



OPEN

DeepInsight-3D architecture for anti-cancer drug response prediction with deep-learning on multi-omics

Alok Sharma^{1,2,5}✉, Artem Lysenko^{1,3,5}✉, Keith A. Boroevich¹ & Tatsuhiko Tsunoda^{1,3,4}✉

Modern oncology offers a wide range of treatments and therefore choosing the best option for particular patient is very important for optimal outcome. Multi-omics profiling in combination with AI-based predictive models have great potential for streamlining these treatment decisions. However, these encouraging developments continue to be hampered by very high dimensionality of the datasets in combination with insufficiently large numbers of annotated samples. Here we proposed a novel deep learning-based method to predict patient-specific anticancer drug response from three types of multi-omics data. The proposed DeepInsight-3D approach relies on structured data-to-image conversion that then allows use of convolutional neural networks, which are particularly robust to high dimensionality of the inputs while retaining capabilities to model highly complex relationships between variables. Of particular note, we demonstrate that in this formalism additional channels of an image can be effectively used to accommodate data from different omics layers while implicitly encoding the connection between them. DeepInsight-3D was able to outperform other state-of-the-art methods applied to this task. The proposed improvements can facilitate the development of better personalized treatment strategies for different cancers in the future.

Precision oncology is rapidly developing. However, only a very small percentage of patients can currently take advantage of it¹. The risks of side effects can be reduced by improving the prediction rate of drug response from targeted therapy which would undoubtedly improve patients' treatment. In this respect, drug prediction on in vivo datasets is a step towards clinical applicability. However, since in vivo data like The Cancer Genome Atlas (TCGA) repository (<https://www.cancer.gov/tcga>) has a scarcity of patient records and drug responses, it is difficult to train a model on in vivo datasets. Considering this challenge, Sharifi-Noghabi et al.² trained their computational model on in vitro data and obtained predictability on in vivo data. In vitro projects have compiled datasets such as Genomics of Drug Sensitivity in Cancer (GDSC)³ and Cancer Cell Line Encyclopedia (CCLE)⁴. These datasets consist of multi-omics profiles such as gene expression, copy number alteration (CNA) and somatic mutations. Although gene expression datasets have shown to be very useful^{3,5}, adding more omic layers could improve the predictability of pan-cancer models². Nevertheless, the platforms of these datasets are different. Furthermore, there exists a scarcity of related samples in GDSC datasets. This augments the problem of adequately estimating a model on the training set of the data.

The application of modern machine learning (ML) methods to problems in biomarker discovery has allowed better prediction of response to treatment. Full scope of these advancements is too extensive to cover here, and the following recent review provides a more complete overview of proposed approaches⁶. To position this work among previous efforts, the field of anticancer drug response prediction can be viewed from the perspective of either the types of data used or the methodology applied. From the data-centric perspective, our method attempts to leverage patient multi-omics data, and as such, is closest to the following two approaches. In², the MOLI method is proposed and learned on datasets from GDSC. It predicts response to a drug from TCGA and

¹Laboratory for Medical Science Mathematics, RIKEN Center for Integrative Medical Sciences, Yokohama, Japan. ²Institute for Integrated and Intelligent Systems, Griffith University, Brisbane, Australia. ³Laboratory for Medical Science Mathematics, Department of Biological Sciences, School of Science, The University of Tokyo, Tokyo, Japan. ⁴Laboratory for Medical Science Mathematics, Department of Computational Biology and Medical Sciences, Graduate School of Frontier Sciences, The University of Tokyo, Tokyo, Japan. ⁵These authors contributed equally: Alok Sharma and Artem Lysenko. ✉email: alok.fj@gmail.com; alysenko@g.ecc.u-tokyo.ac.jp; tsunoda@bs.s.u-tokyo.ac.jp

patient-derived xenograft (PDX)⁷ datasets. These collated datasets have multi-omics profiles: gene expression, CNA, and somatic mutation. The MOLI method reported the area under the curve (AUC) on seven datasets from TCGA and PDX. Their average AUC over the seven datasets was 0.63. Park et al.⁸ also constructed training and test dataset from GDSC and TCGA/PDX resulting in more than seven datasets and proposed the Super.FELT method. The performance of their method also included the seven datasets used in², though some of the resulting test samples slightly differed due to methodology. Nonetheless, their average AUC over the seven datasets was reported to be 0.68. When we reimplemented the Super.FELT method on the MOLI test sets, an average AUC of 0.65 was obtained.

From the methodological angle, our approach is distinguished by the idea of applying convolutional neural networks to tabular data, more specifically by converting it to images. Some of the drug response prediction applications with similar methodology are described as follows. Omics data as images was used with convolutional neural networks (CNNs) in OmicsMapNet⁹. However, OmicsMapNet method was limited in that it relied on pathway maps for creating images. This means that the genes not mapped to particular pathways would be excluded. Another application of CNNs to drug response prediction was DeepIC50¹⁰, and this one-dimensional (1D) CNN method achieved the best result on the evaluated gastric cancer-specific dataset. As the idea of tabular image transformation gained popularity, more specialized methods for optimizing this transformation were developed, notably IGTD¹¹, which was also applied to the task of drug response prediction. In contrast with these previous approaches that predominantly used a single type of data to create an image, this work specifically evaluates the potential of incorporating multiple, connected 'omics layers within the same image through the use of multiple color channels to capture these relationships. Additionally, we illustrate how application of class-activation maps (CAMs) feature extraction can be used to facilitate interpretation of constructed CNN models. The theoretical basis of these proposed improvements is briefly explained below.

Typically, for tabular data, traditional ML methods are used for feature extraction, feature selection and classification problems. Essentially, a column vector of size $d \times 1$ is processed via ML methods for mining and extracting relevant information to perform classification or regression tasks. The complexity of the data in the medical industry is ever-increasing, stretching the boundaries of ML techniques for phenotype identification relevant to disease diagnosis and analysis. In this regard, choosing a small group of crucial genes or elements from a larger set has become a critical step. Furthermore, determining phenotypes via feature extraction and classification processes plays a vital role. Feature selection is crucial in many different types of research and is not just used in analyzing genomic data. Therefore, the procedures of feature selection, feature extraction, and classification have a major role in determining how reliable ML algorithms are at identifying a subset of genes with the correct phenotypes. However, traditional ML approaches ignore the neighbourhood information and presumptively consider each component of a sample to be mutually independent.

Two-dimensional CNNs, on the other hand, are a class of deep learning architectures that have demonstrated promising results and attracted considerable interest in all forms of image analysis. Deep learning techniques are well-known and applied in various research fields, including biological studies^{12–19}. Contrary to traditional ML methods, CNNs undertake feature extraction and classification through their convolutional layers using an input image (a $p \times q$ feature matrix). CNNs are highly efficient, automatically extract features from spatially coherent pixels, requires fewer memory footprints enabling deeper network with fewer parameters because of sharing of weights across nodes²⁰, detect higher-order statistics and non-linear correlations, and use less samples to provide promising performance. An image in a local region is made up of spatially coherent pixels, which means that the nearby pixels share similar information. This neighbourhood information is extracted from the adjacent pixels by the layers of CNNs. Therefore, the adjacent pixels of a 2D input feature matrix should have adequate coherence in order to achieve meaningful output from a CNN.

Unlike other ML approaches, the order of nearby pixels in an image used by CNN is no longer independent. In order to arrange tabular data, e.g. omics data, to an image, DeepInsight²¹ pioneered through the element arrangement step and followed by mappings. Thereafter, automatic feature extraction and classification processes via CNNs are conducted. In the 2D pixel frame, the elements or genes are located according to how similar they are, and then the values are mapped to these locations to create the desired layout. With this method, any non-image sample can be converted into a set of images that can be used by CNNs. According to our understanding, it was the first method for transforming different types of non-image input into picture forms for the use of CNN architecture. This method of transforming tabular data to images has been applied and/or further developed in various fields^{22–28}. DeepInsight was a component in the Kaggle.com competition hosted by MIT and Harvard University that secured rank 1 on the leaderboard²⁹.

The proposed method is based on the DeepInsight method²¹ and the DeepFeature method³⁰. The DeepInsight-3D model considers the limitations of DeepInsight and DeepFeature and further extends their utility for multi-layered data. The overview of the proposed model is given in Fig. 1 (see Methods for the details). DeepInsight can convert a feature vector of size $d \times 1$ to an image of size $p \times q$ following its pipeline. It can effectively deal with data of one kind. However, multi-omics data has several layers and, in this case, DeepInsight is restrictive to performing the transformation. We looked into this limitation and expanded the algorithm for multi-layers. Therefore, a sample s_1 having L layers of column vectors $c_1, c_2 \dots c_L$ of size $d \times 1$ each, can now be transformed to a tensor-like shape of size $p \times q \times L$ using DeepInsight-3D. This enables the element arrangement and feature mapping for multi-layered cases. If the number of layers is restricted to three, then it will give a colored image of a sample s_1 . Using dimensionality reduction techniques, such as t-SNE³¹, DeepInsight-3D arranges similar elements from multi-layered data together in a 2D pixel frame and then performs element mappings of all the three layers. DeepInsight-3D can transform in two ways: (1) transformation using the dominant layer, where the most informative layer dictates the pixel locations and the values of other layers are mapped to these pixel locations. (2) The transformation by the equal contribution of layers. This approach finds the pixel locations in two stages; i.e., the element arrangement step is performed two times. In this way, the proposed DeepInsight-3D

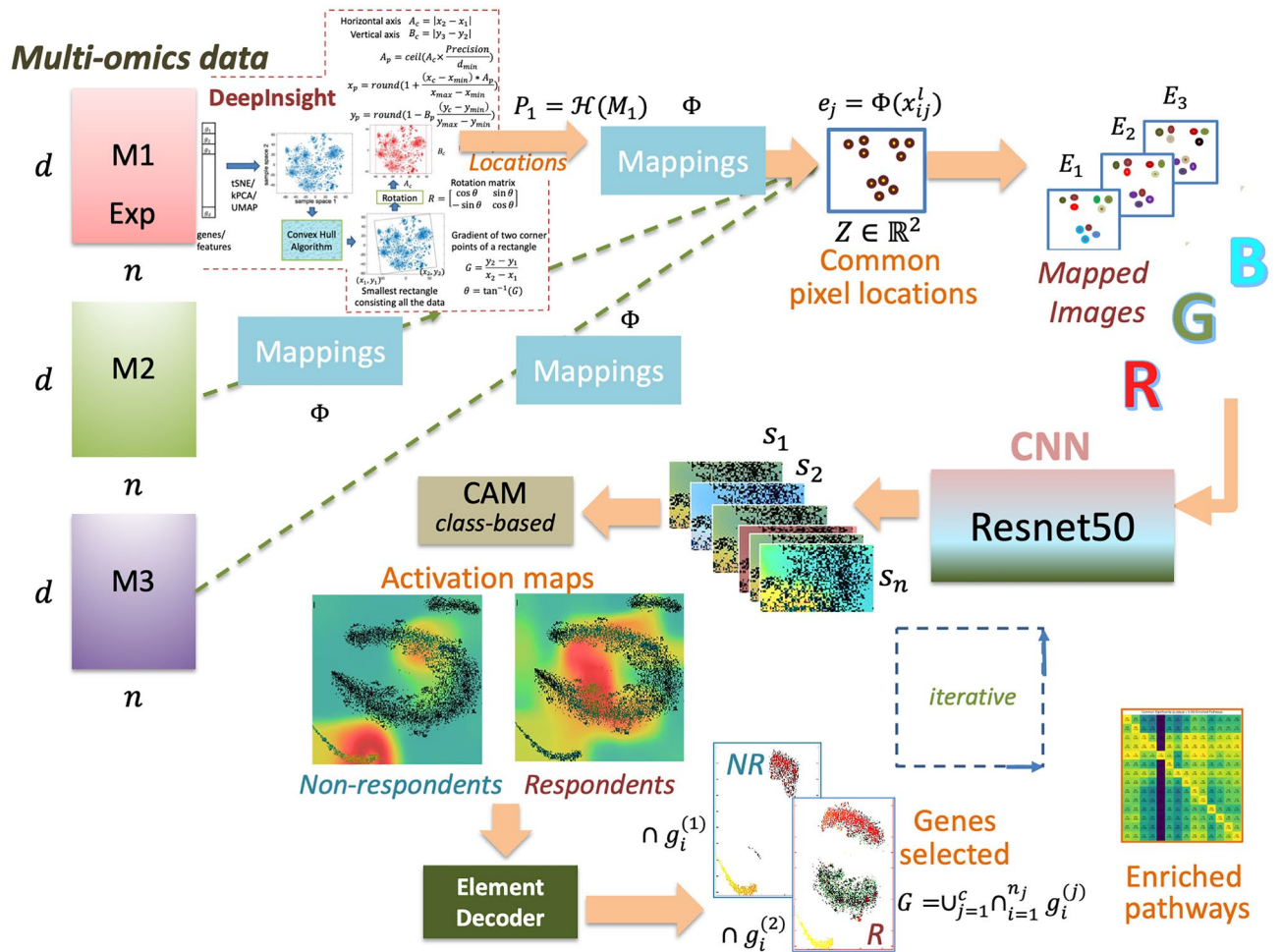


Figure 1. An overview of the DeepInsight-3D model. From the left multi-omics layers are processed via DeepInsight methodology and common pixel locations are found. After mapping omics data, corresponding images are constructed, which are processed to a convolutional neural network. Afterwards, CAM is used to find activation regions and element decoder is used to find a subset of genes.

model, first converts multi-omic or multi-layered data into corresponding images that are organized and colored, and then applies a convolutional neural network (CNN) with automatic feature extraction ability. Moreover, the element or feature selection of the transformed data ($p \times q \times L$) is carried out via the class-activation maps³² and the Element Decoder step. The feature selection is achieved by decoding activations obtained by the last ReLU layer of the CNN architecture.

Considering a multi-omics dataset M having n samples, d dimensions (number of genes or elements) and 3 layers, where $d \gg n$. This leads to a small sample size problem. However, since DeepInsight-3D converts omics vectors to colored images using transformation \mathcal{H} which handles groups of genes as features via element arrangement and direct mapping of genes to pixel locations, making the model robust compared to the limited sample size. For CNN the weights are not independent and it can bypass the curse of dimensionality problem³³. While the sample sizes of datasets are minimal, DeepInsight-3D with CNN can still perform well if the model is appropriately tuned.

Anticancer drug response is highly complex because not only tissue of origin leads to high diversity of subtypes, but also an ever-evolving landscape of somatic mutations make each cancer essentially unique. Combinations of mutations give rise to different strengths and weaknesses and therefore modern treatments commonly rely on biomarkers to identify which drugs are most likely to be effective³⁴. Single-gene biomarkers are now widely and successfully used, most notably BRCA1/BRCA2, ER and PR gene mutations for breast cancers³⁵, BRAF and KRAS mutations in colorectal and lung cancers³⁶ and PD-1/PD-L1 expression for targeting immunotherapy³⁷. However, biological processes often arise through complex interactions of different genes, cell types and signaling molecules that are far more challenging to profile, but vital for making further improvements to treatment³⁸. Based on these previous results it is now clear that modern machine learning approaches can indeed deliver better predictive performance, but as a trade-off, recovering biological mechanisms that underpin such models still remains challenging. More advanced methods, like deep learning, can automatically perform feature extraction and engineering to deal with high dimensionality of the input data as well as discover interactions, and the potential of such methods for complex biomarker discovery is now well-recognized³⁹. Interpretation

of these discovered relationships is important both for increasing confidence of doctors in using such tools in the clinic and for identifying research leads for future improvements. In this work we demonstrate that some of this complexity can be interpreted by mapping feature importance extracted by commonly-used methods like CAM to the biological pathways. Importantly, unlike in classical pathway enrichment analysis, the genes are extracted based on the magnitude of their contribution to the “answer” of the model rather than simply the level of expression. Our analysis shows that the model does indeed extract groups of genes that are non-random and have meaningful correspondence to different types of pathways, some of which were previously linked to anti-cancer drug response mechanisms.

The contributions of this work are as follows. DeepInsight-3D pipeline is presented where classification and feature selection can be performed for multi-layered non-image samples (or tabular data) through the application of CNNs. Two ways of image construction are introduced, (1) by mapping elements to the pixel locations of the dominant layer (shown in Fig. 1), and (2) by mapping elements to the pixel locations obtained by giving equal importance to all the three layers (implemented in the DeepInsight-3D package as an option). Element decoder is implemented to find genes or elements from the activation maps. We also demonstrate how the developed system can be used to interpret the CNN model and report on the identified key genes and biological processes identified as important for drug response prediction by respective models.

Results

Performance evaluation. The datasets and empirical organization have been summarized here. The details are given in the Experimental Setup section and Supplement Files. Seven datasets from the GDSC resource are used for the training/validation of the model. The number of total training samples for Paclitaxel is 389, Gemcitabine is 844, Cetuximab is 856, Erlotinib is 362, Docetaxel is 829 and Cisplatin is 829. The training set was divided into 90:10 ratio for separate smaller training set and validation set. The independent test sets are taken from PDX/TCGA resource. The number of test samples for Paclitaxel (PDX) is 43, Gemcitabine (PDX) is 25, Cetuximab (PDX) is 60, Erlotinib (PDX) is 21, Docetaxel (TCGA) is 16, Cisplatin (TCGA) is 66 and Gemcitabine (TCGA) is 57. Multi-omics data with expression, CNA and mutation layers are transformed into colored images by Deepinsight-3D. In all the datasets, the number of genes was over 13,000, which was much larger than the number of samples (less than 900). This leads to the high-dimensionality problem. However, since for CNN the weights are not independent, it is not affected by the curse of dimensionality³³. Information loss occurs when two or more than two pixels coincide in the same position of the image framework, compelling two or more than two gene values to attain an average in this particular pixel location.

Next, since expression data contains the majority of information, it has been used to find the pixel locations (Fig. 1). All the layers are utilized to map their values on these common pixel locations. DeepInsight-3D offers to use of manifold techniques, such as t-SNE, UMAP, kernel-PCA and PCA. In our experiments, we found t-SNE showing better results than other techniques, which is why t-SNE was selected. The technique, t-SNE, is not applied in the usual manner, as the visualization of samples in the 2D plane is not required. We want to find genes in the 2D plane, so transposition process is performed prior to applying t-SNE. This would provide points in the Cartesian coordinates corresponding to the genes or elements. After that, the convex-hull algorithm is used to find the minimum square around the plotted points (see Fig. 1). However, the square and the points within, need to be rotated to align with the horizontal and vertical axes. For this reason, a rotation matrix has been used for alignment. Afterwards, the rotated Cartesian coordinates framework is constructed into pixel coordinates (with fixed rows and columns). This process would give pixel locations from the expression layer. Then the expression data will be mapped to these locations. These steps are not followed for the other two layers (CNA and mutation). However, following the previous steps, their values will be mapped on the acquired pixel location. Once the mapping of all the three layers is completed, a colored image corresponding to a sample of multi-omics data can be visualized. In this work, a pre-trained ResNet-50 has been used to train the transformed images from the seven multi-omics datasets. The hyperparameters were adopted from DeepInsight version 2 (please see the Methods section and Supplement File 1 for more details). A Bayesian optimization technique was also applied for a drug dataset (Cisplatin) to improve the classification performance.

Two recently developed methods, MOLFI and Super.FELT, were used as benchmark methods. Other than these methods, we have also compared with non-negative matrix factorization (NMF), feed-forward net, and Geeleher et al.⁴⁰ as reported in². Moreover, a comparison was made with autoencoder (AE), artificial neural network after feature selection (ANNF), AutoBoruta Random Forest (AutoBorutaRF)⁴¹ and SVM⁴² as reported in⁸. All these methods were compared with many preceding algorithms and showed superior performance. The test set configurations (in terms of the number of samples) were kept the same for a fairer comparison. However, the test samples may slightly differ for the methods used in the Super.FELT work. The AUCs were computed for all possible drugs-method combinations and are given in Table 1.

It can be observed from Table 1 that for Paclitaxel, MOLFI and DeepInsight-3D produced promising AUCs. For Cisplatin, Super.FELT had the highest, and for Gemcitabine (TCGA), MOLFI produced the highest. For the remaining 4 drugs, Gemcitabine (PDX), Cetuximab, Erlotinib and Docetaxel, DeepInsight-3D produced the highest AUCs. The average AUC over all the seven datasets for the MOLFI method was 0.63 and for Super.FELT was 0.65. DeepInsight-3D produced an encouraging average AUC of 0.72. For other baseline methods, the average AUC reported are 0.43 for NMF, 0.53 for the feed-forward net, 0.59 for Geeleher, 0.45 for AE, 0.61 for ANNF, 0.38 for AutoBorutaRF and 0.53 for SVM. The confusion matrix of the DeepInsight-3D results over the seven datasets can be seen in Table S3 (Supplement File 1). In brief, the true positive rate (TPR) for Paclitaxel (PDX) was 24/38, Gemcitabine (PDX) was 10/18, Cetuximab (PDX) was 28/55, Erlotinib (PDX) was 13/18, Docetaxel (TCGA) was 6/8, Cisplatin was 1/6 and Gemcitabine was 10/36. The true negative rate (TNR) in the same order was 3/5, 5/7, 5/5, 2/3, 4/8, 46/60 and 18/21. The additional evaluation parameters such as average

Drug	Paclitaxel	Gemcitabine (PDX)	Cetuximab	Erlotinib	Docetaxel	Cisplatin	Gemcitabine (TCGA)	Average
Early inetgration via NMF	0.24	0.56	0.53	0.28	0.39	0.40	0.58	0.43
Feed forward net	0.68	0.48	0.43	0.37	0.69	0.44	0.65	0.53
Geeleher	0.52	0.59	0.58	0.67	0.59	0.62	0.53	0.59
AE*	0.44	0.47	0.42	0.33	0.50	0.46	0.50	0.45
ANNF*	0.64	0.69	0.43	0.65	0.64	0.68	0.57	0.61
AutoBorutaRF*	0.46	0.45	0.17	0.17	0.42	0.45	0.53	0.38
SVM*	0.49	0.64	0.41	0.67	0.53	0.47	0.47	0.53
MOLI	0.74	0.64	0.53	0.63	0.58	0.66	0.65	0.63
Super.FELT	0.64	0.65	0.55	0.76	0.64	0.73	0.61	0.65
DeepInsight-3D	0.74	0.72	0.71	0.85	0.78	0.68	0.53	0.72

Table 1. A comparison of DeepInsight-3D for drug response prediction with multi-omics profiles using test AUCs. The test samples are the same for all the comparators. The highest results are marked with bold fonts. *These methods are executed on a slightly different test sets (but similar enough for comparison) as reported in the Super.FELT paper.

accuracy, sensitivity, specificity, and F1-score over all the drug datasets was 0.62, 0.52, 0.73 and 0.58, respectively, by the DeepInsight-3D method (please refer to Table S7, Supplement File 1 for details). For uncertainty analysis, repeating the same configuration would not significantly affect the results and the results remain same as reported in Table 1. Nonetheless, the AUCs over three repetitions for drugs Paclitaxel (PDX), Gemcitabine (PDX), Cetuximab (PDX), Erlotinib (PDX), Docetaxel (TCGA) and Gemcitabine (TCGA) were 0.74, 0.72, 0.71, 0.86, 0.78, and 0.53. For Cisplatin (TCGA), Bayesian optimization technique with 10 objective functions were used and the best, for which the validation error was minimum, was selected. However, when we change the configuration of DeepInsight-3D method as an ablation study, we observe the variation in the results, which are reported in the summarized way in Figure S6 (Supplement File 1) and in details in Supplement File 3.

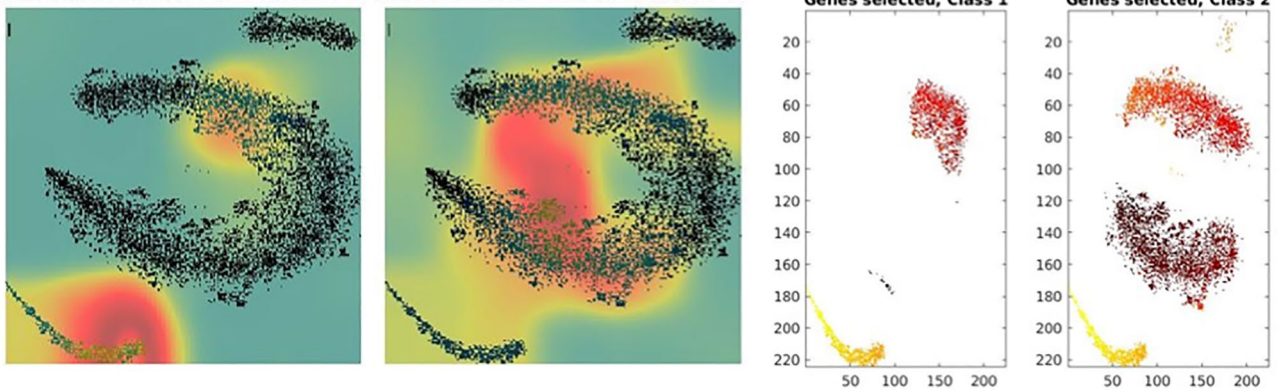
Feature selection to identify genes of interest. DeepInsight-3D can also perform feature selection via class-activation maps (CAMs) to identify genes of interest for each dataset. Since the data dimensionality is very large compared to the number of samples available, there is a high chance of producing an unstable model estimate. Furthermore, not all genes can be well represented in a limited pixel-framework. Appropriate feature selection would reveal background scientific mechanisms. Therefore, we applied an iterative way of conducting feature selection. Gene selection can be performed in 3 ways, (1) considering CAM values for every training sample, (2) taking an average of CAM over training samples, and (3) class-based CAM (described in the Methods section) where the average over a particular class is considered. In this work, class-based CAM has been applied for gene selection. Table S5 (Supplement File 1) depicts the number of genes selected for each drug dataset (non-respondents and respondents). For parameters related to feature selection, see Table S4 (Supplement File 1) and feature selection procedure in Figure S1 (Supplement File 1). The activation maps are shown in Fig. 2. Here the activations of Paclitaxel dataset at stage-1 and stage-5 (last stage) are illustrated. Only the training set of Paclitaxel has been used to find the activations. The selected genes are also depicted on the right-hand side of Fig. 2. The activations for all the seven drugs are depicted in Figure S2 (Supplement File 1).

Pathway-centric context of discovered gene sets. Gene sets identified as important by each drug-specific model were mapped to Reactome⁴³ pathways and QIAGEN Ingenuity Pathway Analysis (IPA) knowledgebase⁴⁴, as described in the methods section. This analysis has revealed that there were both a unique as well as a shared component that was recovered as important for all drugs. The overlap between drug-specific gene sets (Figure S3, Supplement File 1) was very low and consequently we have found that there was very low coverage of relevant common pathways in Reactome database, although this analysis did identify a small number of significantly enriched pathways for most of the drugs. Therefore, we have done additional analysis using IPA software, which is backed by a much larger database and has more pathway definitions with higher levels of granularity. In several cases most significantly enriched subsets have previously been reported in literature as being linked to particular drugs or that class of drugs. This suggests that the proposed system does have some functionality to not only improve quality of drug response prediction, but also allow discovery of meaningful biological processes that may be involved. Summary results are available as a Supplement File 2 (Tables S8–S12); and relevant key findings are summarized below.

Overall, seven evaluations were performed across six different drugs and two independent datasets (TCGA and PDX). One drug, gemcitabine, was profiled in both of these datasets. As described in detail in the methods section, our model interpretation approach allows recovery of class-specific important feature sets, therefore in total fourteen sets were generated and analyzed—i.e., separate gene sets were produced for responders and non-responders for each drug/data source combination. Note that for the purposes of interpretation, feature sets identified for the responder and non-responder classes are of equal interest and relevant gene combinations may be allocated to either or even both of the sets. Profiled drugs have covered a range of commonly exploited mechanisms of action: growth factor blockers, microtubule-binding and different forms of DNA damage.

Stage 1

Activation area: Class 1 Activation area: Class 2



Stage 5

Activation area: Class 1 Activation area: Class 2

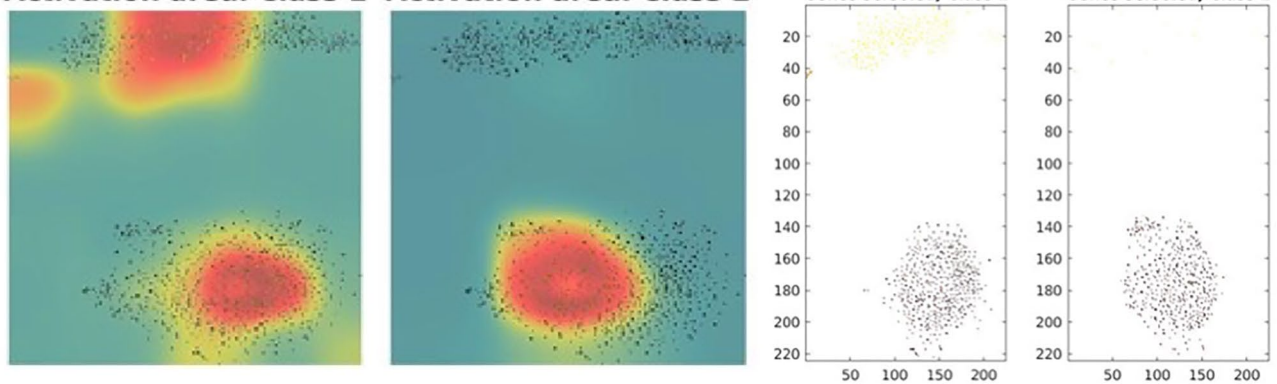


Figure 2. Activations and gene selection on Paclitaxel at Stage 1 and Stage 5. Class 1 is non-responders and Class 2 is responders. The left side of the figure depicts the activations in Class 1 and Class 2, and the right side illustrates the selected genes in the respective classes.

The initial analysis that was based on the Reactome pathway definitions did not identify substantial commonalities between significantly enriched pathways across all of the drugs (Figure S4, Supplement File 1), and furthermore very few results were returned for some of the gene sets. This is likely caused by the relatively small size of the dataset and, in particular, very small number of responders. Although DeepInsight approach does allow for better model generalization in smaller datasets, some effect of noise is still unavoidable in such cases and can lead to greater variability of feature selection from a pool of all potentially usable options. As Reactome analysis did indicate the presence of some significant pathways, a follow-up analysis was done in IPA to explore whether the relatively low number of significant hits was solely due to the challenging dataset or relatively sparser coverage offered by the database used in the evaluation.

Metabolic pathway enrichment analysis using IPA software has identified 269 (responders) and 342 (non-responders) non-redundant canonical pathways that were significant in at least one of the drug-specific sets. None of the pathways were found to be significantly enriched in all 7 of the drug-specific sets in either responder or non-responder categories (Fig. 3). However, a number of key common pathways were recovered for 5–4 sets, suggesting that those may be potentially important mechanisms of multidrug resistance (Fig. 4). Full results of this analysis, including log-transformed significance values for individual pathways are made available in Supplement File 2 (Tables S8–S12). One of the standout features of these results is functional consistency of the results across multiple classifiers. Given the size of the datasets, functions independently identified as important multiple are least likely to be due to overfitting or noise, therefore, we have decided to dissect these functional categories further and check if they have been previously reported to be linked to multidrug resistance. Some notable examples include “STAT3 Pathway” that was significantly enriched among both responder (Docetaxel $p=0.004$; Gemcitabine (PDX) $p=0.017$; Erlotinib $p=0.022$; Paclitaxel $p=0.01$) and non-responder (Gemcitabine (TCGA) $p=0.02$; Docetaxel $p=0.04$; Gemcitabine (PDX) $p=0.0008$; Erlotinib $p=0.011$; Paclitaxel $p=0.049$) gene sets. STAT3 gene is often highly expressed in various cancers and can lead to therapy resistance by upregulating expression of anti-apoptotic proteins, stimulating DNA repair and cell proliferation⁴⁵. Functionally closely related pathways “PI3K/AKT Signaling” (responders: Docetaxel $p=0.004$; Gemcitabine(PDX) $p=0.017$; Erlotinib $p=0.022$; Paclitaxel $p=0.011$; non-responders: Gemcitabine(TCGA) $p=0.004$; Docetaxel $p=6.5 \times 10^{-5}$; Gemcitabine(PDX) $p=0.048$; Erlotinib $p=0.004$; Cetuximab $p=0.021$) and “JAK/STAT Signaling” (responders:

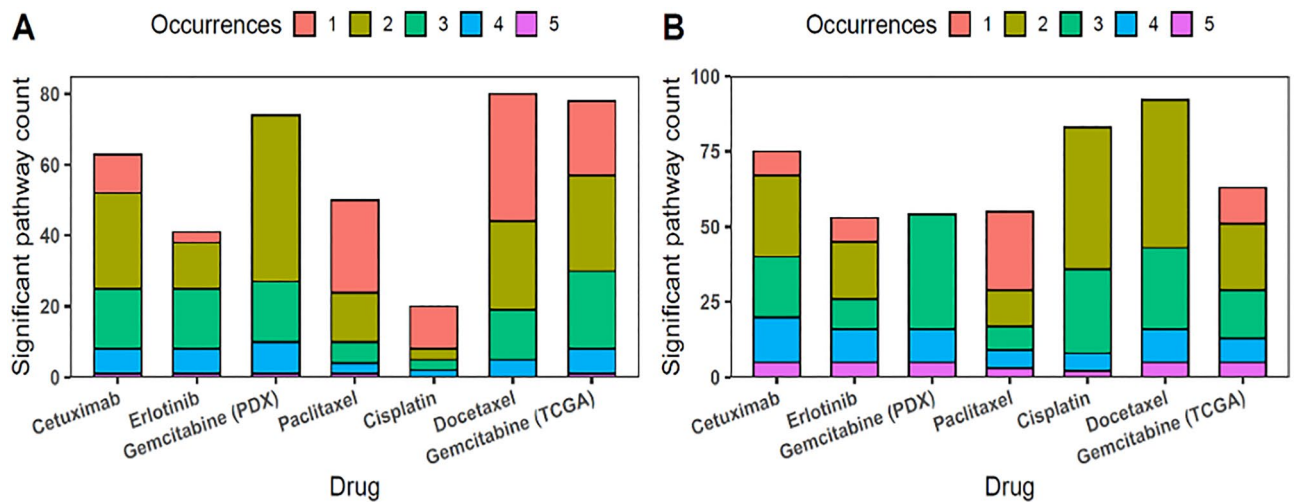


Figure 3. Annotation of selected genes to IPA pathways across all drug-specific classifiers for responder (A) and non-responder (B) classes. Occurrences bars show how many times that part of annotations is also found in sets for other classifiers. Data were analyzed through the use of QIAGEN Ingenuity Pathway Analysis (IPA), 2022 Summer Release (July) data, with client 01–20-04 (installed on 14 Sep 2021).

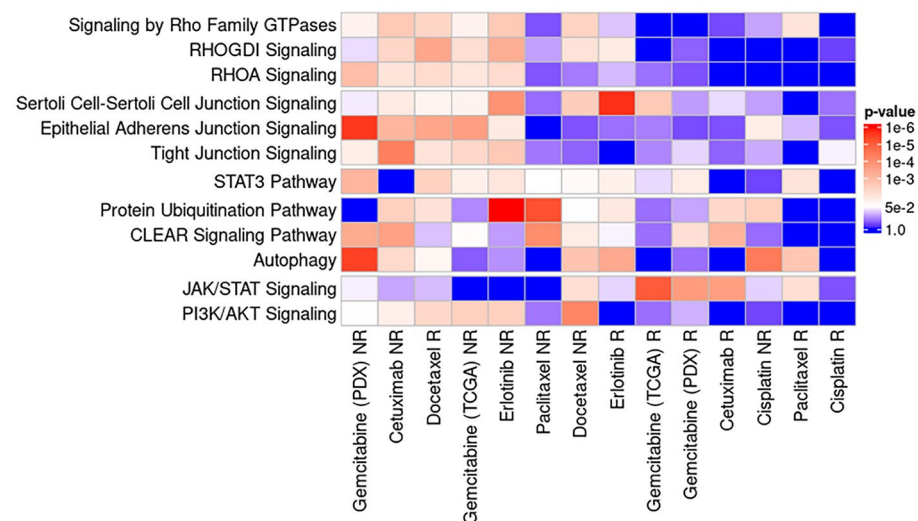


Figure 4. Notable pathways identified as significant by metabolic pathway enrichment using IPA software in gene sets of importance in multiple prediction models. Data were analyzed through the use of QIAGEN Ingenuity Pathway Analysis (IPA), 2022 Summer Release (July) data, with client 01–20-04 (installed on 14 Sep 2021).

Gemcitabine(TCGA) $p = 8.6 \times 10^{-6}$; Gemcitabine(PDX) $p = 2 \times 10^{-4}$; Cetuximab $p = 2.3 \times 10^{-4}$; Paclitaxel $p = 0.008$; non-responders: Docetaxel $p = 0.008$) were also found important in multiple sets and are likewise well-known to be involved in multidrug resistance^{46,47}. All of these pathways are well-known to play a key role in cancer cellular growth, survival, and resistance to treatments. Likewise, three variants of Rho GTPase family signaling pathways were recovered multiple times among the non-responder gene sets, which are an important mechanism for the upstream activation of JAK/STAT pathway⁴⁸ and cancer progression⁴⁹. These results are in line with current understanding of the pivotal role these signaling pathways play in various drug resistance mechanisms⁵⁰.

Another common theme was the potential importance of protein degradation and recycling mechanisms for cancer drug resistance. Functionally-linked ubiquitination (responders: Docetaxel $p = 0.009$; Erlotinib $p = 0.014$; Cetuximab $p = 0.005$; non-responders: Cisplatin $p = 0.003$; Erlotinib $p = 4.8 \times 10^{-7}$; Cetuximab $p = 0.004$; Paclitaxel $p = 5.7 \times 10^{-6}$), autophagy (responders: Docetaxel $p = 0.031$; Erlotinib $p = 4.1 \times 10^{-4}$; Paclitaxel $p = 0.002$; non-responders: Docetaxel $p = 0.002$; Cisplatin $p = 3.7 \times 10^{-3}$; Gemcitabine (PDX) $p = 3.4 \times 10^{-6}$; Cetuximab $p = 0.006$) and lysosome regulation pathways “CLEAR Signaling Pathway” (responders: Gemcitabine (PDX) $p = 0.009$; Cetuximab $p = 7.2 \times 10^{-4}$; non-responders: Gemcitabine (TCGA) $p = 0.044$; Docetaxel $p = 0.017$; Gemcitabine (PDX) $p = 4.4 \times 10^{-4}$; Cetuximab $p = 2.6 \times 10^{-4}$; Paclitaxel $p = 9.3 \times 10^{-5}$) were found multiple times among the

non-responder gene sets. Ubiquitination marks proteins for proteolysis and this mechanism is usually preserved largely intact in cancer cells, suggesting functional essentiality⁵¹. Both lysosome⁵² and ubiquitination-mediated proteasome⁵³ are believed to be involved in multidrug resistance. These processes are part of a broader set of autophagy processes that in turn lead to cellular rejuvenation, stress resistance and medication of anti-drug responses on a cellular level⁵⁴.

The last commonly represented theme was the presence of several pathways to do with extracellular structures and cell adhesion. Some specific examples in this category included “Epithelial Adherens Junction Signaling” (responders: Gemcitabine (PDX) $p = 0.028$; non-responders: Gemcitabine (TCGA) $p = 2.4 \times 10^{-4}$; Cisplatin $p = 0.019$; Gemcitabine (PDX) $p = 2.5 \times 10^{-6}$; Erlotinib $p = 0.015$; Cetuximab $p = 8.1 \times 10^{-4}$), “Sertoli Cell-Sertoli Cell Junction Signaling” (responders: Gemcitabine (TCGA) $p = 0.002$; Docetaxel $p = 0.026$; Erlotinib $p = 2 \times 10^{-6}$; non-responders: Gemcitabine (TCGA) $p = 0.026$; Docetaxel $p = 0.003$; Erlotinib $p = 1.2 \times 10^{-4}$; Cetuximab $p = 0.016$) and “Tight Junction Signaling” (responders: Docetaxel $p = 0.010$; non-responders: Gemcitabine (TCGA) $p = 0.005$; Gemcitabine (PDX) $p = 0.020$; Erlotinib $p = 0.002$; Cetuximab $p = 4.7 \times 10^{-5}$). Interestingly, multiple pathways in this broader functional category were identified in both Reactome and IPA-based enrichment analyses. These processes are important for formation of tumor microenvironment, interactions with the immune system as well as cancer cell dissimulation/metastasis. Some of the identified pathways were previously reported to be involved in drug resistance, for example paxillin was reported to be involved in tyrosine kinase inhibition in lung cancer⁵⁵ and was significantly enriched for both of the tyrosine inhibitor drugs sets (Cetuximab and Erlotinib) in our analysis. More generally, tight junctions play a role in drug efflux⁵⁶ and are important for anoikis and metastasis⁵⁷. Extracellular matrix can confer a cell adhesion-mediated drug resistance in some cancers, like glioblastoma⁵⁸ and small cell lung cancer⁵⁹. However, the potential broader roles in anti-cancer drug resistance are currently not very well defined and these results hinting at prominent involvement highlight a need for further study.

Among the pathways found to be important and significant across several drugs two appear to be relatively unexplored with respect to selected drugs. The first example is “Tryptophan Degradation X”, which was significant for Gemcitabine (TCGA), Erlotinib (PDX) and Cetuximab (PDX). This result is of interest because tryptophan metabolism was proposed as a source of potential targets for future anti-cancer drugs^{60,61}, and better understanding of its interplay with established treatments may facilitate these developments. Another pathway of interest was “Clathrin-mediated Endocytosis Signaling”, which was significant for the same set of drugs. Endocytosis is a process by which molecules on the surface of a cell, like nutrients and signaling molecules, are internalized, mostly for subsequent post-processing or degradation. As a number of cancer driver and drug target proteins are located on the cell surface, it plays an important but very complex role in cancer⁶², and clathrin is one of the key components of this process⁶³. Likewise, the role of this pathway in mediating response to these particular drugs still appears to be unclear. Of course, more detailed follow up of these links will be necessary to verify the observed effects and understand their roles further.

Discussion

Response to cancer treatment is highly complex as it is underlain by both genomic and somatic variations⁶⁴, which can interact to give rise to yet more variation across all other ‘omics. This inevitably leads to very high degree of tumor heterogeneity across patients due to their genomic characteristics⁶⁵, which in turn leads to often vast differences in efficacy of specific treatments. Better understanding of these factors is essential for improvements in drug response prediction that will not only allow more patients to benefit through better targeting of available therapies¹ but also avoid potential adverse effects^{66,67}. In order to manage the complexity, increase scale and minimize risks to patients, drug response is often studied on large in vitro datasets⁴. The discovered mechanisms are then taken forward to pre-clinical models and eventually to clinical trials. One of the key aspects, therefore, is translatability of these results to actual better outcomes in cancer patients⁶⁸. Specifically in the case of computational predictive models, the performance achieved on the in vitro training data must be comparable to that on the real patient data. Importantly, although in vitro high-throughput screening is inevitably required, the number of possible genomic factors and drug combinations make it impossible to exhaustively test all possible options⁶⁹. Therefore, statistical and machine learning models are increasingly necessary to mine these huge datasets and discover novel clinically usable patterns⁷⁰. In turn, these types of analyses are also used to streamline experimental work and therefore reduce the amounts of data that needs to be collected. As a possible additional benefit, more advanced machine learning methods can help to identify combinations of factors that contributed to making a particular prediction in each case and in that way facilitate the identification of pharmacological functions of these targets⁷¹.

In this context, DeepInsight-3D offers two important capabilities. First one is an ability to make accurate predictions of drug response from the high-dimensional ‘omics data. As the cost of collecting such samples continues to decrease, this information is now increasingly used by doctors to make better treatment decisions, creating a need for relevant tools. Potentially, ‘omic-based predictions can be combined with other clinical data as part of the decision process. The second capability is the discovery of underlying features of importance that may then facilitate identification of relevant mechanisms. Deep learning methods have an inherent ability to discover non-linear and interacting features when sufficient data is available—and therefore can expand the range of characterizable features beyond what is possible using less complex methods. DeepInsight-3D extends the versatility of applying CNN to multi-layered tabular data. In this work, DeepInsight-3D provided very encouraging results on drug response multi-omics data. DeepInsight-3D was able to produce an average AUC of 0.72 over seven drug response datasets which is encouraging compared to competing methods in the literature.

The promising performance of DeepInsight-3D could be attributed to the following: (1) The inclusion of neighborhood information through the placement of similar elements together and dissimilar ones apart. (2) Three different layers can be combined into colored images, keeping all the possible information as kernels. (3)

The interconnectivity of deeper networks (CNNs) taps into the non-linear connectivity between elements which shallow or traditional machine learning methods could miss. However, at the same time as one of limitations, a limited number of samples affects the confidence of the model, which could lead to unsatisfactory performance. Therefore, it is generally opted to have ample training samples for better estimation of the model.

Deep learning nets, such as CNN, have many merits, such as automatic feature extraction, finding hidden structures from hyper-dimensional data, finding higher-order statistics of image and non-linear correlations, economical use of neurons for large input sizes allowing much deeper networks are plausible with fewer parameters²⁰, and a parsimonious memory footprint. These properties of CNN can be integrated with the inception of DeepInsight-3D for non-image tabular data with multi layers.

In machine learning techniques for tabular data, any two features are considered mutually independent. However, DeepInsight tries to establish a relationship through the element arrangement step by positioning similar elements together and dissimilar ones apart²¹. DeepInsight-3D further extends this property to multi-layered data. Moreover, the application of DeepFeature is extended. DeepFeature enables a powerful means for the identification of biologically relevant gene sets and provides methodological basement for "explainable AI"³⁰. This has been integrated with DeepInsight-3D to simultaneously identify elements for multi-layered data.

Functional annotation of gene sets found to be important for particular classifications was largely in line with pathways and mechanisms previously reported in the literature. Notably, despite very small training sets, there was remarkable degree of consistency across discovered functional categories among different classifiers. These functions also corresponded to well-known multidrug resistance mechanisms as well as core cancer pathways that are linked to general "fitness" of cancerous cells. Since drug response is highly complex and determinants are highly drug-specific these processes would not be expected to be fully predictive of drug response on their own. However, an ability to discover this relevant subset of functions across multiple datasets suggests that DeepInsight transformation may prove particularly useful in comparative analysis across multi-omics datasets and lead to discovery of common mechanisms. Another notable pattern in these results was that relatively few functions were completely unique to particular drugs. This is likely due to the way relatively small sample size of the datasets interacts with convolution done when data is presented as image, which would favor more consistently present features in such cases.

In this work, DeepInsight-3D is used for multi-omics datasets. However, the proposed method is not limited to omics data. It can handle different kinds of multi-layered tabular data (as long as the elements and samples of diverse layers are arranged in the same order). This method does not require any specific biological information such as chromosome locations and visualizes non-image data through multi-layered mappings.

Although the results were promising, the severe scarcity of training and test samples hindered getting a reasonable model estimate. The same was true for MOL and Super.FELT methods, as their results were sensitive to parameter tuning. In general, CNN works very well when the samples are sufficiently large. However, this was not the case in the work. Nonetheless, all these methods provided a good platform in this direction. DeepInsight-3D can perform sufficiently well when the sample size is sufficient such as in the case of single-cell analysis. This would be our future direction of work.

Conclusions

The proposed method, DeepInsight-3D, demonstrates how the data-to-image approach for analysis of biological data can effectively incorporate different types of omics data and preserve the explicit connections between these layers by placing them in the same positions but in different channels of an input image. As was demonstrated in our previous work, once converted to image form data becomes suitable for use with image-specific convolutional neural network architectures. This study is the first to use this type of omics integration and likewise the first to apply this type of approach to the problem of personalized cancer drug response prediction. Our results have shown that DeepInsight-3D can outperform previously proposed methods and can also be very powerful way to discover underlying important genes, which can then be interpreted to understand the decisions made by the classifier and also identify key biological processes of potential interest. In future work, we would extend this approach by using a regression layer in place of a classification layer in the CNN net to enable us to utilize continuous output values. We would also look at integrating the transformer model for our future work. Presently, transformer models are predominantly used in NLP areas where input signals are time-based data. We will also look if this can be taken further to time-independent data.

Methods

This section covers the proposed DeepInsight-3D methodology. The model consists of the following constituents (1) image transformation by DeepInsight-3D, (2) ResNet-50 model of CNN architecture, (3) class-based CAM to find activation maps, and 4) element decoder to decode genes (Fig. 1). These procedures are described hereunder.

DeepInsight-3D: conversion of multi-layered tabular data to image for CNN. Let a multi-layered sample be depicted by x_{ij}^l , where i represents elements or features, j represents samples, and l represents layers. Therefore, an l -th layer data can be depicted as $M_l = x_{ij}^l$ for $i = 1, 2, \dots, d$, $j = 1, 2, \dots, n$ and $l = 1, 2, \dots, L$, where d is the dimensionality of the data, n is the number of samples, and L is the total number of layers. For multi-omics data in this work, $L = 3$, which gives a multi-layered dataset $M = \{M_1, M_2, M_3\} \in \mathbb{R}^{d \times n \times L}$. The DeepInsight model²¹ converts non-image data M_l to image data E_l . The size of an image sample is $p \times q$. The DeepInsight transform consists of manifold techniques such as t-SNE³¹, UMAP⁷² or Kernel PCA⁷³, convex hull algorithm, rotation of Cartesian coordinates, finding pixel locations and mapping of elements to these pixel locations. In the case of t-SNE, it constructs a probability distribution over a pair of samples so that similar samples have greater probabilities and dissimilar samples have lower probabilities. Then similarly the probability

distribution is found in the 2D plane. Thereafter, minimization of Kullback–Leibler divergence between the two distributions are performed.

These manifold techniques are not used in a usual manner as it is not required to visualize samples but the genes or elements. In this respect, the transpose of data M_l is used to find P_l . Many of these techniques can project data to a 2D plane (DeepInsight-3D does not require 3D or higher dimensional projection). Therefore if $d > 2$ and $n > 2$ then it is possible to find 2D framework for M_l . We can obtain pixel locations by

$$P_l = \mathcal{H}(M_l) \quad \text{for } l = 1, \dots, L \quad (1)$$

where P_l is the pixel locations of layer l , \mathcal{H} denotes the DeepInsight transform to find pixel locations and M_l is an l -th layer of the training set (e.g. gene expression data). Note, the transpose in Eq. (1) is not explicitly shown. Once the framework of the locations is discovered using Eq. (1), elements can be mapped to find the corresponding images, such as

$$e_j = \Phi(x_{ij}^l) \quad \text{for } j = 1, \dots, n \quad \text{and } i = 1, \dots, d \quad (2)$$

where Φ maps a non-image sample $x \in \mathbb{R}^d$ to an image sample $e_j \in \mathcal{F}^{p \times q}$, here \mathcal{F} is a pixel-coordinates system, and, p and q are sizes of rows and columns, respectively. For simplicity, the superscript l is ignored on e_j . From Eq. (2) we get $\Phi : x \rightarrow e$. The transformation Φ also normalizes the values between $[0, 1]$ or $[0, 255]$. In this work, norm-2 has been employed which was introduced in²¹.

Thus, the first layer of image data ($l = 1$) obtained from Eq. (2) is

$$E_1 = \{e_1, e_2, \dots, e_n\} \quad (3)$$

However, this dataset obtained from Eq. (3) is for layer $l = 1$. For $l = 2$, we did not compute the transform \mathcal{H} , however, only Eq. (2) has been used to find E_2 . Similarly, for $l = 3$, we can obtain the dataset E_3 from Eq. (2). Therefore, for $l = 1, \dots, L$, we get a multi-layered image dataset with common pixel locations P_1 . In this work, $L = 3$, so we get a 3D colored image of a multi-omics sample.

In the above model, it has been assumed that information from layer 1 is more than the other two layers, and that's why all the other samples of the remaining two layers also mapped on P_1 . If it cannot be determined which layer has more information compared to others, then all the layers can be used simultaneously to find common pixel locations. In that case, transform \mathcal{H} for $l = 1, \dots, L$ will be applied. However, it would produce multiple pixel locations ($P_1 \dots P_L$) and we need to find the common pixel locations from these pixel locations. This requires a two-stage process and is implemented in the DeepInsight-3D package by setting up the parameter *Parm.FeatureMap* to '0'. This option has not been used in this work, nonetheless, it has been detailed further in Figure S5 (Supplement File 1).

CNN architecture for classification and feature selection. In this work, ResNet-50 has been used for CNN. For feature selection, we have incorporated class-activation maps (CAMs)³². However, other series nets supported by CAM can be used. ResNet-50 has a fixed input image size of $224 \times 224 \times 3$. However, different image sizes can be used, as package resizes and corrects the size according to the requirements of the net. The last ReLU layer has been used to find activation maps. The activation maps express the region of interest for decision making. It provides 3 colored layers in order of importance as red, yellow and blue. Since the red zone is the most informative, it has been used for feature selection purposes using the element decoder (Fig. 1). The training set and validation set are used to estimate and validate the model. The test set is used to evaluate the performance of the trained model. For CAM, only the training set has to be used to compute activations. The default values of hypermeters of CNN net, such as momentum, L2 regularization and initial learning rate have been used (as per version 2 of the DeepInsight package <https://alok-ai-lab.github.io/DeepInsight/>). However, the Bayesian optimization technique has been employed for Cisplatin to tune the hyperparameters. Further description is given in Supplement File 1.

Class activation maps (CAMs) and element decoder. CAMs are computed for each image sample from the training set e_j . CAM produce 3 colors and if we denote R_j as the computed CAM values of the red zone for a sample e_j , then $R_j > \text{threshold}$ depicts a region of interest for this sample. Since samples e_j falls in different classes (here respondents and non-respondents), we can take an average of R_j over the samples of a class. Therefore, class-based CAM can be computed as

$$\text{avg}R_i = \frac{1}{n_i} \sum_{j \in \omega_i} R_j \quad \text{for } i = 1, \dots, c \quad (4)$$

where ω_i denotes i -th class, c is the number of classes (here 2), and n_i is the number of training samples in this class.

For class-based CAMs, instead of taking $R_j > \text{threshold}$, one can consider $\text{avg}R_i > \text{threshold}$ from Eq. (4). Under this activated region, element decoder finds the gene subset. The decoder will locate the argument or index of a pixel falling under this region. A pixel p_k , located at (a_k, b_k) is defined by normalized value $[0, 1]$. However, depending upon the compression, it may contain one gene, more than one gene, or no gene. Searching all the pixels under the activated region (as defined by Eq. (4)), would reveal a list of selected genes. This procedure will provide class-based features (or genes or elements), however, some elements could be common across different classes.

Let G_i be the gene subset found from the i -th class, then the overall selected genes are denoted as

$$G = \cup_{i=1}^c G_i \quad (5)$$

Experimental setup. We used the same setup of datasets as done in², where training sets were collated from GDSC cell lines resource³. The test sets were collated from TCGA patients with the drug response⁷⁴ and PDX encyclopedia resource⁷.

The data was downloaded from the Zenodo repository (<https://zenodo.org/record/4036592>) and correlated into the seven testing and training using R.

The training sets were from GDSC and the independent test were from PDX/TCGA. To train and validate the model, seven datasets from the GDSC collection are employed. There are 389 training samples overall for Paclitaxel, 844 for Gemcitabine, 856 for Cetuximab, 362 for Erlotinib, 829 for Docetaxel, and 829 for Cisplatin. The training set was split into two smaller training set and a validation set in a ratio of 90:10. The PDX/TCGA resource is used to obtain the independent test sets. The number of test samples for Paclitaxel (PDX) is 43, Gemcitabine (PDX) is 25, Cetuximab (PDX) is 60, Erlotinib (PDX) is 21, Docetaxel (TCGA) is 16, Cisplatin (TCGA) is 66 and Gemcitabine (TCGA) is 57. The training sets were subdivided into 90:10 ratio for training the model and validating the trained model. The separate test set was never used during the training and/or validation process. It has been used at the end to provide the evaluation parameters from the trained model. All the three layers, expression, CNA and mutation, are used to transform into colored images by the Deepinsight-3D model and ResNet-50 is used to find the classification performance.

Various parameters such as t-SNE distance, different types of CNN nets and maximum epochs were estimated using the validation set of a drug dataset (see Table S1, Supplement File 1). The hyperparameters (L2regularization, momentum and initial learning rate) were obtained from DeepInsight version 2 (where the Bayesian optimization technique had been used to tune the hyperparameters on an RNAseq dataset). Thereafter, applied to all the datasets. For Cisplatin, the results were unsatisfactory; therefore, the Bayesian optimization technique was applied with 10 maximum objectives. The hyperparameters corresponding to the best performance objective function on the validation set was used to evaluate the test set (see Table S2, Supplement File 1). The training and validation sets were artificially augmented during the training phase of CNN net (ResNet-50). For augmentation, a sample is generated by averaging any two samples having the same class label.

Test samples have two labels, non-responders (NR) and responders (R). The test set labels are exactly the same as² and are shown in Table S6 (Supplement File 1). The total samples used for training models are also derived from GDSC resource, same as², however, the number of NR and R may be different. For all the training sets, first we applied a median of log₂IC50 to separate NR and R labels. This attempt balanced the NR and R samples in the training sets. However, in the case of Cisplatin, the validation accuracy was not promising, and so we then applied 'mean' to separate the labels.

Table S6 (Supplement File 1) shows that all the datasets used in this paper have gene count between 13,039 and 15,500. On the other hand, the number of training samples is between 362 and 856. This leads to the 'curse of dimensionality' problem or high-dimensional problem. Figure 1 provided a pipeline to solve this problem (and its related discussion) that high-dimensional tabular data is first converted to colored images by DeepInsight-3D. After that, CNN is used to get the prediction.

As it can be observed that the number of samples is very limited for all the drug response data, we augmented the training and validation sets during the training phase of CNN.

All the experiments were done on Intel Xeon Gold 5220R Server (2.2 GHz) with 24 CPU cores and 2 parallel NVIDIA A100 PCIe GPUs (CUDA cores: 6912 with 40 GB GPU memory on each A100 GPU). The operating system used was Linux (Ubuntu Desktop version 20.04).

Pre-processing of mutation data. Cancer mutation data is most often extremely sparse, meaning that only a small number of different genes have consequential mutations in each sample. This presents a unique challenge when using it with a CNN classifier—as most inputs in this channel would be zero, it can result in inefficient use of information in that layer due to “dead” artificial neurons⁷⁵. To counter this, we have used guilt-by-association principle to propagate the likely impact of mutations by using protein–protein interaction network. Briefly, the goal of this approach was to assign some part of an “impact” for each actual mutation to proximal genes in the network, as these are likely to be involved in similar biological functions. In this way, some meaningful value is assigned to each gene in all situations, while the information about actual mutations is still preserved by assigning to them the highest possible score. Note that the fine calibration of the impact score is not necessary for this use-case, as neural network is able to discover its own optimal weighting as long as the generated distribution is consistent across all of the training set.

This was done by mapping all of the genes in the dataset to the corresponding proteins of the protein interaction network obtained from STRING database v11.0⁷⁶. A diffusion state distance matrix was calculated for the network based on the original definition of this distance metric⁷⁷. Then, each node was assigned a score equal to the normalized inverse distance value of the closest mutated gene. In this way, the approach has facilitated the identification of possible functionally equivalent mutations as well as mutation hotspots, which have also been demonstrated to be an important network-based feature potentially predictive of clinical outcomes⁷⁸.

Model evaluation. In order to validate DeepInsight-3D, the training sets (besides the test sets) were subdivided into two sets with a 90:10 ratio. The larger set was employed to estimate the model and the smaller set was applied to validate it. The AUC was computed on the test set. In general, the default parameters of DeepInsight

(version 2) were employed (<https://alok-ai-lab.github.io/DeepInsight/>) for this method with a few variations (see Table S1, Supplement File 1 for details). Some important parameters were norm-2 normalization (log transform)²¹, t-SNE to obtain a 2D plane for gene expression data, and that CNA and mutations were mapped to the 2D plane obtained by gene expression, as it is generally considered that gene expression has more information compared to the other profiles. For CNN, we applied a pre-trained ResNet-50. This transfer learning helped to achieve promising results. In order to have faster training, default parameters were applied for all the datasets, and the obtained performance was satisfactory. However, for Cisplatin, we did not get promising results. Therefore, for Cisplatin, the Bayesian optimization technique of hyperparameter tuning was applied for ResNet-50. The hyperparameters that best performed over the validation set have been used for the test set (see Table S2, Supplement File 1 for details).

Finding gene subsets through an iterative process. The number of genes in genomic or multi-omics data is typically very large, making it difficult to put all of them into a finite image size due to fixed technology limits. In this instance, quantized images are unavoidable, meaning that specific image pixels will carry several genes in a single spot. This leads to another issue of selecting a gene from those batch genes (where batch gene refers to a set of two or more genes having the same pixel location in the frame). To address this overlapping issue up to some extent, DeepInsight-3D can be run iteratively to gradually select the elements. The initial iteration will identify a subset of elements that can be utilized as input in subsequent iterations to find a smaller subset of genes or elements.

Functional annotation and interpretation of identified gene sets. The analysis described above resulted in two gene lists (one each for responder and non-responder class) from each trained model that contained the genes identified as important for classifying training samples into a respective category. Functional interpretation of the recovered gene subsets was done individually, by mapping them onto metabolic and signaling pathways as defined by Reactome. This was followed by gene set enrichment analysis done using Fisher's exact test with a Benjamini–Hochberg false discovery rate correction. The analysis was further extended using Ingenuity Pathway Analysis software from QIAGEN Digital Insights, which is backed by an extensive manually-curated knowledgebase. Owing to its more extensive collection of data and more detailed pathway definitions, IPA has allowed identification of a much larger set of additional significant pathways. This analysis was done both for the metabolic and phosphorylation pathways that were filtered to only include direct links with experimental evidence.

Data availability

All the datasets used in this paper can be downloaded from the Zenodo repository <https://zenodo.org/record/4036592>.

Code availability

DeepInsight-3D software package (in Matlab), a dataset, installation instructions and user-manual are available from the GitHub link https://github.com/alok-ai-lab/DeepInsight3D_pkg. The example PDX_Paclitaxel dataset is also separately available from the link http://emu.src.riken.jp/DeepInsight/download_files/dataset1.mat, note the size is 88 MB. The following links for other related packages can be accessed via <http://www.alok-ai-lab.com/tools.php> and/or <http://emu.src.riken.jp/>.

Received: 21 November 2022; Accepted: 8 February 2023

Published online: 11 February 2023

References

- Marquart, J., Chen, E. Y. & Prasad, V. Estimation of the percentage of US patients with cancer who benefit from genome-driven oncology. *JAMA Oncol.* **4**(8), 1093–1098 (2018).
- Sharifi-Noghabi, H. *et al.* MOLI: Multi-omics late integration with deep neural networks for drug response prediction. *Bioinformatics* **35**(14), i501–i509 (2019).
- Iorio, F. *et al.* A landscape of pharmacogenomic interactions in cancer. *Cell* **166**(3), 740–754 (2016).
- Barretina, J. *et al.* The cancer cell line encyclopedia enables predictive modelling of anticancer drug sensitivity. *Nature* **483**(7391), 603–607 (2012).
- Graim, K. *et al.* PLATYPUS: A multiple—view learning predictive framework for cancer drug sensitivity prediction, *Pac. Symp. Biocomput.* **24**, 136–147 (2019).
- Ballester, P. J. *et al.* Artificial intelligence for drug response prediction in disease models. *Brief Bioinform.* **23**(1), bbab450 (2022).
- Gao, H. *et al.* High-throughput screening using patient-derived tumor xenografts to predict clinical trial drug response. *Nat. Med.* **21**(11), 1318–1325 (2015).
- Park, S., Soh, J. & Lee, H. Super.FELT: Supervised feature extraction learning using triplet loss for drug response prediction with multi-omics data. *BMC Bioinform.* **22**(1), 269 (2021).
- Ma, S. & Zhang, Z. OmicsMapNet: Transforming omics data to take advantage of Deep Convolutional Neural Network for discovery. [arXiv:1804.05283](https://arxiv.org/abs/1804.05283), <https://doi.org/10.48550/arXiv.1804.05283> (2018).
- Joo, M., *et al.* A deep learning model for cell growth inhibition IC50 prediction and its application for gastric cancer patients. *Int. J. Mol. Sci.* **20**(24), 6276 (2019).
- Zhu, Y. *et al.* Converting tabular data into images for deep learning with convolutional neural networks. *Sci. Rep.* **11**(1), 11325 (2021).
- Kobayashi, K. *et al.* Fully-connected neural networks with reduced parameterization for predicting histological types of lung cancer from somatic mutations. *Biomolecules* **10**(9), 1249 (2020).
- Stoean, R. *et al.* Transfer learning with convolutional neural networks for cancer survival prediction using gene-expression data. *Plos One* **15**(3), e0230536 (2020).

14. Tanabe, K. *et al.* Comprehensive serum glycopeptide spectra analysis combined with artificial intelligence (CSGSA-AI) to diagnose early-stage ovarian cancer. *Cancers* **12**(9), 2373 (2020).
15. Buturović, L. & Miljković, D. A novel method for classification of tabular data using convolutional neural networks. bioRxiv. <https://doi.org/10.1101/2020.05.02.074203> (2020).
16. Sua, J. N. *et al.* Incorporating convolutional neural networks and sequence graph transform for identifying multilabel protein Lysine PTM sites. *Chemomet. Intell. Lab. Syst.* **206**, 104171 (2020).
17. Le, N. Q. K. *et al.* DeepETC: A deep convolutional neural network architecture for investigating and classifying electron transport chain's complexes. *Neurocomputing* **375**, 71–79 (2020).
18. Manavi, F. *et al.* CNN-Pred: Prediction of single-stranded and double-stranded DNA-binding protein using convolutional neural networks. *Gene* **853**, 147045 (2023).
19. Ballester, P. J. *et al.* Artificial intelligence for drug response prediction in disease models. *Brief. Bioinform.* **23**(1), bbab450 (2021).
20. Habibi Aghdam, H. & Jahani Heravi, E. *Guide to convolutional neural networks: A practical application to traffic-sign detection and classification.* (Springer International Publishing, Switzerland, 2017).
21. Sharma, A. *et al.* DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture. *Sci. Rep.* **9**(1), 11399 (2019).
22. Kanber, B., *Sparse data to structured imageset transformation.* arXiv:2005.10045, 2020.
23. Laguna, R., *et al.* Early detection of gearbox damage, development of ANN forecast maintenance tool, in *2020 IEEE PES Transmission & Distribution Conference & Exhibition - Latin America (T&D LA)*. 2020. p. 1–6.
24. Pasquadibisceglie, V. *et al.* ORANGE: Outcome-oriented predictive process monitoring based on image encoding and CNNs. *IEEE Access* **8**, 184073–184086 (2020).
25. Rodrigues, N. M. *et al.*, Plotting time: On the usage of CNNs for time series classification. arXiv preprint arXiv:2102.04179v1, 2021.
26. Tajmirriahi, M. *et al.* An interpretable convolutional neural network for P300 detection: Analysis of time frequency features for limited data. *IEEE Sens. J.* **22**(9), 8685–8692 (2022).
27. Rahim, M. A. & Hassan, H. M. A deep learning based traffic crash severity prediction framework. *Accid. Anal. Prevent.* **154**, 106090 (2021).
28. Tran, T. P., *et al.*, DeepInsight-Convolutional Neural Network for Intrusion Detection Systems, in *2021 8th NAFOSTED Conference on Information and Computer Science (NICS)*. 2021. p. 120–125.
29. Peng, T.-C., Dhankhar, N. and Aizawa, T. *1st Place Winning Solution - Hungry for Gold*. 2020, Laboratory for Innovation Science at Harvard: Mechanisms of Action (MoA) Prediction competition.
30. Sharma, A. *et al.* DeepFeature: Feature selection in nonimage data using convolutional neural network. *Brief. Bioinform.* **22**(6), bbab297 (2021).
31. Maaten, L. V. D. & Hinton, G. Visualizing data using t-SNE. *J. Mach. Learn. Res.* **9**, 2579–2605 (2008).
32. Zhou, B., *et al.* Learning deep features for discriminative localization. in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*. 2016.
33. Abend, K. How Convolutional Neural Networks Defy the Curse of Dimensionality: Deep Learning Explained. *TechRxiv.* <https://doi.org/10.36227/techrxiv.18316439.v1>, (2022).
34. Ab, D. A. L. & Seo, M. K. Has the development of cancer biomarkers to guide treatment improved health outcomes?. *Eur. J. Health Econ.* **22**(5), 789–810 (2021).
35. Barzaman, K. *et al.* Breast cancer: Biology, biomarkers, and treatments. *Int. Immunopharmacol.* **84**, 106535 (2020).
36. Vakiani, E. & Solit, D. B. KRAS and BRAF: Drug targets and predictive biomarkers. *J. Pathol.* **223**(2), 219–229 (2011).
37. Yi, M. *et al.* Biomarkers for predicting efficacy of PD-1/PD-L1 inhibitors. *Mol. Cancer* **17**(1), 129 (2018).
38. Yu, C. Y. & Mitrofanova, A. Mechanism-centric approaches for biomarker detection and precision therapeutics in cancer. *Front. Genet.* **12**, 687813 (2021).
39. Echle, A. *et al.* Deep learning in cancer pathology: A new generation of clinical biomarkers. *Br J Cancer* **124**(4), 686–696 (2021).
40. Geelheer, P., Cox, N. J. & Huang, R. S. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *Genome Biol.* **15**(3), R47 (2014).
41. Kursu, M. B. & Rudnicki, W. R. Feature selection with theborutapackage. *J. Stat. Softw.* **36**(11), 1–13 (2010).
42. Liu, B. *et al.* Open source machine-learning algorithms for the prediction of optimal cancer drug therapies. *PLoS ONE* **12**(10), e0186906 (2017).
43. Gillespie, M. *et al.* The reactome pathway knowledgebase 2022. *Nucl. Acids Res.* **50**(D1), D687–D692 (2022).
44. Krämer, A. *et al.* Causal analysis approaches in Ingenuity Pathway Analysis. *Bioinformatics* **30**(4), 523–530 (2014).
45. Yang, P.-L. *et al.* STAT3, the challenge for chemotherapeutic and radiotherapeutic efficacy. *Cancers* **12**(9), 2459 (2020).
46. Liu, R. *et al.* PI3K/AKT pathway as a key link modulates the multidrug resistance of cancers. *Cell Death Dis.* **11**(9), 797 (2020).
47. Chen, L. *et al.* Combination of gemcitabine and erlotinib inhibits recurrent pancreatic cancer growth in mice via the JAK-STAT pathway. *Oncol. Rep.* **39**(3), 1081–1089 (2018).
48. Pelletier, S. *et al.* Rho family GTPases are required for activation of Jak/STAT signaling by G protein-coupled receptors. *Mol. Cell Biol.* **23**(4), 1316–1333 (2003).
49. Crosas-Molist, E. *et al.* Rho GTPase signaling in cancer progression and dissemination. *Physiol. Rev.* **102**(1), 455–510 (2022).
50. Sabaawy, H. E. *et al.* JAK/STAT of all trades: Linking inflammation with cancer development, tumor progression and therapy resistance. *Carcinogenesis* **42**(12), 1411–1419 (2021).
51. Deng, L. *et al.* The role of ubiquitination in tumorigenesis and targeted drug discovery. *Signal Transduct. Target Ther.* **5**(1), 11 (2020).
52. Zhitomirsky, B. & Assaraf, Y. G. Lysosomes as mediators of drug resistance in cancer. *Drug Resist. Updat.* **24**, 23–33 (2016).
53. Narayanan, S. *et al.* Targeting the ubiquitin-proteasome pathway to overcome anti-cancer drug resistance. *Drug Resist Updat.* **48**, 100663 (2020).
54. Li, X. *et al.* Autophagy: A novel mechanism of chemoresistance in cancers. *Biomed. Pharmacother.* **119**, 109415 (2019).
55. Wu, D. W. *et al.* Paxillin confers resistance to tyrosine kinase inhibitors in EGFR-mutant lung cancers via modulating BIM and Mcl-1 protein stability. *Oncogene* **35**(5), 621–630 (2016).
56. Brunner, J., Ragupathy, S. & Borchard, G. Target specific tight junction modulators. *Adv. Drug Deliv. Rev.* **171**, 266–288 (2021).
57. Martin, T. A. & Jiang, W. G. Loss of tight junction barrier function and its role in cancer metastasis. *Biochim Biophys. Acta* **1788**(4), 872–891 (2009).
58. Yu, Q. *et al.* Extracellular matrix proteins confer cell adhesion-mediated drug resistance through integrin alpha v in glioblastoma cells. *Front. Cell Dev. Biol.* **9**, 616580 (2021).
59. Sethi, T. *et al.* Extracellular matrix proteins protect small cell lung cancer cells against apoptosis: A mechanism for small cell lung cancer growth and drug resistance in vivo. *Nat. Med.* **5**(6), 662–668 (1999).
60. Platten, M. *et al.* Tryptophan metabolism as a common therapeutic target in cancer, neurodegeneration and beyond. *Nat. Rev. Drug Discov.* **18**(5), 379–401 (2019).
61. Opitz, C. A. *et al.* The therapeutic potential of targeting tryptophan catabolism in cancer. *Br. J. Cancer* **122**(1), 30–44 (2020).
62. Mellman, I. & Yarden, Y. Endocytosis and cancer. *Cold Spring Harb. Perspect. Biol.* **5**(12), a016949 (2013).
63. Kaksonen, M. & Roux, A. Mechanisms of clathrin-mediated endocytosis. *Nat. Rev. Mol. Cell Biol.* **19**(5), 313–326 (2018).

64. Lee, J. K. *et al.* Pharmacogenomic landscape of patient-derived tumor cells informs precision oncology therapy. *Nat. Genet.* **50**(10), 1399–1411 (2018).
65. Hutter, C. & Zenklusen, J. C. The cancer genome atlas: Creating lasting value beyond its data. *Cell* **173**(2), 283–285 (2018).
66. Gavan, S. P., Thompson, A. J. & Payne, K. The economic case for precision medicine. *Expert Rev. Precis. Med. Drug. Dev.* **3**(1), 1–9 (2018).
67. Mishra, A. & Verma, M. Cancer biomarkers: Are we ready for the prime time?. *Cancers (Basel)* **2**(1), 190–208 (2010).
68. Geeleher, P., Cox, N. J. & Huang, R. S. Clinical drug response can be predicted using baseline gene expression levels and in vitro drug sensitivity in cell lines. *Genome Biol.* **15**(3), R47 (2014).
69. Bansal, M. *et al.* A community computational challenge to predict the activity of pairs of compounds. *Nat. Biotechnol.* **32**(12), 1213–1222 (2014).
70. Menden, M. P. *et al.* Community assessment to advance computational prediction of cancer drug combinations in a pharmacogenomic screen. *Nat. Commun.* **10**(1), 2674 (2019).
71. Chen, Y. & Zhang, L. How much can deep learning improve prediction of the responses to drugs in cancer cell lines? *Brief. Bioinform.* **23**(1), bbab378 (2022).
72. McInnes, L., Healy, J. & Melville, J. UMAP: Uniform Manifold Approximation and Projection for Dimension Reduction. [arXiv:1802.03426](https://arxiv.org/abs/1802.03426), <https://doi.org/10.48550/arXiv.1802.03426>, (2018).
73. Schölkopf, B., Smola, A. & Müller, K.-R. Nonlinear component analysis as a Kernel Eigenvalue problem. *Neural Comput.* **10**(5), 1299–1319 (1998).
74. Ding, Z., Zu, S. & Gu, J. Evaluating the molecule-based prediction of clinical drug responses in cancer. *Bioinformatics* **32**(19), 2891–2895 (2016).
75. Sharma, O. A new activation function for deep neural network. in *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*. 2019.
76. Szklarczyk, D. *et al.* The STRING database in 2021: Customizable protein-protein networks, and functional characterization of user-uploaded gene/measurement sets. *Nucl. Acids Res.* **49**(D1), D605–D612 (2021).
77. Cao, M. *et al.* Going the distance for protein function prediction: A new distance metric for protein interaction networks. *PLoS ONE* **8**(10), e76339 (2013).
78. Reyna, M. A., Leiserson, M. D. M. & Raphael, B. J. Hierarchical HotNet: identifying hierarchies of altered subnetworks. *Bioinformatics* **34**(17), i972–i980 (2018).

Acknowledgements

This work was funded by MEXT/JSPS KAKENHI Grant Number JP20H03240.

Author contributions

AS perceived, designed the classification and feature selection models, and wrote the first draft and contributed in the subsequent versions of the manuscript. AL designed the enrichment analysis, improving selection models and contributed in the first draft and thereafter of the manuscript. KAB aided in the enrichment analysis, prepared the data, feature selection tools and helped in the manuscript writeup. TT perceived and contributed in the manuscript writeup. All authors read and approved the manuscript.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary Information The online version contains supplementary material available at <https://doi.org/10.1038/s41598-023-29644-3>.

Correspondence and requests for materials should be addressed to A.S., A.L. or T.T.

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.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, 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 licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence 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. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2023