Optimization of a Deep-Learning Method Based on the Classification of Images Generated by Parameterized Deep Snap a Novel Molecular-Image-Input Technique for Quantitative Structure–Activity Relationship (QSAR) Analysis

Yasunari Matsuzaka and Yoshihiro Uesawa*

Department of Medical Molecular Informatics, Meiji Pharmaceutical University, Tokyo, Japan

Numerous chemical compounds are distributed around the world and may affect the homeostasis of the endocrine system by disrupting the normal functions of hormone receptors. Although the risks associated with these compounds have been evaluated by acute toxicity testing in mammalian models, the chronic toxicity of many chemicals remains due to high cost of the compounds and the testing, etc. However, computational approaches may be promising alternatives and reduce these evaluations. Recently, deep learning (DL) has been shown to be promising prediction models with high accuracy for recognition of images, speech, signals, and videos since it greatly benefits from large datasets. Recently, a novel DL-based technique called DeepSnap was developed to conduct QSAR analysis using three-dimensional images of chemical structures. It can be used to predict the potential toxicity of many different chemicals to various receptors without extraction of descriptors. DeepSnap has been shown to have a very high capacity in tests using Tox21 quantitative qHTP datasets. Numerous parameters must be adjusted to use the DeepSnap method but they have not been optimized. In this study, the effects of these parameters on the performance of the DL prediction model were evaluated in terms of the loss in validation as an indicator for evaluating the performance of the DL using the toxicity information in the Tox21 qHTP database. The relations of the parameters of DeepSnap such as (1) number of molecules per SDF split into (2) zoom factor percentage, (3) atom size for van der waals percentage, (4) bond radius, (5) minimum bond distance, and (6) bond tolerance, with the validation loss following quadratic function curves, which suggests that optimal thresholds exist to attain the best performance with these prediction models. Using the parameter values set with

the best performance, the prediction model of chemical compounds for CAR agonist was built using 64 images, at 105° angle, with AUC of 0.791. Thus, based on these parameters, the proposed DeepSnap-DL approach will be highly reliable and beneficial to establish models to assess the risk associated with various chemicals.

INTRODUCTION

The traditional human-safety assessment of chemical compounds involves repetitive-dosage subacute toxicity testing *in vivo* using animal models. However, the risk remains that such compounds could pose major public health concerns to humans by potentially disrupting normal endocrine functions with various hormone receptors upon long-term exposure (Genuis and Kyrillos, 2017; Heindel et al., 2017; Manibusan and Touart, 2017; Sifakis et al., 2017; Tapia-Orozco et al., 2017; Heindel, 2018; Marty et al., 2018). However, since some molecular mechanisms differ between species and depend on environmental factors, it is often difficult to apply the outcomes of animal testing to predict the effects on human health (Brockmeier et al., 2017; Leist et al., 2017; Fay et al., 2018). Moreover, a large number of chemical substances need to be studied to identify the adverse effects on development, metabolic homeostasis, reproduction, cytotoxicity, etc. (Zhu et al., 2014; Bell et al., 2017; Insel et al., 2017; Juberg et al., 2017; Clark and Steger-Hartmann, 2018; Mortensen et al., 2018). Thus, high-throughput (HTP) assays and economical methods are required (Tollefsen et al., 2014; Chen et al., 2015; Wang et al., 2015; Richard et al., 2016). Alternative computational prediction methods based on *in-silico* experiments are essential for conducting safety evaluations of high-risk chemical substances (Malloy et al., 2017; Lo et al., 2018; Luechtefeld et al., 2018; Zhang et al., 2018). Among these, quantitative structure–activity relationship (QSAR) analysis can predict physiological activity, toxicity, enzymatic reactions, receptor agonist/antagonist activity, environmental fate, etc. (Bloomingdale et al., 2017; Polishchuk, 2017; Halder et al., 2018; Khan and Roy, 2018; Simões et al., 2018). This analysis is conducted based on a formulation of established rules for the relationship between the chemical structure of a compound and its activity and relies on the structural, quantum chemical, and physicochemical features, which are represented as various numerical molecular descriptors (Dougall, 2001; Fang et al., 2003; Roy and Das, 2014; Silva and Trossini, 2014). However, there are limited programs that can precisely evaluate the response patterns of cellular signaling molecules due to various chemical compounds.

These days, machine learning has been applied in extensive toxicological fields, and it is highly effective for risk assessment (Ambe et al., 2018; Banerjee et al., 2018; Luechtefeld et al., 2018; Cipullo et al., 2019). More recently, deep learning (DL), a machine-learning method designed to extract and recognize discriminative information patterns and rules, has been proposed to identify features by several flexible fully-connected layers of a neural network (NN) (Li S. et al., 2017; Qiu et al., 2017;

Hu et al., 2018; Li H. et al., 2018; Luechtefeld et al., 2018; Mayr et al., 2018). Until today, support vector machine, random forest, and artificial NN were needed to select a reasonable combination of features (corresponding to chemical structure descriptors in QSAR analysis) manually when learning (feature selection techniques). In many cases, it is extremely difficult to find the optimal solutions, since myriad (Manallack et al., 2010; Talevi et al., 2012; Guimarães et al., 2016; Fang et al., 2017). Therefore, various approximation methods have been developed to obtain an optimal combination for an approximate solution (Yap et al., 2007; Kulkarni et al., 2009). However, since there is no completely trustworthy approximation method, complicated craftsmanship procedures are required to extract effective features in conventional machine learning.

On the other hand, a convolutional neural network (CNN) that constitutes DL has a function of feature expression learning that makes it automatically extract features and unnecessary to manually extract features (Fernandez et al., 2018; Lumini and Nanni, 2018). Unlike the conventional method, which is essential for extraction of a molecular structure descriptor, it is able to identify the most informative features required automatically, which is useful for prediction from the input information of the entire molecule "without supervision" by hierarchically decomposing an image so that the CNN learns to recognize higher-quality features while maintaining their spatial relationships (Ma et al., 2015; Ragoza et al., 2017; Xu et al., 2017; Ghasemi et al., 2018; Liu R. et al., 2018; Peng et al., 2018). These layer structures of the DL consist of input, hidden intermediate, and output layers of a NN, which is an algorithm designed for pattern recognition where information flows and is referred to as a deep neural network (DNN) (LeCun et al., 2015; Mallat, 2016; Suárez-Paniagua and Segura-Bedmar, 2018; Voulodimos et al., 2018). In this DNN, it is possible to directly learn feature quantity contained in a large amount of input data without human intervention at each layer (Azimi et al., 2018). Moreover, it poses a capacity to improve the prediction accuracy for very complicated image recognition by increasing the information transmission and processing ability using a large number of hidden layers and some techniques such as dropout, data augmentation, Rectified Linear Units (ReLUs), and multiple graphics processing units (GPUs) (Rawat and Wang, 2017; Gawehn et al., 2018; Ha et al., 2018; Hussain et al., 2018; Poernomo and Kang, 2018; Qiao et al., 2018; Saha et al., 2018; Sato et al., 2018; Shen et al., 2018; Steven and Han, 2018; Tustison et al., 2018; Vakli et al., 2018; Wang S. H. et al., 2018). Therefore, it is also possible to cope with the deviation and the deformation of the position of input image data for detecting on the edge region

(Krizhevsky et al., 2012). However, since the result depends on the size of the filter, the moving width, and settings such as padding (the process of filling that allocates the end of region with 0 to pad out the number of convolutions of the edge region of the image) (Szegedy et al., 2014; Johnson and Zhang, 2015). In addition, CNNs appropriate combinations of extracted constituent elements and data orderly to the next layer, so it is possible to efficiently learn feature quantities (Szegedy et al., 2014; Cagli et al., 2017).

Studies have reported very high prediction accuracy DL with highly non-linear hierarchical patterns based on large-scale data, especially in the fields of imaging and toxicology (LeCun et al., 2015; Ma et al., 2015; Mayr et al., 2016; Pastur-Romay et al., 2016; Zhang et al., 2017). In addition, some studies have demonstrated the use of DL in QSAR analysis to calculate feature values from molecular structures without human intervention that three steps: (1) model building from labeled data inputs, (2) evaluation and tuning of the model, and (3) training the final model to perform prediction (Bengio et al., 2013; LeCun et al., 2015; Ma et al., 2015; Mayr et al., 2016; Pastur-Romay et al., 2016; Pham et al., 2017; Zhang et al., 2017). However, since for delivering information on the whole molecule sufficiently established most of the cases where DL is applied to QSAR on conventional descriptor calculation at present. Therefore, further work is required to increase prediction accuracy for applications DL for QSAR analysis. First, a systematic and suitable input is required for complicated data such as the three-dimensional (3D) structures of chemical compounds. Moreover, as a result of the insufficient amount of chemical compounds, there is a lack of training data. To address these issues, a novel QSAR model using DL based on 3D molecular images of chemical compounds was previously developed (Uesawa, 2018).

Deep Snap is a procedure of generating an omnidirectional snapshot portraying 3D structures of chemical compounds using a drawing software (Jmol; Hanson, 2016) based on the Structure Data File (SDF) format (**Figure 1**). The 3D information is input into the DL model without calculating structural descriptors. For example, when the 3D molecular structure is rotated in 45° increments on the x-, y-, and z-axes and photographed, a total of 512 images are captured for each molecule and saved in the portable network graphics (PNG) format. This allows for combining digital information regarding the 2D plane location of the atoms with pixel-level data representing the three primary colors (RGB) (**Figure 1**; Uesawa, 2018). Then, these images are used in inputs of the DL model after a resolution of 256 × 256 pixels images of the 3D molecular structure are represented as a ball-and-stick model for each atomic composition with different colors representing different atoms (Uesawa, 2018). We refer to this omnidirectional snapshot capturing procedure for 3D structures of compounds as "Deep Snap."

In the Tox21 data challenge in 2014, a crowd-sourced QSAR competition for chemical risk assessment held by the National Institutes of Health (NIH) in the United States (Tox21 Data Challenge., 2014), approximately 7,000–9,000 different chemical structures depending on the target type. This data was split evenly into training and validation datasets (a 50% of training and a 50% of validation) that were created for the purpose of developing high-performance prediction models for various adverse-outcome pathways (Attene-Ramos et al., 2013; Tox21 Data Challenge., 2014. Recently, using a set of these chemicals (containing a total of 7,320 different molecules with 3,660 reserved for training and 3,660 reserved for validation), the Deep Snap procedure was applied to successfully predict which chemical compounds disrupt the potential of the mitochondrial membrane (MMP), which play pivotal roles in apoptosis, oxidative phosphorylation, calcium homeostasis, and cellular metabolism such as heme, fatty acid, and steroid synthesis (Midzak et al., 2011; Hua et al., 2012; Bolisetty et al., 2013; Shaughnessy et al., 2014; Li A. X. et al., 2017; Liu et al., 2017; Yun et al., 2017; Wang C. et al., 2018). Individual compounds well-known inhibitors for complex between uncouplers (e.g., Carbonyl cyanide-p-trifluoromethoxyphenylhydrazone: FCCP) and particular protein/complex in the transporter chain (rotenone and antimycin A) have been detected in 76 structurally related clusters from the Tox21 10K library (Attene-Ramos et al., 2015; Xia et al., 2018). As potential mitochondrial toxicants, these compounds were found to cause significant reduction of the MMP using an MMP assay in HepG2 cells and rat hepatocytes (Attene-Ramos et al., 2015; Xia et al., 2018). Using transfer learning techniques and an unmodified version of the AlexNet network, the prediction model developed by the Deep Snap-DL method showed area under the ROC curve (AUC) value of 0.921 in the external validation, which included only 647 of the chemical structures employed previously by the Tox 21 Data Challenge 2014 (Uesawa, 2018). At the Tox 21 Data Challenge 2014 competition, the best AUC = 0.95 (Abdelaziz et al., 2016). The prediction performance (AUC = 0.921) by the Deep Snap-DL method is equal to top 10th in the Tox 21 Data Challenge 2014 competition (Tox21 Data Challenge., 2014; Uesawa, 2018. The result suggests that the DL approach based on Deep Snap is suitable for modeling to support toxicological assessments. However, further improvements are required for speed, automation, optimization, and efficiency. Despite the requirement for these improvements, herein, we examine the parameters for Deep Snap and DL to characterize how they affect the DNNs.

MATERIALS AND METHODS

Data

Chemical substance profiles for cellular toxicity were collected from the publicly available Tox21 10K chemical library, 12,500 chemical substances, including pesticides, industrial, food-use, and drugs, procured from commercial sources screened by the Toxicology in the 21st Century (Tox21) program, a multi-agency collaboration between the U.S. Environmental Protection Agency, the National Institute of Environmental Health Sciences, National Toxicology Program, NIH Chemical Genomics Center, National Center for Advancing Translational Sciences, and the US Food and Drug Administration (1) incorporate advances in molecular systems by identifying patterns of chemical compounds-induced biological response, (2) prioritize compounds for more extensive toxicological evaluation, and (3) develop predictive models for biological response in human
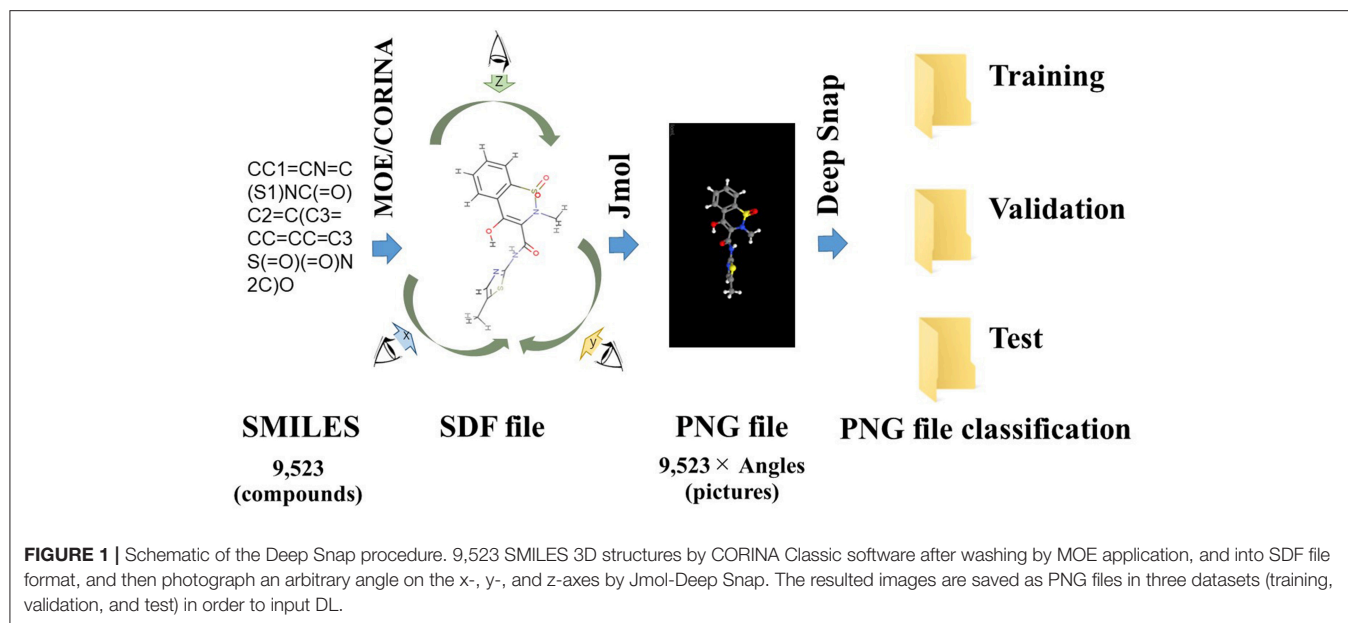
FIGURE 1 | Schematic of the Deep Snap procedure. 9,523 SMILES 3D structures by CORINA Classic software after washing by MOE application, and into SDF file format, and then photograph an arbitrary angle on the x-, y-, and z-axes by Jmol-Deep Snap. The resulted images are saved as PNG files in three datasets (training, validation, and test) in order to input DL.

(NRC., 2007 Collins et al., 2008; Kavlock et al., 2009; Huang et al., 2011, 2014, 2016; Attene-Ramos et al., 2013; Tice et al., 2013; Chen et al., 2015; Hsieh et al., 2015, 2017; Merrick et al., 2015; Huang and Xia, 2017; Sipes et al., 2017). Their structures and the corresponding activities were used to determine agonist of a constitutive androstane receptor (CAR; NR1l3), which is a member of the ligand-activated superfamily of nuclear receptors transcriptionally activated genes predominantly expressed in the liver such as *CYP2B6* and *CYP3A4* involved in not only all phases of drug metabolism, transport, detoxification, and disposition about 50% of the drug metabolization potential in the body but also energy metabolism, tumor progression, cholesterol homeostasis, and glucose metabolism (Qatanani and Moore, 2005; Kobayashi et al., 2015; McMahon et al., 2018).

Deep Snap Procedure: Creation of Molecular Image Files

A total of 9,667 of the chemical structures and the corresponding labeled activity scores were downloaded in the SMILES (Simplified molecular input line entry system) format (Weininger, 1988; Putz and Dudaş, 2013; Achary, 2014; Kumar and Chauhan, 2018) from the PubChem database (AID 1224892) derived from Tox21 10k library, the activity scores defined as the Pubchem_activity_scores (zero and scores between 1 and 100 were represented as inactive and active compounds, respectively, by cell viability and agonist activity screenings of the CAR signaling pathway). Then, by eliminating non-organic compounds, a total of 9,523 of the chemical compounds were selected (**Table 1**; **Supplementary Table 1**). After structure cleaning and standardization (removing salts, counterions, and fragments) by conformational import that is a high-throughput conformer generation method for large numbers of molecules using the MOE application software program (but no treatment of protonation states) (Chen and Foloppe, 2008; Molecular Operating Environment, Chemical

Computing Group, Canada) (**Supplementary Table 1**), one 3D chemical structure per compound which have "rotatable torsions" was curated and optimized to generate a single low energy conformation using CORINA Classic software (Molecular Networks GmbH, Nürnberg, Germany, https://www.mn-am.com/products/corina) has been licensed in the past to predict 3D structures for some of the molecules in the main large public databases of small molecules such as PubChem a data-based commercial 3D molecular model builder with high accuracy and high speed for the 3D-structures of organic and metal-organic (also known as organometallic) molecules high coverage for nearly all organics but approximately half of the organometallics (Sadowski et al., 1994; Reitz et al., 2004; Tetko et al., 2005; Renner et al., 2006; Wang et al., 2009; Schwab, 2010; Andronico et al., 2011; Sayers et al., 2018; 3D Structure Generator CORINA Classic., 2019). Finally, these chemical structures were converted to the SDF file format. During the Deep Snap process, when the number of molecules described in the SDF file is large, the power required for the describing. Therefore, in order to improve the depiction speed, it is possible to multiple processes to be executed simultaneously by partitioning of the input data. The size of PNG file is different depending on the number of per SDF file. Moreover, the csv file including annotation data numbers, activity score, and dataset types that was divided randomly into training (4,761 chemicals), validation (2,381 chemicals), and testing (2,381 chemicals) datasets (**Table 1**; **Supplementary Table 1**) was used as the source for labeling each sample. Since the 3D-chemical structures can rotate 360° on each snapshots were captured at a range of fixed increments based on the SDF molecular structure file and the using a novel technique to capture generated images by their description function without human intervention saved as 256 × 256 (pixels resolution) PNG files (RGB) organized by their annotation data numbers (**Figure 1**). In this study, the 3D structure data was preliminarily portrayed as ball-and-stick

TABLE 1 | Number of chemical compounds in train, validation, and test datasets used in optimization of parameter of Deep Snap.

Activity score	Training	Validation	Test	Sum
0: Non-toxic	3,651	1,858	1,878	7,387
1: Toxic	1,110	523	503	2,136
Sum	4,761	2,381	2,381	9,523

structures in four types of increments on the x-, y-, and z-axes: first was (0,0,0), second was (0,0,0), (0,90,0), and (0,0,90), third was (0,0,0), (180,0,0), (0,180,0), and (0,0,180), fourth was (0,0,0), (180,0,0), (0,180,0), (0,0,180), (0,180,180), (180,0,180), (180,180,0), and (180,180,180) included 4 overlapped images automatically and manually obtained from the Deep Snap process, respectively to assess the systematic and suitable input of the 3D structures of chemical compounds and optimization Deep Snap (**Figures 2A–H**). The 3D ball-and-stick model with different colors to different atoms represented by which uses a unique algorithm to calculate surfaces (Jmol, Herráez, 2006; Cammer, 2007; Hanson, 2016; Scalfani et al., 2016; Hanson and Lu, 2017). More detailed technical information is available at the Jmol website[1] As for the depiction process in Deep Snap, it is possible to design a setting cfg file that can specify arbitrary of the Jmol script such as image pixel size, image format (png or jpg), number of molecules per sdf file to split into (MPS), zoom factor (ZF, %), atom size for van der waals (AT, %), bond radius (BR) (mÅ), minimum bond distance (MBD), bond tolerance (BT), etc. Finally, using 64 pictures 105° angle and (MPS:100, ZF:100, AT:23, MBD:0.4, BT:0.8) as permutation test to assess non-specific activity score, they were randomly reassigned based on the activity scores without changing training, validation, and test datasets. Using a total of 10 different datasets, the prediction models were constructed by Deep Snap-DL method with the parameter values for the best performance optimized in this study eight pictures at 180° angle.

Machine-Learning Models Based on DL

All the two-dimensional (2D) images contained digitized information data about plane configuration and the corresponded to the type of atom for the chemical structure produced by Deep Snap were resized by DIGITS version 4.0.0 software to a fixed resolution of 256 × 256 pixels and input into DL model to build the prediction models, which were trained based on the activity scores of chemical compounds and the corresponding 2D chemical-structure images. In this study, the total number of training epochs was 30, snapshot interval in epochs 1, validation interval in epochs 1, random seed 1, solver type stochastic gradient descent, base learning rate 0.01. Training, testing, and validation were performed using the dataset described in **Table 1** and **Supplementary Table 2**. Finally, the performance of the prediction model was evaluated using one test dataset not used for validation. For the DL, a pre-trained implemented the open-source DL framework was used to build

---

[1]Jmol: An Open-Source Java Viewer for Chemical Structures in 3D. Available online at: http://www.jmol.org/

and train the DL models transfer learning (Jia et al., 2014). AlexNet is a convolutional neural network constructed by the University of Toronto (Krizhevsky et al., 2012). The fundamental architecture of this CNN constituted eight pre-trained layers, including five convolutional/pooling that convolution of feature volume and reduces layers by compressing images using max pooling compresses by selecting the maximum value in each region as a representative value convolutional/pooling layer I converts the previous volume (224 × 224 × 3) to (11 × 11 × 3) convolutional/pooling layer II converts the result of layer I to (5 × 5 × 48) convolutional/pooling layer III converts the result of layer II to (3 × 3 × 256) convolutional/pooling layer IV converts the result of layer III to (3 × 3 × 192) convolutional/pooling layer V converts the result of layer IV to (3 × 3 × 192) fully-connected layers that make final connections between feature values and force to zero to suppress overfitting (dropout) total 4,096 neurons. Since AlexNet has 60 million parameters, their optimization was essential to avoid overfitting (**Figure 3**; Krizhevsky et al., 2012; Szegedy et al., 2014; Cagli et al., 2017; Rawat and Wang, 2017; Aggarwal et al., 2018; Ha et al., 2018; Vakli et al., 2018). The non-saturating nonlinearity $f(x) = \max(0, x)$ as the function instead of such as sigmoid function $f(x) = (1+e^{-x})^{-1}$ or $f(x) = \tanh(x)$ because the training time with gradient descent ReLUs much faster than that associated with if the input is negative, there is no contribution to other units (Nair and Hinton, 2010; Krizhevsky et al., 2012; Elfwing et al., 2018; Saha et al., 2018; Wang S. H. et al., 2018). Furthermore, adding a layer of local response normalization (LRN) between the pooling layer and the convolutional layer increases accuracy. The LRN is capable of handling a large number of CNNs with a large learning capacity that can be controlled by varying their assumptions about the nature of images that (1) the locality of pixel dependencies and (2) the stationarity of statistics.

The loss, which is a summation (not a percentage) of the errors in each dataset as shown below cross entropy error (CEE) with respect to the model's parameters by changing the weight vector values, in construction of the prediction models is calculated on training and validation datasets, where pi and yi correspond to the accuracy label (ground truth vector) and output of softmax (estimate values taken direct from the last layer output) for class *i,* respectively.
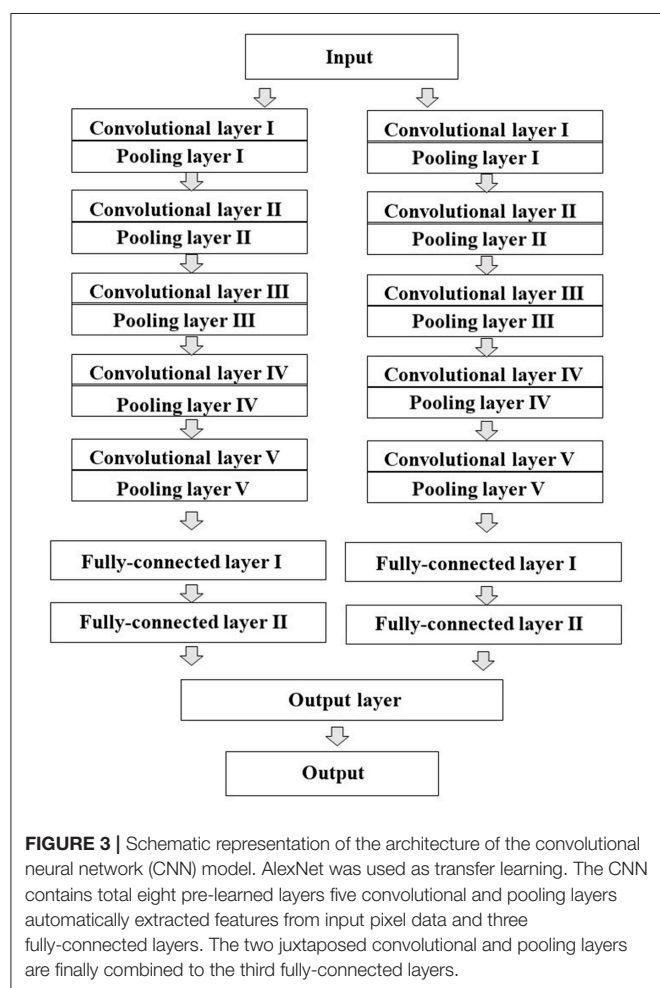
$$CEE = -\Sigma \ (pi) \log(yi)$$

The loss value implies how well or poorly a certain model behaves after each iteration of optimization. Loss is indicative of unless the model has over-fitted with respect to the training data. The accuracy of the model is usually determined after the validation samples are fed to the model and the number of mistakes (zero-one loss) that the model makes recorded. The percentage of misclassification is calculated (Martinez and Stiefelhagen, 2018; Nguyen et al., 2018; Zhang and Sabuncu, 2018; Khened et al., 2019).

Evaluation of the Predictive Models

In this method, it is possible to calculate the prediction result for each of a plurality of images prepared from the x-, y-, and z-axis directions with respect to one molecule. Therefore, the

FIGURE 2 | (A–H) are representative images captured by rotating the 3D structure in 180° increments on Deep Snap. The numbers below the images are the substance identification numbers (SID) provided in the PubChem database and increments of the viewing direction on the x-, y-, and z-axes. Red, yellow, blue, white, and gray colors in the molecular structures indicate the oxygen, sulfur, nitrogen, hydrogen, and carbon atoms, respectively.

FIGURE 3 | Schematic representation of the architecture of the convolutional neural network (CNN) model. AlexNet was used as transfer learning. The CNN contains total eight pre-learned layers five convolutional and pooling layers automatically extracted features from input pixel data and three fully-connected layers. The two juxtaposed convolutional and pooling layers are finally combined to the third fully-connected layers.
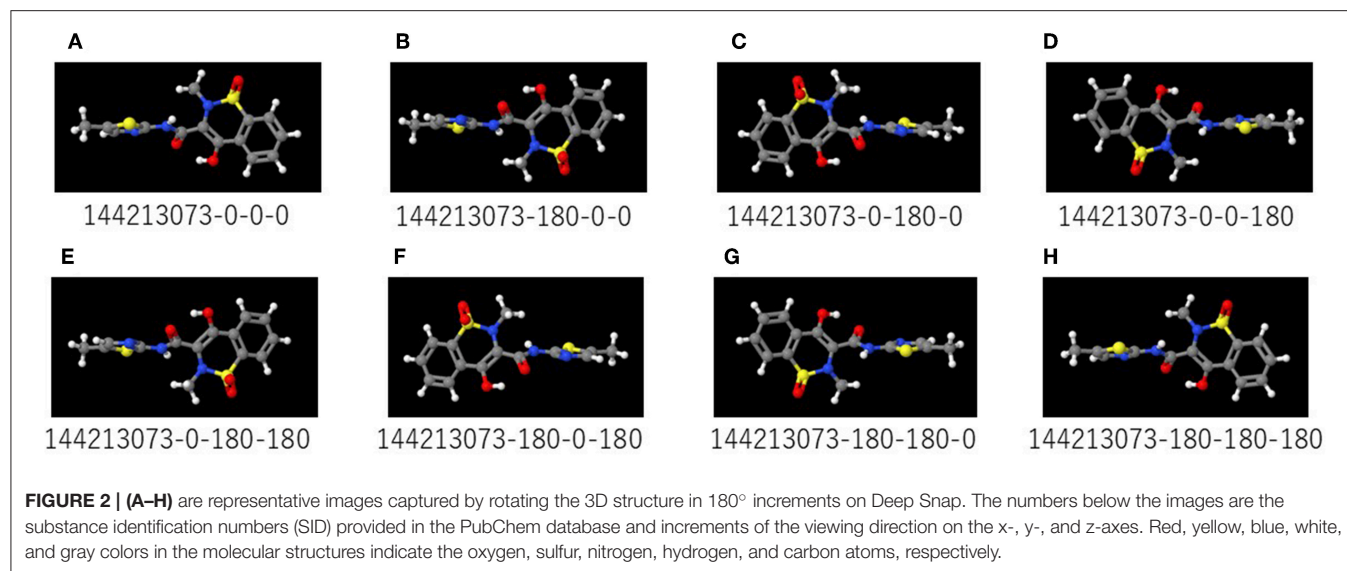
the experimentally determined (true) labels, and the auroc (area under receiver operating characteristic) was calculated using JMP pro 14, statistical discovery software (SAS Institute Inc. NC) to evaluate the predictive models using 3D chemical structures including training (38,088 pictures), validation (19,048 pictures), and testing (19,048 pictures) datasets captured from eight increments on the x-, y-, and z-axes: (0,0,0), (180,0,0), (0,180,0), (0,0,180), (0,180,180), (180,0,180), (180,180,0), and (180,180,180) (**Supplementary Table 2**) (Linden, 2006). Sensitivity describes the true positive rate i.e., the proportion of actual positive samples that were correctly identified as positive for all positive samples including true and false positives.

$$Sensitivity = \Sigma \ True \ Positives/(\Sigma \ True \ Positives + \Sigma \ False \ Positives)$$

Specificity is the true negative rate i.e., the proportion of actual negative samples that were correctly identified as negative for all negative samples including true and false negatives.

$$Specificity = \Sigma \ True \ Negatives/(\Sigma \ True \ Negatives + \Sigma \ False \ Positives)$$

Random Forest

The file, including chemical structures as indicated by SMILES, chemical annotation numbers, activity scores, dataset classes divided into training and validation. Based on this information, the 3D chemical structures were built, descriptors were calculated using the MOE chemical calculation system. Using these descriptors, the prediction model was constructed using JMP pro 14.
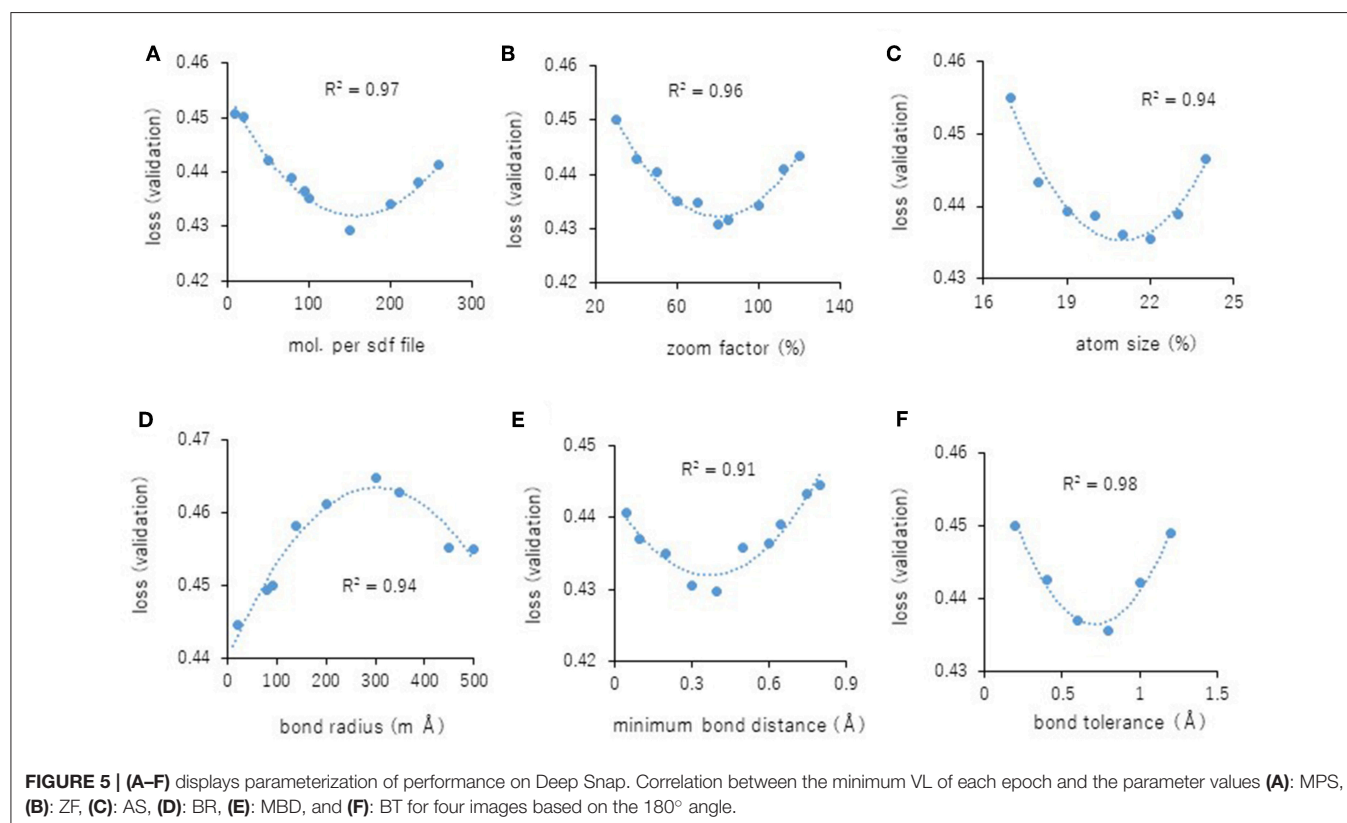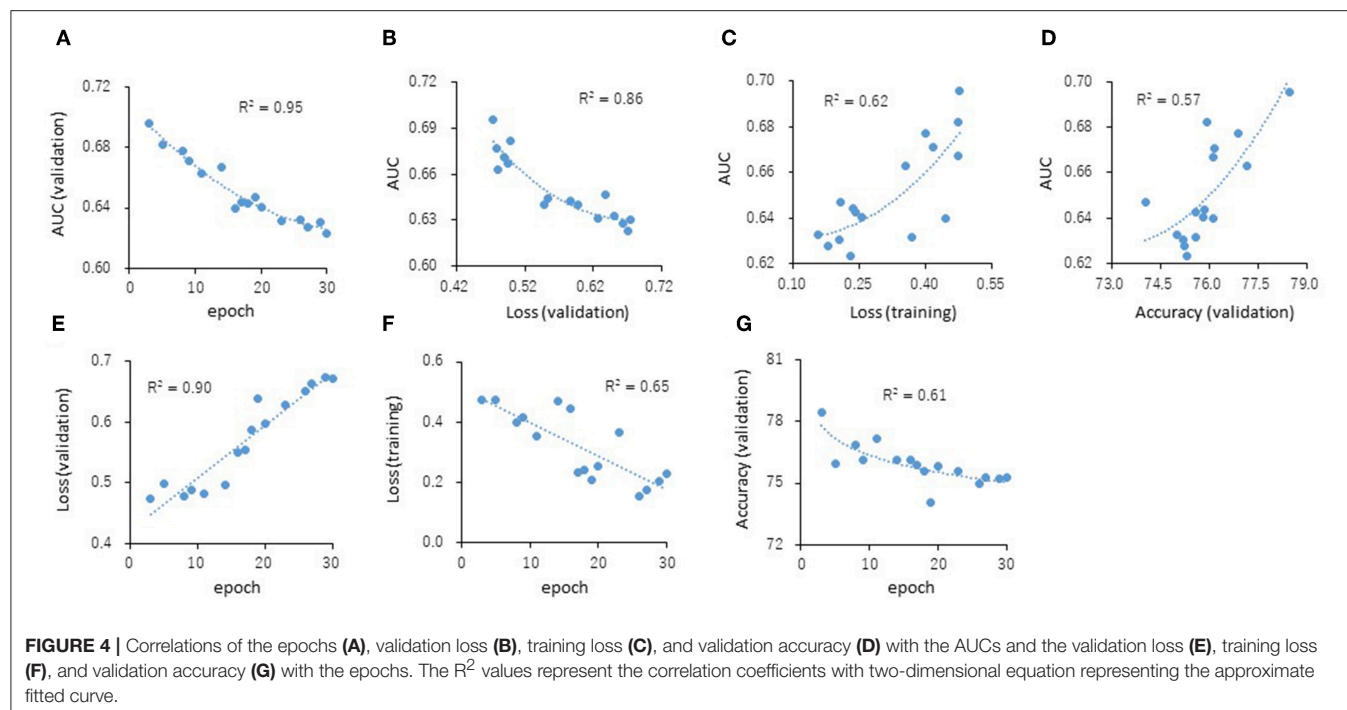
RESULTS AND DISCUSSION

The predictive models for the presence or absence of activity as a CAR agonist and cell viability were built using the open-source Caffe in combination with the Deep Snap approach

median of all these predicted values generated per molecule was used as a representative predicted value for each molecule. The metric was calculated on the basis of the predicted and

were applied to the training (38,088 pictures) and validation (19,048 pictures) datasets 180° angle (**Supplementary Table 2**). The testing dataset (19,048 pictures) was used to measure the performance of each prediction model (**Supplementary Table 2**). The AUC was calculated. The correlations (R^2 values) of the AUC with each epoch were 0.95 (**Figure 4A**). The correlations (R^2 values) were calculated from the testing datasets with validation loss (VL), training loss (TL), and validation accuracy (VA). VL is an error summation not a percentage obtained from how well the model is doing for. TL is an error summation which by attempting to determine good values for all the weights and biases (an empirical risk minimization). VA is the percentage of correct answers based on the results obtained from. As results, these R^2 values with AUCs were 0.86 (VL), 0.62 (TL), and 0.57 (VA), respectively (**Figures 4B–D**). Moreover, the R^2 values of the VL, TL, and VA each epochs were 0.90, 0.65, and 0.61, respectively (**Figures 4E–G**). These findings suggest that VL is the most important parameter of those considered here for evaluating the performance of a DL model.

Next, the parameters for capturing Jmol-generated images on Deep Snap were optimized by assessing the DL models using the same procedure based on the VL using four pictures on the x-, y-, and z-axes: (0,0,0), (180,0,0), (0,180,0), and (0,0,180) in the training (19,044 pictures), validation (9,524 pictures), and test (9,524 pictures) datasets (**Figures 2A–D** and **Supplementary Table 2**). The following parameters were considered: (1) the number of molecules per SDF file: MPS, (2) the zoom factor: ZF, (3) the atom size for Van der Waals interactions: AT, (4) the bond radius: BR, (5) the minimum bond distance: MBD, and (6) the bond tolerance: BT. The parameter values (and corresponding minimum VL values) for the best model are as follows: (1) MPS: 150 (0.430), (2) ZF: 80% (0.431), (3) AT: 22% (0.435), (4) BR: 20 mÅ (0.425), (5) MBD: 0.4 Å (0.430), and (6) BT: 0.8 Å (0.436) (**Figures 5A–F**). In addition, the R^2 values between these parameters and VLs were more than 0.90, and each of these relations followed quadratic function curves. Also, the R^2 values of the running time (RT) in DL with the above six parameters showed that the RTs were moderately associated with AT ($R^2 = 0.48$), BR ($R^2 = 0.47$), and BT ($R^2 = 0.43$) (**Supplementary Figures 1C,D,F**). However, MPS, ZF, and MBD showed no associations (**Supplementary Figures 1A,B,E**). Similarly, the image pixel size (IPS) was examined in the same way as the VL and RT in DL using three pictures on the x-, y-, and z-axes: (0,0,0), (0,90,0), and (0,0,90) in the training (14,283 pictures, 4,761 compounds), validation (7,143 pictures, 2,381 compounds), and test (7,143 pictures, 2,381 compounds) datasets (**Supplementary Table 2**). The IPSs (256×256) and (64×64) exhibited minimum VL (0.440) (**Figure 6A**) and minimum RT (10 min) (**Figure 6B**), respectively. Moreover, the number of cores in the multi-core CPU architecture showed the minimum RT (8 min) in the Jmol-generated images with 70 (**Figure 6C**). Also, we explored the effects of the minimum VL with space-filling, where the atoms are represented by spheres whose radii and center-to-center distances are proportional to the radii of the atoms and the distances between the atomic nuclei using one (0,0,0) or four (0,0,0), (180,0,0), (0,180,0), (0,0,180) image angles (**Figures 2A–D**) on the optimized parameters.
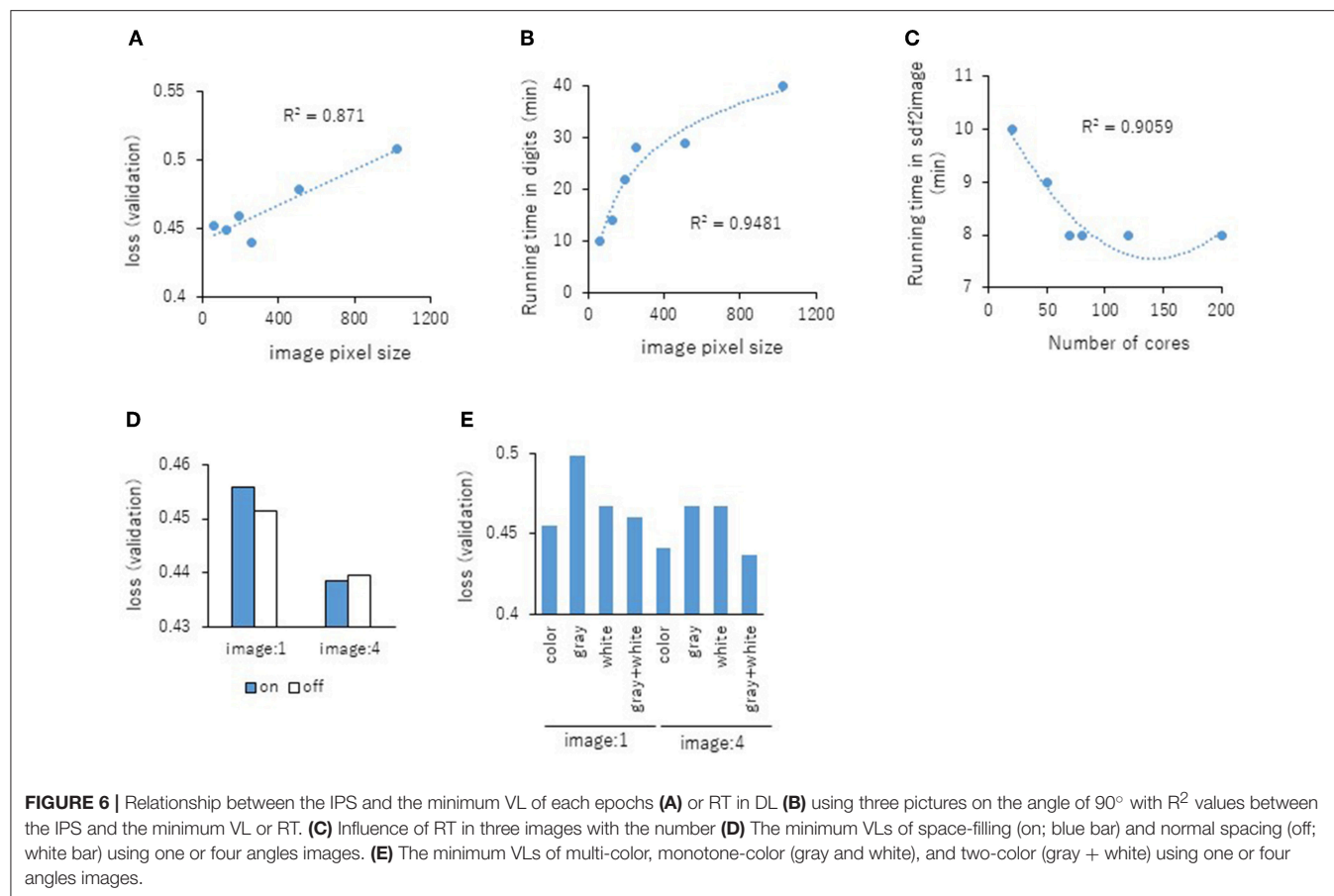
When using one image, space-filling chemical structures into the image slightly increased the minimum VL (0.456) compared with that of normal spacing (0.452) (**Figure 6D**, left). However, there were no minimum VL changes between space-filling and normal spacing when using four image angles (**Figure 6D**, right). Furthermore, we compared the influence of the image color types of chemical structures with the minimum VL by using one or four image angles the optimized parameters, similarly. When the atomic colors of all the structures were changed to monotone (gray or white), these minimum VLs (0.468 or 0.467 for gray and white, respectively) increased to more than that of normal multi-color structures (0.442) using four image angles (**Figure 6E**, right). However, in the structures where the color of all atoms was changed to gray except for hydrogen (two-color: gray + white), the minimum VL (0.437) was decreased slightly compared with that of normal multi-color structures (0.442) using the four images (**Figure 6E**, right). When one angle image was used similarly, increased minimum VL of gray (0.499), white (0.468), or gray + white (0.460) was observed compared with that of normal multi-color (0.455) (**Figure 6E**, left). These findings suggest that optimal thresholds exist to attain the best performance with the prediction model. Finally, using the parameter values for the best performance model, AUCs were calculated using eight images of chemical structures captured at 180° increments on the x-, y-, and z-axes. As a result of optimization, the AUC exhibited 0.764 with minimum VL of 0.432. Furthermore, using 64 images at 105° angle and with default parameter values other than BR 15mÅ, the AUC increased into 0.791.

To assess (1) the suitableness of input as supervised data, (2) sufficient amount of images for training, and (3) adequate training for input dataset of pictures of chemical structure into the DL, the activity scores of the datasets, including training, validation, and test, were randomly assigned keeping the numbers of the three datasets unchanged as permutation test. The calculation of the performed each parameterized values of Deep Snap with each best performance model to capture chemical structures eight pictures at 180° angle using a total of ten different datasets with assignments of various activity scores. As result, the average AUCs were 0.553 (±0.007) with the average minimum VL of 0.522 (±0.014), indicated almost random guessing. These results suggest that the prediction models in this study extracted the CAR agonist activity-specific structural features from chemical compounds. Also, we calculated the AUC random forest as another method the same datasets for the above Deep Snap for CAR agonist and 206 of descriptors to build the prediction model in ROC-AUC value 0.749. Previously, we found that the prediction for the performance of compounds inducing MMP disruption was better 45° angles using 512 pictures for one molecule, with AUCs of 0.921 (Uesawa, 2018). Moreover, using 90° angle which 64 pictures for each, the performance of the prediction model indicated that the ROC-AUC value was 0.898 (Uesawa, 2018). In this study, we have used only 64 pictures based on 105° angle to avoid high computational cost. These results suggested that the prediction performance in the Deep Snap-DL method could be improved by input images due to more information about chemical structures. Also, as for the score

FIGURE 4 | Correlations of the epochs **(A)**, validation loss **(B)**, training loss **(C)**, and validation accuracy **(D)** with the AUCs and the validation loss **(E)**, training loss **(F)**, and validation accuracy **(G)** with the epochs. The R^2 values represent the correlation coefficients with two-dimensional equation representing the approximate fitted curve.

FIGURE 5 | **(A–F)** displays parameterization of performance on Deep Snap. Correlation between the minimum VL of each epoch and the parameter values **(A)**: MPS, **(B)**: ZF, **(C)**: AS, **(D)**: BR, **(E)**: MBD, and **(F)**: BT for four images based on the 180° angle.

activity of the CAR, the chemicals with scores other than 0 were defined as positive in order to secure enough input data in this study. However, in Tox21 program, the obvious activity for the CAR agonist is defined for chemicals with score of more than 40

(PubChem; https://pubchem.ncbi.nlm.nih.gov/#, AID 1224892). Therefore, it is necessary to optimize various types of assignments for the activity scores and/or other datasets in detail to further increase the prediction performance. In addition, a comparison

FIGURE 6 | Relationship between the IPS and the minimum VL of each epochs **(A)** or RT in DL **(B)** using three pictures on the angle of 90° with R^2 values between the IPS and the minimum VL or RT. **(C)** Influence of RT in three images with the number **(D)** The minimum VLs of space-filling (on; blue bar) and normal spacing (off; white bar) using one or four angles images. **(E)** The minimum VLs of multi-color, monotone-color (gray and white), and two-color (gray + white) using one or four angles images.

of the performances between this state-of-the-art Deep Snap and 1,024 of extended-connectivity fingerprint (ECFG) of descriptors calculated from Dragon 7.0 (Kode srl., Pisa, Italy, Rogers and Hahn, 2010; Nikolic et al., 2012; Concu and Cordeiro, 2018; Uesawa, 2018). The prediction model constructed by DL in an H2O 3.2 package, where hidden layers, epochs, and best epochs were 200, 10, and 5, respectively (H20 ai, CA, USA, Chow, 2014) with ECFP showed that the ROC-AUC was 0.888 (Uesawa, 2018). In addition, the random forest in JMP pro 14, in which number of terms and maximum splits per tree were 500 and 256 for fingerprint, and 500 and 29 for 3D descriptors, respectively, predicted the models using the above ECFP descriptors or 3D descriptors with AUC of 0.901 or 0.907 (Uesawa, 2018). Until today, to improve the performance of prediction model, the selection of structural descriptors carried out using the skills and knowledge. Because it is difficult to perfectly preserve the original data, many of these descriptors are irreversible conversions. However, in the DL method using task-specific automatically extracted image information for molecular structures that do not require such high craftsmanship input data, it may demonstrate equal to or better than the above method using descriptors hand-engineered without prior knowledge or assumptions about the features.

When considering applying DL to a compound, whose molecular structure is a variable data format that can have branches and loops, there are problems with how to handle that input or output. To address this issue, graphic-based convolution, which has the ability to handle graph structures, simple encoding of the molecules (atoms, bonds, distances, etc.) represented by edge-connected nodes introducing convolution operations on each nodes non-Euclidean structure was proposed as modifications of DL architectures specialized for molecular fingerprints and models in the terms of structural features, physical properties, and activity (Duvenaud et al., 2015; Gilmer et al., 2017; Zhou and Li, 2017; Fernandez et al., 2018; Li C. et al., 2018). Since a chemical compound can also be represented as an undirected graphs of atoms when an atom is defined as a vertex (node) and a bond is defined as a side (edge), it is possible to construct a highly accurate prediction model by applying a convolution operation to the graph including their physical and chemical properties and extracting meaningful features from the large scale datasets of graph structure (Defferrard et al., 2016; Kipf and Welling, 2016). However, unlike image data, there drawback that a connection relation of peripheral nodes around the attention node of the graph is indefinite for each target node. To solve this difficulty with a heuristic or theoretical approach, graph convolution can be applied to graph Fourier transformation considering the adjacency of nodes by parameterizing weighted and undirected graphs without loops and multiple edges. Fourier conversion decomposes a waveform

signal component by frequency component, but graph Fourier conversion decomposed a signal defined on a graph into "gentle signal" or "steep signal." As for chemical structure, the graph signal converts into a graph spectral region assigning feature vectors to each atom in a chemical substance and their interaction between atoms. Thus, it is very well-adapted to prediction of local molecular structure-dependent physiological activity. In the case of definitions derived from the graph Fourier transform, for technical reasons, it needs to undirected and weighted graph without loops and multiple edges. On the other hand, by defining graph convolution more directly from only the connection relationship of nodes and edges, it is possible to introduce a more complicated structure such as a directed graph, multiple edges, and loops to graph convolution (Schlichtkrull et al., 2017). That is, for each node, its adjacent nodes are classified according to how they are connected, and then the sum (or average) of the signals of the neighboring nodes is taken for each neighborhood according to the manner of connection and according to how it is connected. However, since this method relied on edge and/or node information, the graph structures from the 3D conformational flexibility and the diversity of many features on the edge and/or node, such as shape, electrostatics, quantum effects, and other properties emerged from the molecular graph essential to clearly represent the biological systems and their relationship for the molecular activity and to consistently outperform other models (Kearnes et al., 2016). Additionally, since this graph structured format is heterogeneous among molecules, many learning algorithms how to process the complex graph effectively, except homogeneous input features. Therefore, to resolve issues, data transformings for the graph structure of the molecules to fix data size and format (Duvenaud et al., 2015; Liu K. et al., 2018). In addition, representations by the SMILES (Weininger, 1988; Putz and Dudaş, 2013; Achary, 2014; Jastrzebski et al., 2018; Kumar and Chauhan, 2018) do not encode bond lengths and mutual orientation of atom in space, meaning that they lack information for the molecular conformations, such as 3D atomic arrangements and some molecule stereoisomers.

Also, 3D-CNN, convolutional layers extended to 3D filter that move 3-directions (x, y, z) extract spatiotemporal features from moving objects proposed as a method applied to motion image recognition (Ji et al., 2013; Blendowski and Heinrich, 2018; Lu et al., 2018). It has been successfully used to extract against the temporal change of the spatial structure data as a feature expression of 3D volume space such as cuboid output using the node locally connected to all the images within a certain time width (Ji et al., 2013; Maturana and Scherer, 2015). In this method, although the temporal change such as event detection in videos, 3D images etc. is considered in the extracted feature, it depends on the size in the time direction of the filter. Therefore, when recognizing an operation longer than the filter size, selection and combination processing of those features must be performed. As for chemical compounds, the 3D-CNN has been successfully shown to able to handle the data with spatial structure such as 3D-structures, on the choice of the data representation (Ji et al., 2013; Maturana and Scherer, 2015; Blendowski and Heinrich, 2018; Kuzminykh et al., 2018). If a suitable representation used, the most critical

information efficiently captured. In addition, the chemical compounds induced conformational changes target interactions is possible to a number of conformations or orientations (Tuffery and Derreumaux, 2017; Salmaso and Moro, 2018). Furthermore, the conformational changes of target proteins by ligands and protein-ligands interactions have been studied computational (Yang et al., 2016; Hollingsworth and Dror, 2018; Nusrat and Khan, 2018). Therefore, the 3D-CNN could be a very useful method for extracting structural features based on molecular dynamics, which the dynamic behavior of molecular system as a function of time. However, since a data in non-euclidean spaces, such as spherical data is difficult to trivially apply for direct 3D representation, the suitable conditions such as scaling and required number of input samples have not been cleared completely, which leads to poor performance by sparsity and redundancy in the data and increased complexity in the convolution process (Ji et al., 2013; Maturana and Scherer, 2015; Blendowski and Heinrich, 2018; Kuzminykh et al., 2018). In additions, 3D-CNNs requires more 3D matrix and more calculations than 2D. Thus, the scaling for the CNNs to 3D representations is not straightforward due to the sparsity in input data and the complexity in the convolution operations (Ji et al., 2013; Maturana and Scherer, 2015; Blendowski and Heinrich, 2018; Kuzminykh et al., 2018). Therefore, even now, 3D-CNN need shape descriptors by hand, such as light field descriptors (Pu and Ramani, 2006), mesh DOG (Zaharescu et al., 2009), spin images (Johnson and Hebert, 1999), heat kernel signatures (Xiang et al., 2014), and spherical harmonics high performance (Kazhdan et al., 2003). To alleviate this problem, although Gaussian blur representation was proposed to reduce the sparsity and the redundancy of input, convolving with the Gaussian kernel leads to information loss (Kuzminykh et al., 2018).

Previously, it was ascertained that the Deep Snap-DL method yields the corresponding predicted values for different physiological activities between optical R/S isomers (Uesawa, 2018). This report indicated that Deep Snap-DL accurately extract physiological activities depending on molecular conformation-specificity optimization for various conformations is necessary to maintain high performance of the prediction model. In this research, to define the steric conformation of the molecular structure, CORINA Classic software was used. However, if more suitable definition of 3D steric structures of chemical compounds directly or indirectly related to biological activity, mechanisms, and molecular pathways such as determination of 3D structure for a protein receptor with apparent ligand affinity pocket were established based on the molecular dynamics stimulation, the Deep Snap-DL procedure would be outperformed.

On the other hand, there are some problems that need to be improved so far in this Deep Snap-DL method. At first, in principle, this strategy to capture more detail and greater amount of information chemical structures using more molecular images from 3D-rotation (Uesawa, 2018). In supervised learning, output data corresponding to input data can be obtained, but learning is performed for the purpose of minimizing the error by comparing the output to new data. Therefore, the correction of misclassification for a large amount of labeled input data

is difficult. If the classification criteria within image data could be clarified using proposed visual explanations technique (Simonyan et al., 2013; Mahendran and Vedaldi, 2014; Selvaraju et al., 2016; Smilkov et al., 2017; Zhen et al., 2017; Philbrick et al., 2018), it may be useful for estimation of 3D structure important for physiological activity of the compound and would more reduction of calculation cost by reducing the number of images used. Furthermore, by parameters for Deep Snap in this study, the calculation time was reduced the relatively high performance of the prediction model for the CAR agonist activity. In agreement with previous report although DL able to accurately predict for a molecule with just close neighbors in the training dataset, a hitherto unexamined chemical was predicted close to the average of all training molecule activities, which the lack of ability to learn beyond the training dataset (Liu R. et al., 2018). Deep Snap-DL method indicated the performances of prediction models depending on input datasets produced by various conditions including bonds, spacing, angles, colors, atom size, etc. Moreover, the AUCs were reduced by random permutation of the activity scores of datasets consisting training, validations, and test as non-endpoint activity. These findings suggested that the task-specific improvement of Deep Snap-DL technique by adjustments of input data with the representations of chemical structure such as bonds, space, atom size etc. could be more available approach than conventional methods. Taken together, by combining the Deep Snap strategy with parts of graph-CNN or 3D-CNN functions. Overall, the novel approach Deep Snap not only would fill a gap between chemical structure and toxicological prediction, but also may be useful for constructing an *in silico* prediction model of appropriate chemical risk assessment replace.

In summary, the relations of the parameters of Deep Snap such as (1) number of molecules per SDF files split into (2) zoom factor percentage, (3) atom size for van der waals percentage, (4) bond radius, (5) minimum bond distance, and (6) bond tolerance with the VLs as indicator for evaluating the performance of the DL following quadratic function curves, suggesting that optimal thresholds exist to attain the best performance with these prediction models. Using the parameter values the best performance with the prediction model, the prediction model for CAR agonist was built using 64 images at 105° angle AUCs of 0.791. The results of this study feature the possible power of novel DL-based QSAR approach for prediction of potential toxicity of large datasets of any chemical compounds.

AUTHOR CONTRIBUTIONS

YU initiated and supervised the work, designed the experiments, collected the information about chemical compounds, and edited the manuscript. YM drafted the manuscript. YU and YM read and approved the final manuscript.

FUNDING

SUPPLEMENTARY MATERIAL

The Supplementary Material for this article can be found online at: https://www.frontiersin.org/articles/10.3389/fbioe.2019.00065/full#supplementary-material

REFERENCES

3D Structure Generator CORINA Classic. (2019). *3D Structure Generator CORINA Classic*. Nürnberg: Molecular Networks GmbH. Available online at: www.mn-am.com

Abdelaziz, A., Spahn-Langguth, H., Schramm, K -W., and Tetko, I. V. (2016). Consensus modeling for HTS assays using *in silico* descriptors calculates the best balanced accuracy in Tox21 challenge. *Front. Environ. Sci.* 4:2. doi: 10.3389/fenvs.2016.00002

Achary, P. G. (2014). Simplified molecular input line entry system-based optimal descriptors: QSARmodelling for voltage-gated potassium channel subunit Kv7.2. *SAR QSAR Environ. Res.* 25, 73–90. doi: 10.1080/1062936X.2013.842930

Aggarwal, H. K., Mani, M. P., and Jacob, M. (2018). MoDL: model based deep learning architecture for inverse problems. *IEEE Trans. Med. Imaging.* 38, 394–405. doi: 10.1109/TMI.2018.2865356

Ambe, K., Ishihara, K., Ochibe, T., Ohya, K., Tamura, S., Inoue, K., et al. (2018). *In silico* prediction of chemical-induced hepatocellular hypertrophy using molecular descriptors. *Toxicol. Sci.* 162, 667–675. doi: 10.1093/toxsci/kfx287

Andronico, A., Randall, A., Benz, R. W., and Baldi, P. (2011). Data-driven high-throughput prediction of the 3-D structure of small molecules: review and progress. *J. Chem. Inf. Model.* 51, 760–776. doi: 10.1021/ci100223t

Attene-Ramos, M. S., Huang, R., Michael, S., Witt, K. L., Richard, A., Tice, R. R., et al. (2015). Profiling of the Tox21 chemical collection for mitochondrial function to identify compounds that acutely decrease mitochondrial membrane potential. *Environ. Health Perspect.* 123, 49–56. doi: 10.1289/ehp.1408642

Attene-Ramos, M. S., Miller, N., Huang, R., Michael, S., Itkin, M., Kavlock, R. J., et al. (2013). The Tox21 robotic platform for the assessment of environmental chemicals from vision to reality. *Drug Discov. Today.* 18, 716–723. doi: 10.1016/j.drudis.2013.05.015

Azimi, S. M., Britz, D., Engstler, M., Fritz, M., and Mücklich, F. (2018). Advanced steel microstructural classification by methods. *Sci. Rep.* 8:2128. doi: 10.1038/s41598-018-20037-5

Banerjee, P., Eckert, A. O., Schrey, A. K., and Preissner, R. (2018). ProTox-II: a webserver for the prediction of toxicity of chemicals. *Nucleic Acids Res.* 46, W257–W263. doi: 10.1093/nar/gky318

Bell, S. M., Phillips, J., Sedykh, A., Tandon, A., Sprankle, C., Morefield, S. Q., et al. (2017). An integrated chemical environment to support 21st-century toxicology. *Environ. Health Perspect.* 125:054501. doi: 10.1289/EHP1759

Bengio, Y., Courville, A., and Vincent, P. (2013). Representation learning: a review and new perspectives. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 1798–1828. doi: 10.1109/TPAMI.2013.50

Blendowski, M., and Heinrich, M. P. (2018). Combining MRF-based deformable registration and deep binary 3D-CNN descriptors for large lung motion estimation in COPD patients. *Int. J. Comput. Assist. Radiol Surg.* 14, 43–52. doi: 10.1007/s11548-018-1888-2

Bloomingdale, P., Housand, C., Apgar, J. F., Millard, B. L., Mager, D. E., Burke, J. M., et al. (2017). Quantitative systems toxicology. *Curr. Opin. Toxicol.* 4, 79–87. doi: 10.1016/j.cotox.2017.07.003

Bolisetty, S., Traylor, A., Zarjou, A., Johnson, M. S., Benavides, G. A., Ricart, K., et al. (2013). Mitochondria-targeted heme oxygenase-1 decreases oxidative stress in renal epithelial cells. *Am. J. Physiol. Renal. Physiol.* 305, F255–F264. doi: 10.1152/ajprenal.00160.2013

Brockmeier, E. K., Hodges, G., Hutchinson, T. H., Butler, E., Hecker, M., Tollefsen, K. E., et al. (2017). The role of omics in the application of adverse

outcome pathways for chemical risk assessment. *Toxicol. Sci.* 158, 252–262. doi: 10.1093/toxsci/kfx097

Cagli, E., Dumas, C., and Prouff, E. (2017). *Convolutional Neural Networks with Data Augmentation against Jitter-Based Countermeasures—Profiling Attacks without Pre-Processing*. Cryptology ePrint Archive: Report 2017/740.

Cammer, S. (2007). SChiSM2: creating interactive web page annotations of molecular structure models using Jmol. *Bioinformatics.* 23, 383–384. doi: 10.1093/bioinformatics/btl603

Chen, I. J., and Foloppe, N. (2008). Conformational sampling of druglike molecules with MOE and catalyst: implications for pharmacophore modeling and virtual screening. *J. Chem. Inf. Model.* 48, 1773–1791. doi: 10.1021/ci800130k

Chen, S., Hsieh, J. H., Huang, R., Sakamuru, S., Hsin, L. Y., Xia, M., et al. (2015). Cell-based high-throughput screening for aromatase inhibitors in the Tox21 10K library. *Toxicol. Sci.* 147, 446–457. doi: 10.1093/toxsci/kfv141

Chow, J -F. (2014). *Things to Try After useR!—Part 1: Deep Learning with H2O.* Available online at: http://www.r-bloggers.com/things-to-try-after-user-part-1-deeplearning-with-h2o/ (Accessed August 10, 2017).

Cipullo, S., Snapir, B., Prpich, G., Campo, P., and Coulon, F. (2019). Prediction of bioavailability and toxicity of complex chemical mixtures through machine learning models. *Chemosphere* 215, 388–395. doi: 10.1016/j.chemosphere.2018.10.056

Clark, M., and Steger-Hartmann, T. (2018). A big data approach to the concordance of the toxicity of pharmaceuticals in animals and humans. *Regul. Toxicol. Pharmacol.* 96, 94–105. doi: 10.1016/j.yrtph.2018.04.018

Collins, F. S., Gray, G. M., and Bucher, J. R. (2008). Toxicology. Transforming environmental health protection. *Science* 319, 906–907. doi: 10.1126/science.1154619

Concu, R., and Cordeiro, M. N. D. S. (2018). Looking for new inhibitors for the epidermal growth factor receptor. *Curr. Top. Med. Chem.* 18, 219–232. doi: 10.2174/1568026618666180329123023

Defferrard, M., Bresson, X., and Vandergheynst, P. (2016). Convolutional neural networks on graphs with fast localized spectral filtering. *arXiv:1606.09375v3* [Preprint]. Available online at: https://arxiv.org/pdf/1606.09375.pdf

Dougall, L. G. (2001). Functional methods for quantifying agonists and antagonists. *J. Recept. Signal Transduct. Res.* 21, 117–137. doi: 10.1081/RRS-100107425

Duvenaud, D., Maclaurin, D., Aguilera-Iparraguirre, J., Gómez-Bombarelli, R., Hirzel, T., Aspuru-Guzik, A., et al. (2015). Convolutional networks on graphs for learning molecular fingerprints. *arXiv:1509.09292v2.* Available online at: https://arxiv.org/pdf/1509.09292.pdf

Elfwing, S., Uchibe, E., and Doya, K. (2018). Sigmoid-weighted linear units for neural network function approximation in reinforcement learning. *Neural. Netw.* 107, 3–11. doi: 10.1016/j.neunet.2017.12.012

Fang, H., Tong, W., Welsh, W. J., and Sheehan, D. M. (2003). QSAR models in receptor mediated effects: the nuclear receptor superfamily. *J. Mol. Struct.* 622, 113–125. doi: 10.1016/S0166-1280(02)00623-1

Fang, X., Bagui, S., and Bagui, S. (2017). Improving virtual screening predictive accuracy of Human kallikrein 5 inhibitors using machine learning models. *Comput. Biol. Chem.* 69, 110–119. doi: 10.1016/j.compbiolchem.2017.05.007

Fay, K. A., Villeneuve, D. L., Swintek, J., Edwards, S. W., Nelms, M. D., Blackwell, B. R., et al. (2018). Differentiating pathway-specific from nonspecific effects in high-throughput toxicity data: a foundation for prioritizing adverse outcome pathway development. *Toxicol. Sci.* 163, 500–515. doi: 10.1093/toxsci/kfy049

Fernandez, M., Ban, F., Woo, G., Hsing, M., Yamazaki, T., LeBlanc, E., et al. (2018). Toxic colors: the use of deep learning for predicting toxicity of compounds merely from their graphic images. *J. Chem. Inf. Model.* 58, 1533–1543. doi: 10.1021/acs.jcim.8b00338

Gawehn, E., Hiss, J. A., Brown, J. B., and Schneider, G. (2018). Advancing drug discovery via GPU-based deep learning. *Expert Opin. Drug Discov.* 13, 579–582. doi: 10.1080/17460441.2018.1465407

Genuis, S. J., and Kyrillos, E. (2017). The chemical disruption of human metabolism. *Toxicol. Mech. Methods.* 27, 477–500. doi: 10.1080/15376516.2017.1323986

Ghasemi, F., Mehridehnavi, A., Pérez-Garrido, A., and Pérez-Sánchez, H. (2018). Neural network and deep-learning algorithms used in QSAR studies: merits and drawbacks. *Drug Discov. Today.* 23, 1784–1790. doi: 10.1016/j.drudis.2018.06.016

Gilmer, J., Schoenholz, S. S., Riley. P. F., VInyals, O., and Dahl, G. E. (2017). Neural message passing for quantum chemistry. *arXiv:1704.01212v2* [Preprint]. Available online at: https://arxiv.org/pdf/1704.01212.pdf

Guimarães, M. C., Duarte, M. H., Silla, J. M., and Freitas, M. P. (2016). Is conformation a fundamental descriptor in QSAR? A case for halogenated anesthetics. *Beilstein J. Org. Chem.* 12, 760–768. doi: 10.3762/bjoc.12.76

Ha, R., Chang, P., Karcich, J., Mutasa, S., Fardanesh, R., Wynn, R. T., et al. (2018). Axillary lymph node evaluation utilizing convolutional neural networks using MRI dataset. *J. Digit Imaging.* 31, 851–856. doi: 10.1007/s10278-018-0086-7

Halder, A. K., Moura, A. S., and Cordeiro, M. N. D. S. (2018). QSAR modelling: a therapeutic patent review 2010-present. *Expert Opin. Ther. Pat.* 28, 467–476. doi: 10.1080/13543776.2018.1475560

Hanson, R. M. (2016). Jmol SMILES and Jmol SMARTS: specifications and applications. *J. Cheminform.* 26:50. doi: 10.1186/s13321-016-0160-4

Hanson, R. M., and Lu, X. J. (2017). DSSR-enhanced visualization of nucleic acid structures in Jmol. *Nucleic Acids Res.* 45:W528–W533. doi: 10.1093/nar/gkx365

Heindel, J. J. (2018). The developmental basis of disease: Update on environmental exposures and animal models. *Basic Clin. Pharmacol. Toxicol.* 1–9. doi: 10.1111/bcpt.13118

Heindel, J. J., Skalla, L. A., Joubert, B. R., Dilworth, C. H., and Gray, K. A. (2017). Review of developmental origins of health and disease publications in environmental epidemiology. *Reprod. Toxicol.* 68, 34–48. doi: 10.1016/j.reprotox.2016.11.011

Herráez, A. (2006). Biomolecules in the computer: Jmol to the rescue. *Biochem. Mol. Biol. Educ.* 34, 255–261. doi: 10.1002/bmb.2006.494034042644

Hollingsworth, S. A., and Dror, R. O. (2018). Molecular dynamics simulation for all. *Neuron.* 99, 1129–1143. doi: 10.1016/j.neuron.2018.08.011

Hsieh, J. H., Huang, R., Lin, J. A., Sedykh, A., Zhao, J., Tice, R. R., et al. (2017). Real-time cell toxicity profiling of Tox21 10K compounds reveals cytotoxicity dependent toxicity pathway linkage. *PLoS ONE* 12:e0177902. doi: 10.1371/journal.pone.0177902

Hsieh, J. H., Sedykh, A., Huang, R., Xia, M., and Tice, R. R. (2015). A data analysis pipeline accounting for artifacts in Tox21 quantitative high-throughput screening assays. *J. Biomol. Screen.* 20, 887–897. doi: 10.1177/1087057115581317

Hu, G., Wang, K., Peng, Y., Qiu, M., Shi, J., and Liu, L. (2018). Deep learning methods for underwater target feature extraction and recognition. *Comput. Intell. Neurosci.* 2018:10. doi: 10.1155/2018/1214301

Hua, S., Zhang, H., Song, Y., Li, R., Liu, J., Wang, Y., et al. (2012). High expression of Mfn1 promotes early development of bovine SCNT embryos: improvement of mitochondrial membrane potential and oxidative metabolism. *Mitochondrion.* 12, 320–327. doi: 10.1016/j.mito.2011.12.002

Huang, R., Sakamuru, S., Martin, M. T., Reif, D. M., Judson, R. S., Houck, K. A., et al. (2014). Profiling of the Tox21 10K compound library for agonists and antagonists of the estrogen receptor alpha signaling pathway. *Sci Rep.* 4:5664. doi: 10.1038/srep05664

Huang, R., Southall, N., Wang, Y., Yasgar, A., Shinn, P., Jadhav, A., et al. (2011). The NCGC pharmaceutical collection: a comprehensive resource of clinically approved drugs enabling repurposing and chemical genomics. *Sci. Transl. Med.* 3:80ps16. doi: 10.1126/scitranslmed.3001862

Huang, R., and Xia, M. (2017). Editorial: Tox21 challenge to build predictive models of nuclear receptor and stress response pathways as mediated by exposure to environmental toxicants and drugs. *Front. Environ. Sci.* 5, 1–3. doi: 10.3389/fenvs.2017.00003

Huang, R., Xia, M., Sakamuru, S., Zhao, J., Shahane, S. A., Attene-Ramos, M., et al. (2016). Modelling the Tox21 10 K chemical profiles for *in vivo* toxicity prediction and mechanism characterization. *Nat Commun.* 7:10425. doi: 10.1038/ncomms10425

Hussain, Z., Gimenez, F., Yi, D., and Rubin, D. (2018). Differential data augmentation techniques for medical imaging classification tasks. *AMIA Annu. Symp. Proc.* 2017, 979–984.

Insel, P. A., Amara, S. G., Blaschke, T. F., and Meyer, U. A. (2017). Introduction to the theme "new methods and novel therapeutic approaches in pharmacology and toxicology". *Annu. Rev. Pharmacol. Toxicol.* 57, 13–17. doi: 10.1146/annurev-pharmtox-091616-023708

Jastrzebski, S., Leśniak, D., and Czarnecki, V. M. (2018). Learning to SMILE(S). *arXiv:1602.06289* [Preprint]. Available online at: https://arxiv.org/pdf/1602.06289.pdf

Ji, S., Yang, M., and Yu, K. (2013). 3D convolutional neural networks for human action recognition. *IEEE Trans. Pattern Anal. Mach. Intell.* 35, 221–231. doi: 10.1109/TPAMI.2012.59

Jia, Y., Shelhamer, E., Donahue, J., Karayev, S., Long, J., Girshick, R., et al. (2014). Caffe: convolutional architecture for fast feature embedding. *CVPR* 675–678. Available online at: https://ucb-icsi-vision-group.github.io/caffe-paper/caffe.pdf

Johnson, A. E., and Hebert, M. (1999). Using spin images for efficient object recognition in cluttered 3D scenes. *IEEE Trans. Pattern Anal. Machine Intelligence* 21, 433–449. doi: 10.1109/34.765655

Johnson, R., and Zhang, T. (2015). Semi-supervised convolutional neural networks for text categorization via region embedding. *Adv. Neural. Inf. Process. Syst.* 28, 919–927. Available online at: https://papers.nips.cc/paper/5849-semi-supervised-convolutional-neural-networks-for-text-categorization-via-region-embedding.pdf

Juberg, D. R., Knudsen, T. B., Sander, M., Beck, N. B., Faustman, E. M., Mendrick, D. L., et al. (2017). FutureTox III: bridges for translation. *Toxicol. Sci.* 155, 22–31. doi: 10.1093/toxsci/kfw194

Kavlock, R. J., Austin, C. P., and Tice, R. R. (2009). Toxicity testing in the 21st century: implications for human health risk assessment. *Risk Anal.* 29, 485–487; discussion 492-497. doi: 10.1111/j.1539-6924.2008.01168.x

Kazhdan, M., Funkhouser, T., and Rusinkiewicz, S. (2003). Rotation invariant spherical harmonic representation of 3D shape descriptors. *Eurogr. Sympos. Geomet. Process.* 43, 156–165. doi: 10.2312/SGP/SGP03/156-165

Kearnes, S., McCloskey, K., Berndl, M., Pande, V., and Riley, P. (2016). Molecular graph convolutions: moving beyond fingerprints. *J. Comput. Aided Mol. Des.* 30, 595–608. doi: 10.1007/s10822-016-9938-8

Khan, P. M., and Roy, K. (2018). Current approaches for choosing feature selection and learning algorithms in quantitative structure-activity relationships (QSAR). *Expert Opin. Drug Discov.* 29, 1–15. doi: 10.1080/17460441.2018.1542428

Khened, M., Kollerathu, V. A., and Krishnamurthi, G. (2019). Fully convolutional multi-scale residual DenseNets for cardiac segmentation and automated cardiac diagnosis using ensemble of classifiers. *Med. Image Anal.* 51, 21–45. doi: 10.1016/j.media.2018.10.004

Kipf, T. N., and Welling, M. (2016). Semi-supervised classification with graph convolutional networks. *arXiv:1609.02907v4* [Preprint]. Available online at: https://arxiv.org/pdf/1609.02907.pdf

Kobayashi, K., Hashimoto, M., Honkakoski, P., and Negishi, M. (2015). Regulation of gene expression by CAR: an update. *Arch. Toxicol.* 89, 1045–1055. doi: 10.1007/s00204-015-1522-9

Krizhevsky, A., Sutskev, I., and Hinton, G. E. (2012). ImageNet classification with deep convolutional neural networks. *Adv. Neural. Inf. Process. Syst.* 1, 1097–1105. doi: 10.1145/3065386

Kulkarni, A. J., Jayaraman, V. K., and Kulkarni, B. D. (2009). Review on lazy learning regressors and their applications in QSAR. *Comb. Chem. High Throughput Screen.* 12, 440–450. doi: 10.2174/138620709788167908

Kumar, A., and Chauhan, S. (2018). Use of Simplified Molecular Input Line Entry System and molecular graph based descriptors in prediction and design of pancreatic lipase inhibitors. *Future Med. Chem.* 10, 1603–1622. doi: 10.4155/fmc-2018-0024

Kuzminykh, D., Polykovskiy, D., Kadurin, A., Zhebrak, A., Baskov, I., Nikolenko, S., et al. (2018). 3D molecular representations based on the wave transform for convolutional neural networks. *Mol. Pharm.* 15, 4378–4385. doi: 10.1021/acs.molpharmaceut.7b01134

LeCun, Y., Bengio, Y., and Hinton, G. (2015). Deep learning. *Nature* 521, 436–444. doi: 10.1038/nature14539

Leist, M., Ghallab, A., Graepel, R., Marchan, R., Hassan, R., et al. (2017). Adverse outcome pathways: opportunities, limitations and open questions. *Arch. Toxicol.* 91, 3477–3505. doi: 10.1007/s00204-017-2045-3

Li, A. X., Sun, M., and Li, X. (2017). Withaferin-A induces apoptosis in osteosarcoma U2OS cell line via generation of ROS and disruption of mitochondrial membrane potential. *Eur. Rev. Med. Pharmacol. Sci.* 21, 1368–1374. doi: 10.4103/0973-1296.211042

Li, C., Cui, Z., Zheng, W., Xu, C., Ji, R., and Yang, J. (2018). Action-attending graphic neural network. *IEEE Trans. Image Process.* 27, 3657–3670. doi: 10.1109/TIP.2018.2815744

Li, H., Gong, X. J., Yu, H., and Zhou, C. (2018). Deep neural network based predictions of protein interactions using primary sequences. *Molecules* 23:E1923. doi: 10.3390/molecules23081923

Li, S., Jiang, H., and Pang, W. (2017). Joint multiple fully connected convolutional neural network with extreme learning machine for hepatocellular carcinoma nuclei grading. *Comput. Biol. Med.* 1, 156–167. doi: 10.1016/j.compbiomed.2017.03.017

Linden, A. (2006). Measuring diagnostic and predictive accuracy in disease management: an introduction to receiver operating characteristic (ROC) analysis. *J. Eval. Clin. Pract.* 12, 132–139. doi: 10.1111/j.1365-2753.2005.00598.x

Liu, K., Sun, X., Jia, L., Ma, J., Xing, H., Wu, J., et al. (2018). Chemi-Net: a molecular graph convolutional network for accurate drug property prediction. *arXiv:1803.06236v2* [Preprint]. Available online at: https://arxiv.org/abs/1803.06236

Liu, Q., Wang, Q., Xu, C., Shao, W., Zhang, C., Liu, H., et al. (2017). Organochloride pesticides impaired mitochondrial function in hepatocytes and aggravated disorders of fatty acid metabolism. *Sci. Rep.* 7:46339. doi: 10.1038/srep46339

Liu, R., Wang, H., Glover, K. P., Feasel, M. G., and Wallqvist, A. (2018). Dissecting machine-learning prediction of molecular activity: is an applicability domain needed for quantitative structure-activity relationship models based on deep neural networks? *J. Chem. Inf. Model.* 59, 117–126. doi: 10.1021/acs.jcim.8b00348

Lo, Y. C., Rensi, S. E., Torng, W., and Altman, R. B. (2018). Machine learning in chemoinformatics and drug discovery. *Drug Discov. Today* 23, 1538–1546. doi: 10.1016/j.drudis.2018.05.010

Lu, N., Wu, Y., Feng, L., and Song, J. (2018). Deep learning for fall detection: 3D-CNN combined with LSTM on video kinematic data. *IEEE J. Biomed. Health Inform.* 23, 314–323. doi: 10.1109/JBHI.2018.2808281

Luechtefeld, T., Rowlands, C., and Hartung, T. (2018). Big-data and machine learning to revamp computational toxicology and its use in risk assessment. *Toxicol. Res.* 7, 732–744. doi: 10.1039/c8tx00051d

Lumini, A., and Nanni, L. (2018). Convolutional neural networks for ATC classification. *Curr. Pharm. Des.* 24, 4007–4012. doi: 10.2174/1381612824666181112113438

Ma, J., Sheridan, R. P., Liaw, A., Dahl, G. E., and Svetnik, V. (2015). Deep neural nets as a method for quantitative structure-activity relationships. *J. Chem. Inf. Model.* 55, 263–274. doi: 10.1021/ci500747n

Mahendran, A., and Vedaldi, A. (2014). Understanding deep image representations by inverting them. *arXiv:1412.0035v1* [Preprint]. Available online at: https://arxiv.org/pdf/1412.0035.pdf

Mallat, S. (2016). Understanding deep convolutional networks. *Philos. Trans. A Math. Phys. Eng. Sci.* 374:20150203. doi: 10.1098/rsta.2015.0203

Malloy, T., Zaunbrecher, V., Beryt, E., Judson, R., Tice, R., Allard, P., et al. (2017). Advancing alternatives analysis: the role of predictive toxicology in selecting safer chemical products and processes. *Integr. Environ. Assess. Manag.* 13, 915–925. doi: 10.1002/ieam.1923

Manallack, D. T., Burden, F. R., and Winkler, D. A. (2010). Modelling inhalational anaesthetics using bayesian feature selection and QSAR modelling methods. *ChemMedChem.* 5, 1318–1323. doi: 10.1002/cmdc.201000056

Manibusan, M. K., and Touart, L. W. (2017). A comprehensive review of regulatory test methods for endocrine adverse health effects. *Crit. Rev. Toxicol.* 47, 433–481. doi: 10.1080/10408444.2016.1272095

Martinez, M., and Stiefelhagen, R. (2018). Taming the cross entropy loss. *arXiv:1810.05075v1* [Preprint]. Available online at: https://arxiv.org/pdf/1810.05075.pdf

Marty, M. S., Borgert, C., Coady, K., Green, R., Levine, S. L., Mihaich, E., et al. (2018). Distinguishing between endocrine disruption and non-specific effects on endocrine systems. *Regul. Toxicol. Pharmacol.* 99, 142–158. doi: 10.1016/j.yrtph.2018.09.002

Maturana, D., and Scherer, S. (2015). 3D Convolutional Neural Networks for landing zone detection from LiDAR. *IEEE Int. Conf. Robot. Autom.* 2015, 1050–4729. doi: 10.1109/ICRA.2015.7139679

Mayr, A., Klambauer, G., Unterthiner, T., and Hochreiter, S. (2016). DeepTox: toxicity prediction using deep learning. *Front. Environ. Sci.* 3:80. doi: 10.3389/fenvs.2015.00080

Mayr, A., Klambauer, G., Unterthiner, T., Steijaert, M., Wegner, J. K., Ceulemans, H., et al. (2018). Large-scale comparison of machine learning

methods for drug target prediction on ChEMBL. *Chem. Sci.* 9, 5441–5451. doi: 10.1039/c8sc00148k

McMahon, M., Ding, S., Acosta-Jimenez, L. P., Terranova, R., Gerard, M. A., and Vitobello, A., et al. (2018). Constitutive androstane receptor 1 is constitutively bound to chromatin and 'primed' for transactivation in hepatocytes. *Mol. Pharmacol.* 95, 97–105. doi: 10.1124/mol.118.113555

Merrick, B. A., Paules, R. S., and Tice, R. R. (2015). Intersection of toxicogenomics and high throughput screening in the Tox21 program: an NIEHS perspective. *Int. J. Biotechnol.* 14, 7–27. doi: 10.1504/IJBT.2015.074797

Midzak, A. S., Chen, H., Aon, M. A., Papadopoulos, V., and Zirkin, B. R. (2011). ATP synthesis, mitochondrial function, and steroid biosynthesis in rodent primary and tumor Leydig cells. *Biol. Reprod.* 84, 976–985. doi: 10.1095/biolreprod.110.087460

Mortensen, H. M., Chamberlin, J., Joubert, B., Angrish, M., Sipes, N., Lee, J. S., et al. (2018). Leveraging human genetic and adverse outcome pathway (AOP) data to inform susceptibility in human health risk assessment. *Mamm. Genome.* 29, 190–204. doi: 10.1007/s00335-018-9738-7

Nair, V., and Hinton, G. E. (2010). "Rectified linear units improve restricted Boltzmann machines," in *Proceedings of the 27th International Conference on Machine Learning* (Haifa), 807–814.

Nguyen, Q., Mukkamala, M. C., and Hein, M. (2018). On the loss landscape of a class of deep neural networks with no bad local valleys. *arXiv:arXiv:1809.10749v1* [Preprint]. Available online at: https://arxiv.org/abs/1809.10749

Nikolic, K., Filipic, S., and Agbaba, D. (2012). Multi-target QSAR and docking study of steroids binding to corticosteroid-binding globulin and sex hormone-binding globulin. *Curr. Comput. Aided Drug Des.* 8, 296–308. doi: 10.2174/157340912803519642

NRC. (2007). *Toxicity Testing in the 21st Century: A Vision and a Strategy.* Washington, DC: The National Academies Press.

Nusrat, S., and Khan, R. H. (2018). Exploration of ligand-induced protein conformational alteration, aggregate formation, and its inhibition: a biophysical insight. *Prep. Biochem. Biotechnol.* 48, 43–56. doi: 10.1080/10826068.2017.1387561

Pastur-Romay, L. A., Cedron, F., Pazos, A., and Porto-Pazos, A. B. (2016). Deep artificial neural networks and neuromorphic chips for big data analysis: pharmaceutical and bioinformatics applications. *Int. J. Mol. Sci.* 17:1313. doi: 10.3390/ijms17081313

Peng, Y., Rios, A., Kavuluru, R., and Lu, Z. (2018). Extracting chemical-protein relations with ensembles of SVM and deep learning models. *Database.* 2018:bay073. doi: 10.1093/database/bay073

Pham, T., Tran, T., Phung, D., and Venkatesh, S. (2017). Predicting healthcare trajectories from medical records: a deep learning approach. *J. Biomed. Inform.* 69, 218–229. doi: 10.1016/j.jbi.2017.04.001

Philbrick, K. A., Yoshida, K., Inoue, D., Akkus, Z., Kline, T. L., Weston, A. D., et al. (2018). What does deep learning see? Insights from a classifier trained to predict contrast enhancement phase from CT images. *AJR Am. J. Roentgenol.* 211, 1184–1193. doi: 10.2214/AJR.18.20331

Poernomo, A., and Kang, D. K. (2018). Biased dropout and crossmap dropout: learning towards effective dropout regularization in convolutional neural network. *Neural. Netw.* 104, 60–67. doi: 10.1016/j.neunet.2018.03.016

Polishchuk, P. (2017). Interpretation of quantitative structure-activity relationship models: past, present, and future. *J. Chem. Inf. Model.* 57, 2618–2639. doi: 10.1021/acs.jcim.7b00274

Pu, J., and Ramani, K. (2006). On visual similarity based 2D drawing retrieval. *Computer-Aided Design.* 38, 249–259. doi: 10.1016/j.cad.2005.10.009

Putz, M. V., and Dudaş, N. A. (2013). Determining chemical reactivity driving biological activity from SMILES transformations: the bonding mechanism of anti-HIV pyrimidines. *Molecules.* 18, 9061–9116. doi: 10.3390/molecules18089061

Qatanani, M., and Moore, D. D. (2005). CAR, the continuously advancing receptor, in drug metabolism and disease. *Curr. Drug Metab.* 6, 329–339. doi: 10.2174/1389200054633899

Qiao, H., Wu, J., Li, X., Shoreh, M. H., Fan, J., and Dai, Q. (2018). GPU-based deep convolutional neural network for tomographic phase microscopy with l1 fitting and regularization. *J. Biomed. Opt.* 23, 1–7. doi: 10.1117/1.JBO.23.6.066003

Qiu, Y., Yan, S., Gundreddy, R. R., Wang, Y., Cheng, S., Liu, H., et al. (2017). A new approach to develop computer-aided diagnosis scheme of breast mass

classification using deep learning technology. *J. Xray Sci. Technol.* 25, 751–763. doi: 10.3233/XST-16226

Ragoza, M., Hochuli, J., Idrobo, E., Sunseri, J., and Koes, D. R. (2017). Protein-ligand scoring with convolutional neural networks. *J. Chem. Inf. Model.* 57, 942–957. doi: 10.1021/acs.jcim.6b00740

Rawat, W., and Wang, Z. (2017). Deep convolutional neural networks for image classification: a comprehensive review. *Neural. Comput.* 29, 2352–2449. doi: 10.1162/NECO_a_00990

Reitz, M., Sacher, O., Tarkhov, A., Trumbach, D., and Gasteiger, J. (2004). Enabling the exploration of biochemical pathways. *Org. Biomol. Chem.* 2, 3226–3237. doi: 10.1039/B410949J

Renner, S., Schwab, C. H., Gasteiger, J., and Schneider, G. (2006). Impact of conformational flexibility on three-dimensional similarity searching using correlation vectors. *J. Chem. Inf. Model.* 46, 2324–2332. doi: 10.1021/ci050075s

Richard, A. M., Judson, R. S., Houck, K. A., Grulke, C. M., Volarath, P., Thillainadarajah, I., et al. (2016). ToxCast Chemical landscape: paving the road to 21st century toxicology. *Chem. Res. Toxicol.* 29, 1225–1251. doi: 10.1021/acs.chemrestox.6b00135

Rogers, D., and Hahn, M. (2010). Extended-connectivity fingerprints. *J. Chem. Inf. Model.* 50, 742–754. doi: 10.1021/ci100050t

Roy, K., and Das, R. N. (2014). A review on principles, theory and practices of 2D-QSAR. *Curr. Drug Metab.* 15, 346–379. doi: 10.2174/1389200215666140908102230

Sadowski, J., Gasteiger, J., and Klebe, G. (1994). Comparison of automatic three-dimensional model builders using 639 X-ray structures. *J. Chem. Inf. Comput. Sci.* 34, 1000–1008. doi: 10.1021/ci00020a039

Saha, M., Chakraborty, C., and Racoceanu, D. (2018). Efficient deep learning model for mitosis detection using breast histopathology images. *Comput. Med. Imaging Graph.* 64, 29–40. doi: 10.1016/j.compmedimag.2017

Salmaso, V., and Moro, S. (2018). Bridging molecular docking to molecular dynamics in exploring ligand-protein recognition process: an overview. *Front. Pharmacol.* 9:923. doi: 10.3389/fphar.2018.00923

Sato, M., Horie, K., Hara, A., Miyamoto, Y., Kurihara, K., Tomio, K., et al. (2018). Application of deep learning to the classification of images from colposcopy. *Oncol. Lett.* 15, 3518–3523. doi: 10.3892/ol.2018.7762

Sayers, E. W., Agarwala, R., Bolton, E. E., Brister, J. R., Canese, K., Clark, K., et al. (2018). Database resources of the National center for biotechnology information. *Nucleic Acids Res.* 33, D39–D45. doi: 10.1093/nar/gky1069

Scalfani, V. F., Williams, A. J., Tkachenko, V., Karapetyan, K., Pshenichnov, A., Hanson, R. M., et al. (2016). Programmatic conversion of crystal structures into 3D printable files using Jmol. *J. Cheminform.* 8:66. doi: 10.1186/s13321-016-0181-z

Schlichtkrull, M., Kipf, T. N., Bloem, P., van den Berg, R., Titov, I., and Welling, M. (2017). Modeling relational data with graph convolutional networks. *arXiv:1703.06103v4* [Preprint]. Available online at: https://arxiv.org/pdf/1703.06103.pdf

Schwab, C. H. (2010). Conformations and 3D pharmacophore searching. *Drug Discovery Today Technol.* 7, e245–e253. doi: 10.1016/j.ddtec.2010.10.003

Selvaraju, R. R., Cogswell, M., Das, A., Vedantam, R., Parikh, D., and Batra, D. (2016). Grad-CAM: visual explanations from deep networks via gradient-based localization. *arXiv:1610.02391v3* [Preprint]. Available online at: https://arxiv.org/abs/1610.02391

Shaughnessy, D. T., McAllister, K., Worth, L., Haugen, A. C., Meyer, J. N., Domann, F. E., et al. (2014). Mitochondria, energetics, epigenetics, and cellular responses to stress. *Environ. Health Perspect.* 122, 1271–1278. doi: 10.1289/ehp.1408418

Shen, X., Tian, X., Liu, T., Xu, F., and Tao, D. (2018). Continuous dropout. *IEEE Trans. Neural. Netw. Learn Syst.* 29, 3926–3937. doi: 10.1109/TNNLS.2017.2750679

Sifakis, S., Androutsopoulos, V. P., Tsatsakis, A. M., and Spandidos, D. A. (2017). Human exposure to endocrine disrupting chemicals: effects on the male and female reproductive systems. *Environ. Toxicol. Pharmacol.* 51, 56–70. doi: 10.1016/j.etap.2017.02.024

Silva, F. T., and Trossini, G. H. (2014). The survey of the use of QSAR methods to determine intestinal absorption and oral bioavailability during drug design. *Med. Chem.* 10, 441–448. doi: 10.2174/1573406410666140415122115

Simões, R. S., Maltarollo, V. G., Oliveira, P. R., and Honorio, K. M. (2018). Transfer and multi-task learning in QSAR modeling: advances and challenges. *Front. Pharmacol.* 9:74. doi: 10.3389/fphar.2018.00074

Simonyan, K., Vedaldi, A., and Zisserman,. A. (2013). Deep inside convolutional networks: visualising image classification models and saliency maps. *arXiv:1312.6034v2* [Preprint]. Available online at: https://arxiv.org/pdf/1312.6034.pdf

Sipes, N. S., Wambaugh, J. F., Pearce, R., Auerbach, S. S., Wetmore, B. A., Hsieh, J. H., et al. (2017). An intuitive approach for predicting potential human health risk with the Tox21 10k library. *Environ. Sci. Technol.* 51, 10786–10796. doi: 10.1021/acs.est.7b00650

Smilkov, D., Thorat, N., Kim, B., Viegas, F., and Wattenberg, M. (2017). SmoothGrad: removing noise by adding noise. *arXiv:1706.03825v1* [Preprint]. Available online at: https://arxiv.org/pdf/1706.03825.pdf

Steven, E. O., and Han, D. S. (2018). Feature representation and data augmentation for human activity classification based on wearable IMU sensor data using a deep LSTM neural network. *Sensors* 18:E2892. doi: 10.3390/s18092892

Suárez-Paniagua, V., and Segura-Bedmar, I. (2018). Evaluation of pooling operations in convolutional architectures for drug-drug interaction extraction. *BMC Bioinformatics* 19:209. doi: 10.1186/s12859-018-2195-1

Szegedy, C., Liue, W., Jiam, Y., Sermanet, P., Reed, S., Anguelov, D., et al. (2014). Going deeper with convolutions. *arXiv:1409:4842v1* [Preprint]. Available online at: https://arxiv.org/abs/1409.4842v1

Talevi, A., Bellera, C. L., Di Ianni, M., Duchowicz, P. R., Bruno-Blanch, L. E., and Castro, E. A. (2012). An integrated drug development approach applying topological descriptors. *Curr. Comput. Aided Drug Des.* 8, 172–181. doi: 10.2174/157340912801619076

Tapia-Orozco, N., Santiago-Toledo, G., Barrón, V., Espinosa-García, A. M., García-García, J. A., and García-Arrazola, R. (2017). Environmental epigenomics: current approaches to assess epigenetic effects of endocrine disrupting compounds (EDC's) on human health. *Environ. Toxicol. Pharmacol.* 51, 94–99. doi: 10.1016/j.etap.2017.02.004

Tetko, I. V., Gasteiger, J., Todeschini, R., Mauri, A., Livingstone, D., Ertl, P., et al. (2005). Virtual computational chemistry laboratory—design and description. *J. Comput. Aided Mol. Des.* 19, 453–463. doi: 10.1007/s10822-005-8694-y

Tice, R. R., Austin, C. P., Kavlock, R. J., and Bucher, J. R. (2013). Improving the human hazard characterization of chemicals: a Tox21 update. *Environ. Health Perspect.* 121, 756–765. doi: 10.1289/ehp.1205784

Tollefsen, K. E., Scholz, S., Cronin, M. T., Edwards, S. W., de Knecht, J., Crofton, K., et al. (2014). Applying adverse outcome pathways (AOPs) to support integrated approaches to testing and assessment (IATA). *Regul. Toxicol. Pharmacol.* 70, 629–640. doi: 10.1016/j.yrtph.2014.09.009

Tox21 Data Challenge. (2014). *National Center for Advancing Translational Sciences.* Available online at: https://tripod.nih.gov/tox21/challenge/

Tuffery, P., and Derreumaux, P. (2017). Flexibility and binding affinity in protein-ligand, protein-protein and multi-component protein interactions: limitations of current computational approaches. *J. R. Soc. Interface* 9, 20–33. doi: 10.1098/rsif.2011.0584

Tustison, N. J., Avants, B. B., Lin, Z., Feng, X., Cullen, N., Mata, J. F., et al. (2018). Convolutional neural networks with template-based data augmentation for functional lung image quantification. *Acad. Radiol.* 5:e3. doi: 10.1016/j.acra.2018.08.003

Uesawa, Y. (2018). Quantitative structure–activity relationship analysis using deep learning based on a novel molecular image input technique. *Bioorg. Med. Chem. Lett.* 28, 3400–3403. doi: 10.1016/j.bmcl.2018.08.032

Vakli, P., Deák-Meszlényi, R. J., Hermann, P., and Vidnyánszky, Z. (2018). Transfer learning improves resting-state functional connectivity pattern analysis using convolutional neural networks. *Gigascience* 7:e130. doi: 10.1093/gigascience/giy130

Vouldodimos, A., Doulamis, N., Doulamis, A., and Protopapadakis, E. (2018). Deep learning for computer vision: a brief review. *Comput. Intell. Neurosci.* 2018:7068349. doi: 10.1155/2018/7068349

Wang, C., Chen, H., and Ying, W. (2018). Cytosolic aspartate aminotransferase mediates the mitochondrial membrane potential and cell survival by maintaining the calcium homeostasis of BV2 microglia. *Neuroreport.* 29, 99–105. doi: 10.1097/WNR.0000000000000914

Wang, S. H., Phillips, P., Sui, Y., Liu, B., Yang, M., and Cheng, H. (2018). Classification of Alzheimer's disease based on eight-layer convolutional neural network with leaky rectified linear unit and max pooling. *J. Med. Syst.* 42:85. doi: 10.1007/s10916-018-0932-7

Wang, Y., Xiao, J., Suzek, T. O., Zhang, J., Wang, J., and Bryant, S. H. (2009). PubChem: a public information system for analyzing bioactivities of small molecules. *Nucleic Acids Res.* 37, W623–W633. doi: 10.1093/nar/gkp456

Wang, Y., Xing, J., Xu, Y., Zhou, N., Peng, J., Xiong, Z., et al. (2015). *In silico* ADME/T modelling for rational drug design. *Q. Rev. Biophys.* 48, 488–515. doi: 10.1017/S0033583515000190

Weininger, D. (1988). SMILES, a chemical language and information system. 1. Introduction to methodology and encoding rules. *J. Chem. Inf. Comput. Sci.* 28, 31–36. doi: 10.1021/ci00057a005

Xia, M., Huang, R., Shi, Q., Boyd, W. A., Zhao, J., Sun, N., et al. (2018). Comprehensive analyses and prioritization of Tox21 10K chemicals affecting mitochondrial function by in-depth mechanistic studies. *Environ. Health Perspect.* 126:077010. doi: 10.1289/EHP2589

Xiang, Y., Mottaghi, R., and Savarese, S. (2014). "Beyond PASCAL: a benchmark for 3D object detection in the wild," in *IEEE Winter Conference on Applications of Computer Vision* (Steamboat Springs). doi: 10.1109/WACV.2014.6836101

Xu, Y., Pei, J., and Lai, L. (2017). Deep learning based regression and multiclass models for acute oral toxicity prediction with automatic chemical feature extraction. *J. Chem. Inf. Model.* 57, 2672–2685. doi: 10.1021/acs.jcim.7b00244

Yang, L., Zhang, J., Che, X., and Gao, Y. Q. (2016). Simulation studies of protein and small molecule interactions and reaction. *Methods Enzymol.* 578, 169–212. doi: 10.1016/bs.mie.2016.05.031

Yap, C. W., Li, H., Ji, Z. L., and Chen, Y. Z. (2007). Regression methods for developing QSAR and QSPR models to predict compounds of specific pharmacodynamic, pharmacokinetic and toxicological properties. *Mini Rev. Med. Chem.* 7, 1097–1107. doi: 10.2174/138955707782331696

Yun, X., Rao, W., Xiao, C., and Huang, Q. (2017). Apoptosis of leukemia K562 and Molt-4 cells induced by emamectin benzoate involving mitochondrial membrane potential loss and intracellular Ca^{2+} modulation. *Environ. Toxicol. Pharmacol.* 52, 280–287. doi: 10.1016/j.etap.2017.04.013

Zaharescu, A., Boyer, E., Varanasi, K., and Horaud, R. (2009). "Surface feature detection and description with applications to mesh matching," in *2009 IEEE Conference on Computer Vision and Pattern Recognition* (Florida) 373–380. doi: 10.1109/CVPR.2009.5206748

Zhang, L., Tan, J., Han, D., and Zhu, H. (2017). From machine learning to deep learning: progress in machine intelligence for rational drug discovery. *Drug Discov. Today* 22, 1680–1685. doi: 10.1016/j.drudis.2017.08.010

Zhang, Q., Li, J., Middleton, A., Bhattacharya, S., and Conolly, R. B. (2018). Bridging the data gap from *in vitro* toxicity testing to chemical safety assessment through computational modeling. *Front. Public Health* 6:261. doi: 10.3389/fpubh.2018.00261

Zhang, Z., and Sabuncu, M. R. (2018). Generalized cross entropy loss for training deep neural networks with noisy labels. *arXiv:1805.07836v4.* Available online at: https://arxiv.org/pdf/1805.07836.pdf

Zhen, X., Chen, J., Zhong, Z., Hrycushko, B., Zhou, L., Jiang, S., et al. (2017). Deep convolutional neural network with transfer learning for rectum toxicity prediction in cervical cancer radiotherapy: a feasibility study. *Phys. Med. Biol.* 62, 8246–8263. doi: 10.1088/1361-6560/aa8d09

Zhou, Z., and Li, X. (2017). Convolution on graph: a high-orderand adaptive approach. *arXiv:1706.09916v2* [Preprint]. Available online at: https://arxiv.org/pdf/1706.09916.pdf

Zhu, H., Zhang, J., Kim, M. T., Boison, A., Sedykh, A., and Moran, K. (2014). Big data in chemical toxicity research: the use of high-throughput screening assays to identify potential toxicants. *Chem. Res. Toxicol.* 27, 1643–1651. doi: 10.1021/tx500145h