GAT.

unlike GraphDRP and XGDP, tCNN used 1D convolutional layers to encode the SMILES notation of drugs, which renders it infeasible to decode the developed models to investigate structural saliency of drugs upon reaction with cancer cells.

Blind prediction of responses of unknown drugs

In the blind test of response prediction of unknown drugs, we divide the dataset by constraining the existence of drugs exclusively in training, validation, or testing set. Specifically, out of 223 drugs in total, 167 drugs' response data are used for a 3-fold cross-validation, and response data of 56 drugs are preserved for testing. The blind prediction task aims at testing whether the model developed on known drugs has the generalizability to predict responses of unknown drugs.

In the blind test experiment, we compare our method with tCNN, GraphDRP and TGSA. DeepCDR is ignored since the code to flexibly divide the dataset according to drug occurrence is not provided. As shown in Table 2, GAT- and GAT_E-based XGDP remarkably outperform other models. All baseline methods fail to perform well on blind test, especially in terms of R^2, which is in accordance with their original research[11,25,27]. tCNN and TGSA achieves a very small R^2 value (~0.02) and GraphDRP even results in negative R^2 values, which indicates these models are not making a sensible prediction when a brand new drug is given. Nevertheless, GAT-based XGDP models with and without edge features are able to achieve a significant improvement compared with the baselines.

XGDP achieves state-of-the-art performance in both rediscovery and blind test. However, scrutinizing the results of XGDP with various GNN types, it is observed that incorporating chemical bond type as edge features or relation types in relational GNNs does not always give rise to a better performance. Despite that RGCN outperforms GCN in both tasks, GAT-based XGDP suppresses all other edge-enhanced GAT models in Table 1, and in Table 2, only GAT_E performs better than plain GAT convolution. Nonetheless, in the next section, we will demonstrate that, to investigate the structural importance of molecules, it is essential to include edge features as well.

Prediction without cross-attention layers

To investigate the role of the cross-attention layers, we conducted an ablation study to compare XGDP with or without the attention layers. Particularly, we removed the two cross-attention modules following the GNN and CNN, and directly concatenated the features learned by the GNN and CNN modules as the input of the final dense layer. As shown in Table 3, it is evident that the cross-attention layer enhances the performance of drug response prediction and maintains better stability.

Discovery of drug mechanisms

We decode our models with GNNExplainer and Integrated Gradients, and present the attribution results of our best performing GATv2 model in this section. GNNExplainer is leveraged to explain the model's graph convolutional layers, and thus attribute the input molecular graphs. By interpreting a reaction pair of drug and cell line, each node and edge in the molecular graph is assigned with a saliency score. For each drug, we sum and average the saliency scores across all the cell lines for each node and edge, and perform a max-min normalization across the nodes or edges in one molecular graph. The normalized scores range from 0 to 1 and clearly illustrate the importance of a region of substructures to a drug's biochemical reaction. The normalized score is thereby used for a heatmap visualization, where red in Figs. 2, 3, 4, and 5 represents high saliency and blue represents low saliency.

To investigate the gene saliency in the pharmacodynamic process, we aggregate the saliency scores across all the cell lines for each drug in the test set, and thereby rank and select the top 50 genes with highest accumulated scores. Attribution of four drugs are illustrated as examples to support this study in the following sections.

Method	Conv type	RMSE (↓)	PCC (↑)	R^2 (↑)
tCNN[11]	CNN	0.056 ± 0.001	0.356 ± 0.019	0.027 ± 0.010
GraphDRP[25]	GCN	0.063 ± 0.002	0.450 ± 0.026	0.153 ± 0.048
	GAT	0.071 ± 0.003	0.351 ± 0.165	-0.041 ± 0.045
TGSA (exp)[27]	GraphSAGE	2.809 ± 0.035	0.329 ± 0.058	0.026 ± 0.078
XGDP	GCN	0.056 ± 0.000	0.400 ± 0.016	0.048 ± 0.015
	GAT	0.053 ± 0.001	0.448 ± 0.036	0.149 ± 0.052
	GAT_E	**0.052 ± 0.003**	**0.505 ± 0.090**	**0.164 ± 0.043**
	GATv2_E	0.055 ± 0.002	0.442 ± 0.041	0.058 ± 0.024
	RGCN	0.055 ± 0.001	0.405 ± 0.031	0.063 ± 0.045
	RGAT	0.055 ± 0.002	0.257 ± 0.061	0.063 ± 0.060

Table 2. Performance of proposed and baseline models in task of drug-blind prediction. Best performance (marked in bold) is achieved by XGDP-GAT_E.

Method	Conv type	RMSE (↓)	PCC (↑)	R^2 (↑)
XGDP (w/o attn)	GCN	0.045 ± 0.018	0.457 ± 0.476	0.480 ± 0.416
	GAT	0.038 ± 0.000	0.831 ± 0.003	0.679 ± 0.000
	GAT_E	0.037 ± 0.001	0.834 ± 0.010	0.691 ± 0.020
	GATv2_E	0.035 ± 0.000	0.847 ± 0.001	0.718 ± 0.002
XGDP	GCN	0.026 ± 0.000	0.918 ± 0.001	0.843 ± 0.002
	GAT	0.026 ± 0.000	**0.923 ± 0.000**	**0.851 ± 0.001**
	GAT_E	0.026 ± 0.000	0.922 ± 0.001	0.849 ± 0.001
	GATv2_E	0.026 ± 0.000	0.921 ± 0.001	0.846 ± 0.001

Table 3. Performance of proposed and baseline models in task of drug-blind prediction. Best performance (marked in bold) is achieved by XGDP-GAT_E.

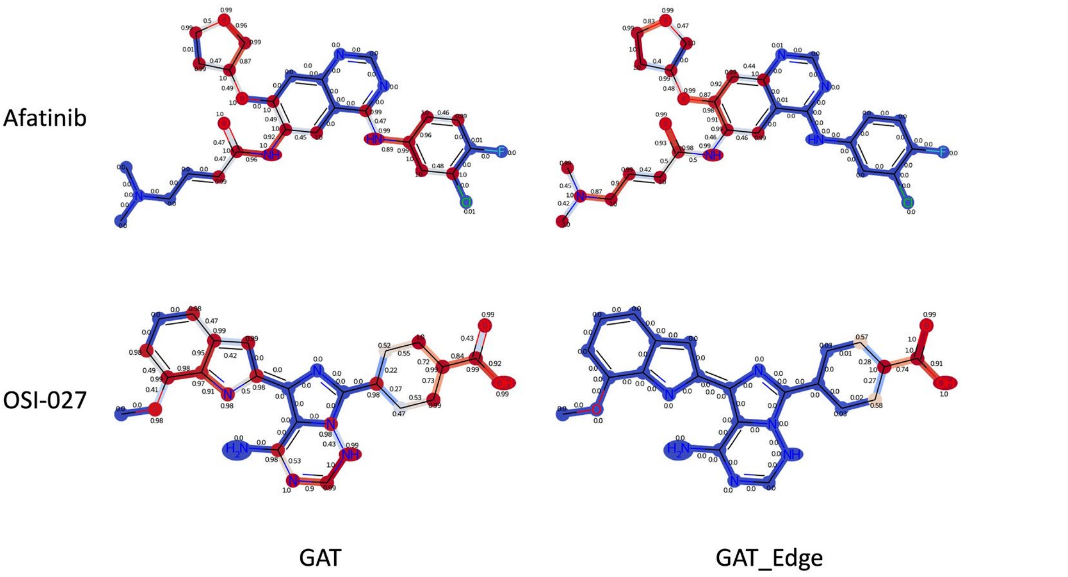

Fig. 3. Comparison of saliency maps generated by XGDP with GAT (left column) and GAT_E (right column). Afatinib (first row) and OSI-027 (second row) are used as examples.

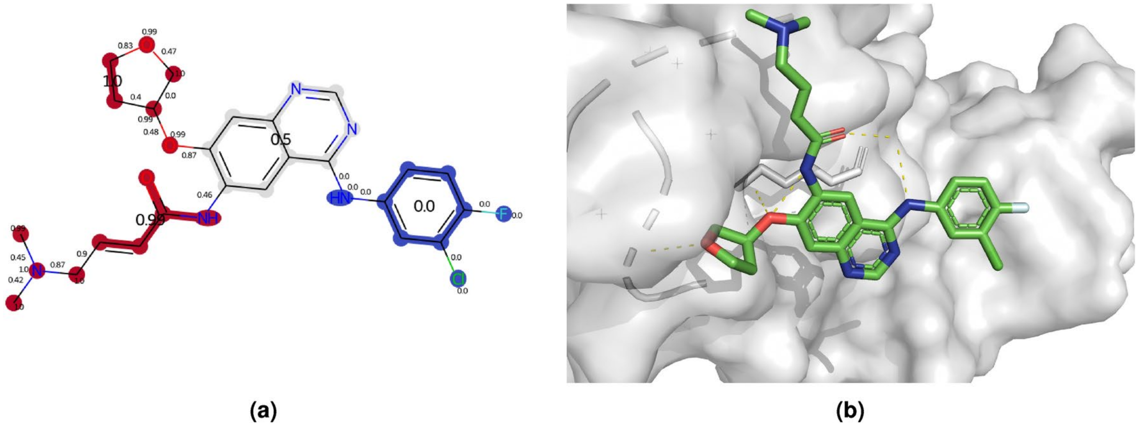

Fig. 4. (**a**) Saliency map of Afatinib, (**b**) binding mode of Afatinib with EGFR.

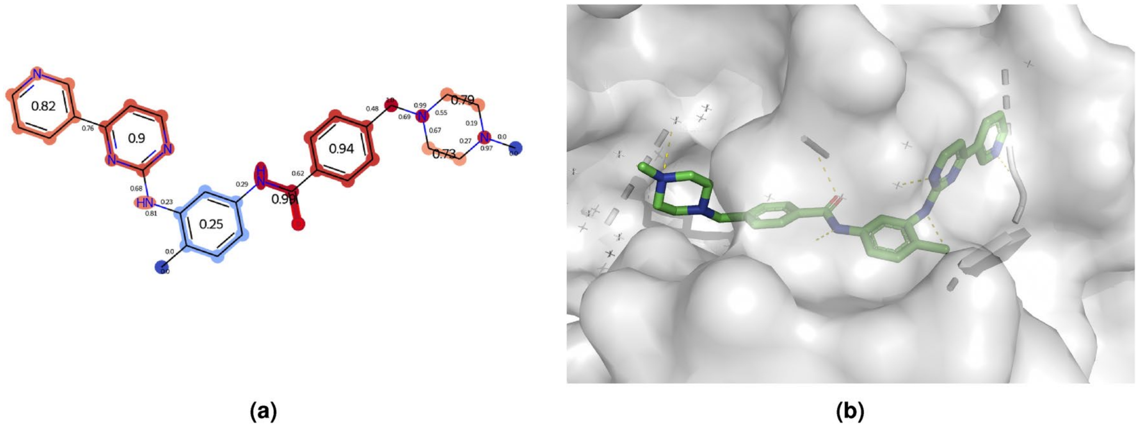

Fig. 5. (**a**) Saliency map of Imatinib, (**b**) binding mode of Imatinib with DDR1.

Necessity of including edge features

In the section of drug response prediction, we demonstrate that GAT-based XGDP obtained the best performance in both rediscovery and blind test. However, we do not observe any benefits of incorporating edge features such as bond types into model development. In this section, we will compare the molecular saliency heatmap obtained by interpreting GAT-XGDP with and without edge features.

Figure 3 presents the saliency maps generated by interpreting XGDP based on GAT and GAT_E. We observe that when edge features are absent in GAT convolutions, the model is likely to assign inconsistent saliency scores to atoms and bonds that are connected. Specifically, the case of atom with a high positive score and bond with a low negative score attached to the atom happens regularly in GAT-based models. This phenomenon thus hinders the study of substructure importance, since attached atom and bond are assigned with highly contrary saliency scores. However, this problem is overcome by GAT_E-based XGDP which incorporates edge features in model training. In the right column of Fig. 3, the significant (red) and insignificant (blue) structures are separated clearly instead of mixed with each other. Therefore, we conclude that edge features are essential for the model to correctly identify salient structures in molecules. The model decoding experiments in the following sections will all be conducted on XGDP-GAT_E model developed in the rediscovery test.

Chemical structure investigation

In this section, we took three drugs, i.e., Afatinib, Imatinib and Sunitinib, as examples to illustrate XGDP's capability of capturing salient substructures in drug reactions. We show the saliency heatmap of each drug and its binding mode with the protein target from the Protein Data Bank (PDB)[54]. For a clearer illustration, we leveraged the Extended Functional Groups (EFG) algorithm[55] to identify the common functional groups in our dataset, and calculated the average of the saliency score of each atom in the functional group to present the importance of each functional group in the drug molecules. In the illustrations of drug protein binding mode, we show the contacts between drug molecule and its surroundings (<5Å).

Afatinib is a famous EGFR inhibitor. According to[56], the acrylamide group in Afatinib is important for its inhibitation to kinanse activity of the ErbB family of proteins. As shown in Fig. 4, this functional group and other binding sites of Afatinib are successfully identified by our model. Imatinib is a DDR1 inhibitor. In Fig. 5, important binding sites, corresponding to the crystal structure of the DDR1 kinase in complex with Imatinib[57], such as the aminopyrimidine group, are assigned with a relatively high saliency score (> 0.9). Sunitinib is a potent PDGFR inhibitor[58]. In the crystal structure of PDGFR in complex with Sunitinib (6JOK), the binding sites have been remarkably identified by our model as shown in Fig. 6.

Biomarker and pathway analysis

Table 4 presents the top genes (ranking < 200 in 956 genes) identified by XGDP that are recorded to have interactions with the corresponding drugs in the drug-gene interaction database[59]. Particularly, ERBB3 and EGFR are ranked 60 and 78, respectively, out of 956 genes for Afatinib, DDR1 is ranked 16 for Imatinib, and PDGFA is ranked 113 for Sunitinib. Their specific interactions can be viewed in Figs. 3, 4, and 5.

Moreover, we perform Gene Set Enrichment Analysis (GSEA)[60] with GSEApy[61] using the attributed saliency scores. The top 5 enriched terms for each of the example drugs are shown in Table 5 together with their enrichment scores (ES) and normalized enrichment scores (NES). The identified pathways are well associated with cancer metastasis and progression. Specifically, epithelial-to-mesenchymal transition (EMT), which is one of the top enriched pathway for all drugs, is responsible for induction of cancer stem cells and immune escape during cancer progression in various cancers such as head and neck squamous-cell carcinoma (HNSC). Upregulation of KRAS signaling, which is usually the second most enriched pathway, is also found to be associated with a number of types of cancers such as breast cancer and pancreatic cancer. Therefore, we claim that the proposed method has the capability of capturing drug reaction mechanism and thus generating trustworthy prediction of drug responses.

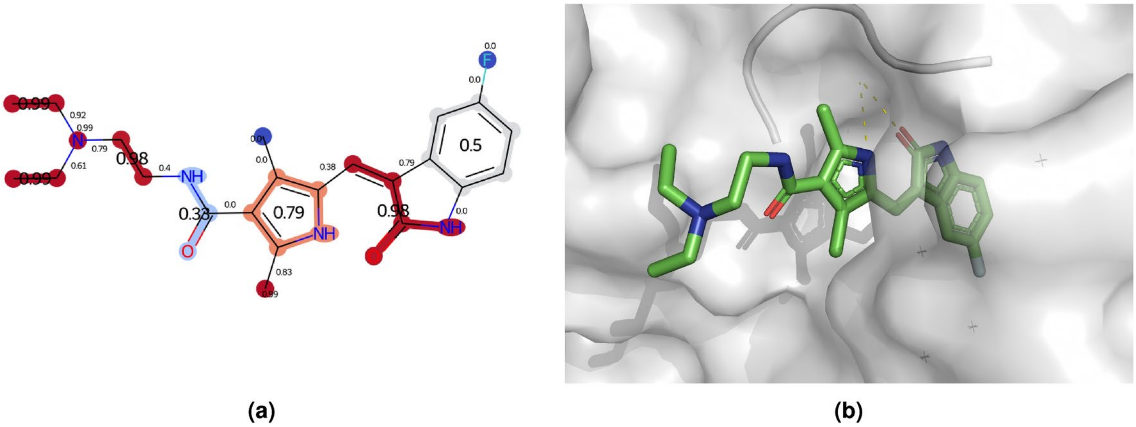

Fig. 6. (**a**) Saliency map of Sunitinib, (**b**) binding mode of Sunitinib with PDGFRA.

Drug	Gene	Rank	Saliency score
Afatinib	ERBB3	60	0.29
Afatinib	EGFR	78	0.26
Afatinib	ERBB2	108	0.23
Imatinib	MYC	2	0.85
Imatinib	SFN	10	0.64
Imatinib	DDR1	16	0.56
Imatinib	CDKN2A	23	0.54
Imatinib	EGFR	31	0.47
Imatinib	IKZF1	93	0.28
Imatinib	SMAD3	166	0.22
Sunitinib	NOS3	7	0.7
Sunitinib	EGFR	19	0.51
Sunitinib	FGFR2	47	0.36
Sunitinib	HMOX1	76	0.3
Sunitinib	PDGFA	113	0.27
Sunitinib	ERBB2	198	0.19

Table 4. Top salient genes identified by XGDP when predicting drug responses for Dasatinib, Erlotinib and Ponatinib.

Drug	Term	ES	NES
Afatinib	HALLMARK_KRAS_SIGNALING_UP	356.8	0.7
	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	350.5	0.68
	HALLMARK_INFLAMMATORY_RESPONSE	316.08	0.62
	HALLMARK_ALLOGRAFT_REJECTION	268.76	0.52
	HALLMARK_APICAL_JUNCTION	268.5	0.52
Imatinib	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	343.83	0.73
	HALLMARK_KRAS_SIGNALING_UP	298.37	0.63
	HALLMARK_APICAL_JUNCTION	289.58	0.62
	HALLMARK_MYOGENESIS	282.84	0.6
	HALLMARK_TNFA_SIGNALING_VIA_NFKB	280.9	0.6
Sunitinib	HALLMARK_EPITHELIAL_MESENCHYMAL_TRANSITION	345.81	0.73
	HALLMARK_KRAS_SIGNALING_UP	307.34	0.65
	HALLMARK_APICAL_JUNCTION	290.69	0.61
	HALLMARK_TNFA_SIGNALING_VIA_NFKB	277.54	0.59
	HALLMARK_UV_RESPONSE_DN	264.91	0.56

Table 5. Enriched pathways from GSEA on attributed saliency scores.

Conclusion

This study introduced a novel framework XGDP to predict response levels of anti-cancer drugs and discover underlying mechanism of action of drugs. To enhance the predictive power of GNN models, first we adapted the Morgan algorithm that is used for computing ECFPs to form our node features. Same procedures as Morgan algorithm were followed to identify the substructures of the molecule but the feature vector of each atom was assigned as the membership of the identified structures. Then we incorporated the type of chemical bonds as the edge features. These strategies enabled us to depict the molecule in a more meticulous manner and was testified to improve the GNN's prediction in terms of RMSE and PCC. Furthermore, we also attempted to explore relational GNN in the drug response prediction task, which describes edges as different relations and develops distinct message passing patterns for them. It was shown that RGCN outperformed GCN without edge features. However, due to the limited GPU resources, we were not able to train the RGAT model with an optimal batch size. This part of experiments is left for future investigations.

Moreover, we leveraged state-of-the-art attribution approaches in deep learning, GNNExplainer and Integrated Gradients, to explain our developed model. The explanations were visualized as saliency maps of both molecules and genes. Remarkably, those saliency maps could be supported by the SAR studies of the drugs. Consequently, we claim that our model is able to capture the significant functional groups of drugs and their potential targeted genes, and thus reveal the comprehensive mechanism of action of drugs. In the future, we intend to extend this study to a multi-omics level. Although genes contain the most vital information of the cause of disease, they do not directly interact with drugs in most cases. Therefore, protein and metabolites data should be considered. In addition, gene mutation and DNA methylation data may have a more direct reflection on the somatic abnormality, which are also expected to be explored in future works.

Data Availability

The drug response data can be downloaded from GDSC. And the gene expression data can be downloaded from CCLE under mRNA expression. Our implementation is released on Github (https://github.com/SCSE-Biomedical-Computing-Group/XGDP). Data preprocessing can be referred to our codes.

Received: 16 June 2024; Accepted: 11 December 2024
Published online: 02 January 2025

References

1. Singh, D. P. & Kaushik, B. A systematic literature review for the prediction of anticancer drug response using various machine learning and deep learning techniques. *Chem. Biol. Drug Des.* (2022).
2. Rafique, R., Islam, S. R. & Kazi, J. U. Machine learning in the prediction of cancer therapy. *Comput. Struct. Biotechnol. J.* **19**, 4003–4017 (2021).
3. Firoozbakht, F., Yousefi, B. & Schwikowski, B. An overview of machine learning methods for monotherapy drug response prediction. *Brief. Bioinform.* **23**, bbab408 (2022).
4. Baptista, D., Ferreira, P. G. & Rocha, M. Deep learning for drug response prediction in cancer. *Brief. Bioinform.* **22**, 360–379 (2021).
5. Partin, A. et al. Deep learning methods for drug response prediction in cancer: Predominant and emerging trends. arXiv preprint arXiv:2211.10442 (2022).
6. Moffat, J. G., Vincent, F., Lee, J. A., Eder, J. & Prunotto, M. Opportunities and challenges in phenotypic drug discovery: An industry perspective. *Nat. Rev. Drug Discov.* **16**, 531–543 (2017).
7. An, X., Chen, X., Yi, D., Li, H. & Guan, Y. Representation of molecules for drug response prediction. *Brief. Bioinform.* **23**, bbab393 (2022).
8. Heller, S. R., McNaught, A., Pletnev, I., Stein, S. & Tchekhovskoi, D. InChI, the IUPAC international chemical identifier. *J. Cheminform.* **7**, 1–34 (2015).
9. Weininger, D. Smiles. A chemical language and information system 1 introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* **28**, 31–36 (1988).
10. Suphavilai, C., Bertrand, D. & Nagarajan, N. Predicting cancer drug response using a recommender system. *Bioinformatics* **34**, 3907–3914 (2018).
11. Liu, P., Li, H., Li, S. & Leung, K.-S. Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network. *BMC Bioinform.* **20**, 1–14 (2019).
12. Chang, Y. et al. Cancer drug response profile scan (CDRscan): A deep learning model that predicts drug effectiveness from cancer genomic signature. *Sci. Rep.* **8**, 1–11 (2018).
13. Durant, J. L., Leland, B. A., Henry, D. R. & Nourse, J. G. Reoptimization of MDL keys for use in drug discovery. *J. Chem. Inf. Comput. Sci.* **42**, 1273–1280 (2002).
14. Reutlinger, M. et al. Chemically advanced template search (cats) for scaffold-hopping and prospective target prediction for 'orphan'molecules. *Mol. Inf.* **32**, 133 (2013).
15. Rogers, D. & Hahn, M. Extended-connectivity fingerprints. *J. Chem. Inf. Model.* **50**, 742–754 (2010).
16. Li, M. et al. DeepDsc: A deep learning method to predict drug sensitivity of cancer cell lines. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **18**, 575–582 (2019).
17. Shao, J. et al. S2dv: Converting smiles to a drug vector for predicting the activity of anti-HBV small molecules. *Brief. Bioinform.* **23**, 593 (2022).
18. Mikolov, T., Chen, K., Corrado, G. & Dean, J. Efficient estimation of word representations in vector space. arXiv preprint arXiv:1301.3781 (2013).
19. Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E. & Svetnik, V. Deep neural nets as a method for quantitative structure-activity relationships. *J. Chem. Inf. Model.* **55**, 263–274 (2015).
20. Sun, M. et al. Graph convolutional networks for computational drug development and discovery. *Brief. Bioinform.* **21**, 919–935 (2020).
21. Hu, L. et al. Dual-channel hypergraph convolutional network for predicting herb-disease associations. *Brief. Bioinform.* **25**, bbae067 (2024).
22. Zhao, B.-W. et al. A geometric deep learning framework for drug repositioning over heterogeneous information networks. *Brief. Bioinform.* **23**, bbac384 (2022).
23. Korolev, V., Mitrofanov, A., Korotcov, A. & Tkachenko, V. Graph convolutional neural networks as "general-purpose" property predictors: The universality and limits of applicability. *J. Chem. Inf. Model.* **60**, 22–28 (2019).

24. Xiong, Z. et al. Pushing the boundaries of molecular representation for drug discovery with the graph attention mechanism. *J. Med. Chem.* **63**, 8749–8760 (2019).
25. Nguyen, T., Nguyen, G. T., Nguyen, T. & Le, D.-H. Graph convolutional networks for drug response prediction. *IEEE/ACM Trans. Comput. Biol. Bioinform.* **19**, 146–154 (2021).
26. Liu, Q., Hu, Z., Jiang, R. & Zhou, M. DeepCDR: A hybrid graph convolutional network for predicting cancer drug response. *Bioinformatics* **36**, i911–i918 (2020).
27. Zhu, Y. et al. TGSA: Protein–Protein association-based twin graph neural networks for drug response prediction with similarity augmentation. *Bioinformatics* **38**, 461–468 (2022).
28. Ma, T. et al. DualGCN: A dual graph convolutional network model to predict cancer drug response. *BMC Bioinform.* **23**, 129 (2022).
29. Zuo, Z. et al. SWnet: A deep learning model for drug response prediction from cancer genomic signatures and compound chemical structures. *BMC Bioinform.* **22**, 1–16 (2021).
30. Chen, D. et al. Algebraic graph-assisted bidirectional transformers for molecular property prediction. *Nat. Commun.* **12**, 1–9 (2021).
31. Bishop, C. M., Svensén, M. & Williams, C. K. GTM: The generative topographic mapping. *Neural Comput.* **10**, 215–234 (1998).
32. Yoshimori, A. Prediction of molecular properties using molecular topographic map. *Molecules* **26**, 4475 (2021).
33. Ying, Z., Bourgeois, D., You, J., Zitnik, M. & Leskovec, J. Gnnexplainer: Generating explanations for graph neural networks. *Advances in Neural Information Processing Systems* **32** (2019).
34. Sundararajan, M., Taly, A. & Yan, Q. Axiomatic attribution for deep networks. In *International Conference on Machine Learning*, 3319–3328 (PMLR, 2017).
35. Yang, W. et al. Genomics of drug sensitivity in cancer (GDSC): A resource for therapeutic biomarker discovery in cancer cells. *Nucleic Acids Res.* **41**, D955–D961 (2012).
36. Cancer Cell Line Encyclopedia Consortium; Genomics of Drug Sensitivity in Cancer Consortium. Pharmacogenomic agreement between two cancer cell line data sets. *Nature* **528**, 84–87 (2015).
37. Wang, Y. et al. PubChem: A public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* **37**, W623–W633 (2009).
38. RDKit: Open-source cheminformatics. http://www.rdkit.org. [Online; Accessed 11-Apr.-2013].
39. Duan, Q. et al. L1000cds2: Lincs 1000 characteristic direction signatures search engine. *NPJ Syst. Biol. Appl.* **2**, 1–12 (2016).
40. Ramsundar, B. et al. *Deep Learning for the Life Sciences* (O'Reilly Media, 2019). https://www.amazon.com/Deep-Learning-Life-Sciences-Microscopy/dp/1492039837.
41. Vaswani, A. et al. Attention is all you need. *Advances in Neural Information Processing Systems* **30** (2017).
42. Kipf, T. N. & Welling, M. Semi-supervised classification with graph convolutional networks. arXiv preprint arXiv:1609.02907 (2016).
43. Veličković, P. et al. Graph attention networks. arXiv preprint arXiv:1710.10903 (2017).
44. Schlichtkrull, M. et al. Modeling relational data with graph convolutional networks. In *European Semantic Web Conference*, 593–607 (Springer, 2018).
45. Busbridge, D., Sherburn, D., Cavallo, P. & Hammerla, N. Y. Relational graph attention networks. arXiv preprint arXiv:1904.05811 (2019).
46. Brody, S., Alon, U. & Yahav, E. How attentive are graph attention networks? arXiv preprint arXiv:2105.14491 (2021).
47. Kokhlikyan, N. *et al.* Captum: A unified and generic model interpretability library for pytorch. arXiv preprint arXiv:2009.07896 (2020).
48. Shrikumar, A., Greenside, P., Shcherbina, A. & Kundaje, A. Not just a black box: Learning important features through propagating activation differences. arXiv preprint arXiv:1605.01713 (2016).
49. Bach, S. et al. On pixel-wise explanations for non-linear classifier decisions by layer-wise relevance propagation. *PloS One* **10**, e0130140 (2015).
50. Shrikumar, A., Greenside, P. & Kundaje, A. Learning important features through propagating activation differences. In *International Conference on Machine Learning*, 3145–3153 (PMLR, 2017).
51. Yuan, H., Yu, H., Wang, J., Li, K. & Ji, S. On explainability of graph neural networks via subgraph explorations. In *International Conference on Machine Learning*, 12241–12252 (PMLR, 2021).
52. Paszke, A. *et al.* Pytorch: An imperative style, high-performance deep learning library. In Wallach, H. et al. (eds.) *Advances in Neural Information Processing Systems 32*, 8024–8035 (Curran Associates, Inc., 2019).
53. Fey, M. & Lenssen, J. E. Fast graph representation learning with pytorch geometric. arXiv preprint arXiv:1903.02428 (2019).
54. Berman, H. M. et al. The protein data bank. *Nucleic Acids Res.* **28**, 235–242 (2000).
55. Lu, J., Xia, S., Lu, J. & Zhang, Y. Dataset construction to explore chemical space with 3d geometry and deep learning. *J. Chem. Inf. Model.* **61**, 1095–1104 (2021).
56. Solca, F. et al. Target binding properties and cellular activity of afatinib (BIBW 2992), an irreversible ErbB family blocker. *J. Pharmacol. Exp. Ther.* **343**, 342–350 (2012).
57. Canning, P. et al. Structural mechanisms determining inhibition of the collagen receptor ddr1 by selective and multi-targeted type ii kinase inhibitors. *J. Mol. Biol.* **426**, 2457–2470 (2014).
58. Abouantoun, T. J., Castellino, R. C. & MacDonald, T. J. Sunitinib induces PTEN expression and inhibits PDGFR signaling and migration of medulloblastoma cells. *J. Neuro-oncol.* **101**, 215–226 (2011).
59. Freshour, S. L. et al. Integration of the drug–gene interaction database (DGIdb 4.0) with open crowdsource efforts. *Nucleic Acids Res.* **49**, D1144–D1151 (2021).
60. Subramanian, A. et al. Gene set enrichment analysis: A knowledge-based approach for interpreting genome-wide expression profiles. *Proc. Natl. Acad. Sci.* **102**, 15545–15550 (2005).
61. Fang, Z., Liu, X. & Peltz, G. GSEApy: A comprehensive package for performing gene set enrichment analysis in python. *Bioinformatics* **39**, btac757 (2023).

Acknowledgements

This research was supported by AcRF Tier-1 grant RG14/23 of Ministry of Education, Singapore.

Author contributions

C.W., A.K.G., and J.C.R. conceived the experiment(s), C.W. and A.K.G. conducted the experiment(s), C.W., J.C.R., and A.K.G. analysed the results. C.W. and J.C.R. wrote and reviewed the manuscript.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to J.C.R.

Reprints and permissions information is available at www.nature.com/reprints.

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.