Review Article

Convolutional Neural Network: Potential applications in the pharmaceutical sector

Anju Sharma, Prabha Garg*

Department of Pharmacoinformatics,
National Institute of Pharmaceutical Education and Research (NIPER), S.A.S. Nagar
Sector 67, Phase 10, S.A.S. Nagar, Punjab-160062, India
prabhagarg@niper.ac.in; gargprabha@yahoo.com

In the past ten years, convolutional neural networks (CNN) have radically altered the landscape of computer vision and image processing and have attained cutting-edge performance in comparison to the traditional machine learning algorithms. Due to their ability to automatically learn hierarchical representations from raw input, CNNs have become a powerful tool for analyzing complex information in a wide variety of domains, from object recognition, classification, detection, and medical image analysis, to autonomous driving. The rising popularity has led to their incorporation into the pharmaceutical industry to evaluate massive chemical libraries, anticipate drug-target interactions, optimize drug design, and find innovative drug candidates that are both effective and specific. In this article, a broad overview of CNN's applications is provided in different pharmaceutical-associated domains.

Introduction

Convolutional neural networks (CNNs), deep learning (DL), and multilayer neural networks based on neuroscience findings have revolutionized the field of computer vision in recent years, achieving state-of-the-art results in tasks such as image classification, object detection, and segmentation.[1] Before CNNs, artificial neural networks (ANNs) required the use of laborious, manual feature extraction techniques to identify objects in images or perform classification. CNN, an extended version of ANN, offers a more scalable method for object recognition and image classification by extracting features from grid-like datasets, implementing matrix multiplication and other concepts from linear algebra to find patterns and excel at analyzing inputs such as images, text, speech, audio or video, distinguishing them from other neural networks.[2] CNNs process the images by assigning weights and biases to several aspects of the input images using four fundamental processes: convolution, inclusion of nonlinearity, pooling, and classification.[3] Figure 1 is an illustration of the architecture of CNN.

The convolutional layer (CL) is CNN's main building unit where most computing occurs. It requires input data (a color image composed of a 3D pixel matrix where an image's three dimensions — height, width, and depth — correspond to RGB (red, green, blue)), a filter, and a feature map.[4] A filter, also referred to as a feature detector or kernel, is a two-dimensional array (n * n) of weights that determines the receptive field size. A filter traverses through the image to check for the feature by calculating a dot product between the filter weights and input pixels and carries out various operations (blurring, sharpening, edge detection, or detecting an explicit feature) to capture the temporal and spatial dependencies of the image.[5] The dot product is then fed into an output array followed by a filter shift using a step (stride). The procedure is repeated till the filter has swept across the entire image. The final filter data, i.e., a series of dot products, generates a feature map (2D array of output values), also known as an activation map or convolved feature. This entire process is called convolution. After each convolution operation, CNNs add nonlinearity to the feature map using rectified linear unit (ReLU) transformation.[6] ReLU, the most popular activation function with biological stimulation, transforms each neuron’s output and maps it to the highest possible value or zero if negative.[7] The first

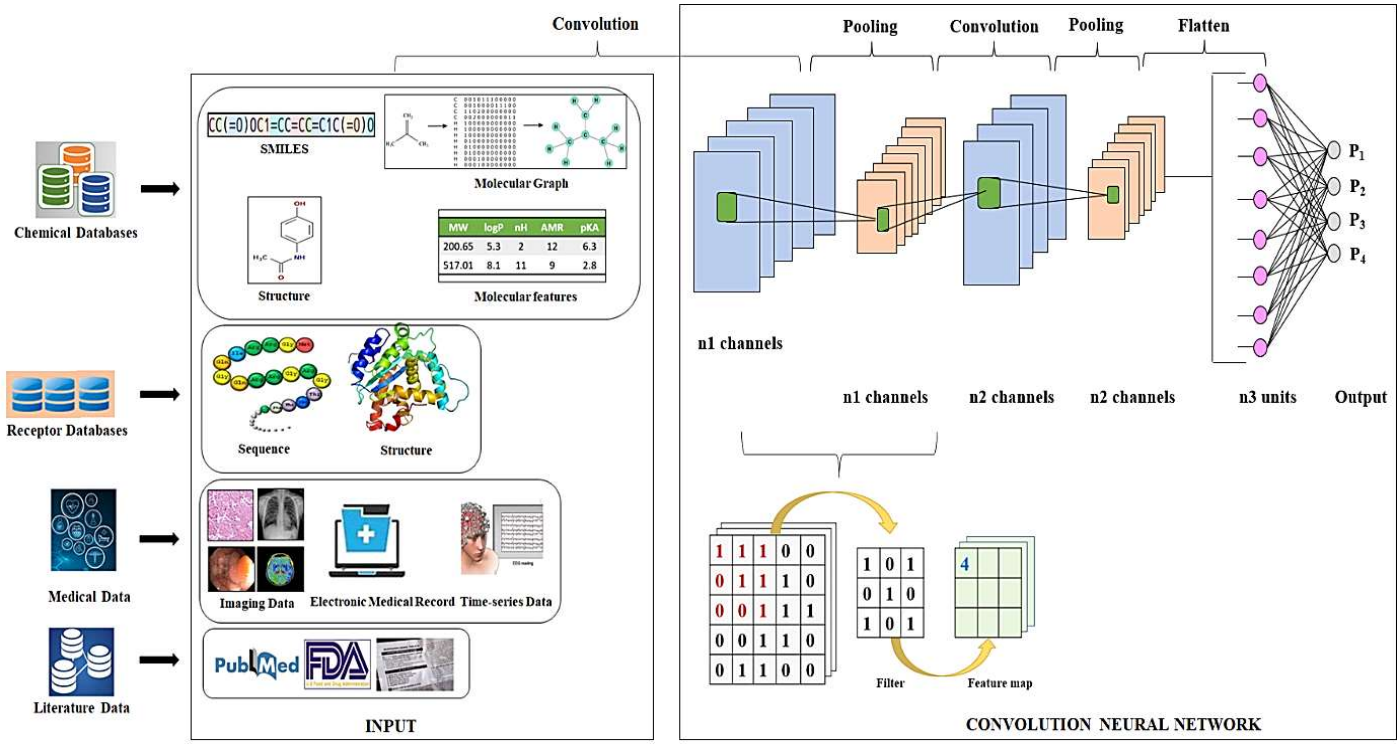

Figure 1: A simplified diagram of the applicability domain of CNN architecture.

CL can be followed by another CL, thus making the CNN's structure hierarchical, where later layers can perceive pixels from previous layers' receptive fields. Pooling layers (PL), commonly known as downsampling layers, reduce the dimensions by reducing the number of input parameters. The pooling operation sweeps a filter across the entire input, akin to the CL; however, unlike the CL, the filter does not have any weights associated with it. Instead, the filter applies an aggregation function on the values that are included inside the receptive field and returns only the most significant information from the feature map, which is passed to successive layers until the last layer.[8] The most common forms of pooling are maximum (Max) and average pooling. Max pooling, most frequently used, selects the pixel with the maximum value to send to the output array as the filter advances over the input. The average pooling method, as the name suggests, computes the average value inside the receptive field as it traverses the input and sends that information to the output array. Even though the pooling layer loses information, it improves CNN's efficiency and reduces the complexity and overfitting risk.[9] The fully connected layer (FCL) has each output layer node connecting directly to a previous layer node, in contrast to previous partially connected layers where input pixel values are not directly coupled to the output layer. The FCL gathers (flattens) all features extracted by the preceding layers and their filters to perform the classification task using softmax activation functions, i.e., to categorize inputs with a probability from 0 to 1. There are three different types of CNNs, including 1D, 2D, and 3D CNNs, depending on the input data or problem. The 1D CNN, often used with time series data or molecular features of compounds, shifts the filter in a single direction. The 2D CNNs are employed in image processing and classification tasks where the filter traverses the image in two directions. The 3D CNN utilizes 3D images from MRI or CT scans and has a kernel that moves in three directions.

The drug design and development cycle for novel drugs (small-molecule) faces several obstacles, including high cost-to-market, minimal success in clinical trials, and lengthy cycle periods.[10] Despite high expenses, the pharmaceutical industry's drug productivity continues to decline. Several factors contribute to this trend, including the complexity of obtaining approvals for novel chemical compounds, market saturation, and the tendency to spend in both developed and developing markets, among others. The potential of CNNs goes beyond computer vision, as evidenced by their successful application in a diverse range of fields, including the pharmaceutical industry.[11] Their ability to analyze complex datasets and make predictions on a range of molecular properties has led to their widespread use in the pharmaceutical industry over

Figure 2. A diagrammatic illustration of CNN applicability in a pharmaceutical domain.

the past decade, where they are used for everything from drug design and discovery to medication safety reviews (Fig 2). This review article explores various applications of CNNs in pharmaceutical sciences, including drug discovery and design, pharmacokinetics and pharmacodynamics, drug formulation and delivery, pharmaceutical manufacturing, diagnosis of disease, and classification of its progression.

Drug discovery and design

Drug design is a multifaceted process that involves designing molecules, identifying and optimizing lead compounds that are capable of interacting with specific targets in the human body or possess desirable drug-like properties. Drug development relies heavily on protein-ligand scoring because it allows researchers to reduce the vast chemical space to a manageable number of molecules for further study. CNNs have been applied in various stages of drug design and discovery, including virtual screening, de novo drug design, and structure-activity relationship (SAR) analysis as well as in generating novel compounds with desired properties. The input to the CNN is usually a SMILES notation or 3D representation of the drug molecule or target receptor, and the output is a predicted affinity score or novel compounds with the desired structure or property. Once a CNN model is trained on a large dataset of known drug-target interactions or compounds, it can be used to screen a large database of compounds to discover potential drug candidates that are likely to bind to the target or similar to the training set compounds. This can greatly speed up the process by plummeting the number of compounds that need to be experimentally tested. Several prediction models have been reported, each capable of designing or generating molecules based on the training set data. Segler et al. used SMILES notations of more than two lakh drug-like molecules from the ZINC database to generate novel molecules that were similar to the training set. The generated molecules were successfully evaluated based on their similarity to the training set and their drug-like properties.[12] In a similar study, a 3D chemical structure was used to train cycle-consistent adversarial networks (CycleGAN) for successfully generating novel compounds with similar attributes to those of the training set and predicting their drug-likeness properties.[13] CycelGAN is an image-to-image translation model that comprises two generators and two discriminators. The generator is a neural network that learns to create new data samples similar to a training dataset from noise data, while the discriminator is another neural network that learns to distinguish between real and generated data. Together, they are trained in an adversarial process to produce high-quality synthetic data. A prototype-driven diversity network, also called a generative chemistry architecture, was proposed by Harel et al. that incorporates an encoder (responsible for mapping real data samples to a lower-dimensional space), CNN, and recurrent neural network (RNN; a feedback-connected network that processes sequential data by maintaining state information from prior inputs) components to build diverse molecules with molecular template-like features without explicit prior chemical knowledge.14 Maziarka et al. introduced a novel approach called Mol-CycleGAN for molecular optimization in drug discovery. The authors proposed the use of a CycleGAN to learn the mapping between molecular structures and their corresponding properties. By training the Mol-CycleGAN on a dataset of molecules with known properties, the model generated new molecules with desired properties, thereby showcasing its potential as a powerful tool for molecular optimization in drug design and discovery.

There has been extensive usage of CNN algorithms in virtual screening to automatically extract features from the two- or three-dimensional structure of a receptor (protein or gene) to predict its binding affinities to a ligand and identify potential drug candidates.15 AtomNet, the first structure-based deep CNN application to predict binding affinity, was trained with ChEMBL dataset (comprising 78,000 actives and 2,000,000 decoys, spanning 290 targets)

using a 3D grid technique to encrypt the binding site surroundings of individual atoms into voxelized feature vectors.[16] The voxelized feature vector is a 3D grid of voxels, each of which represents an object's color, texture, material, and other characteristics. Another structure-based virtual screening study employed densely connected CNNs and a transfer learning approach (use of a pre-trained model to start a new task and allows the model to leverage knowledge learned from one domain to another) to create protein family-specific models that outperformed the machine learning (ML) benchmark models.[17] An ensemble model (combination of multiple models) based on CNN was reported by Paul et al. to improve structure-based drug design using a dataset of 22.5 million ligand poses docked into various binding locations across the Protein Data Bank (CrossDocked2020 dataset). The ensemble model classified binding positions and selected poses appropriately.[18] Shayakhmetov et al. constructed a CNN model for gene-specific virtual screening using the SMILES strings of 3,000 chemicals from the ChEMBL database to identify chemicals that activate specific genes in a transcriptome with an accuracy of 80%.[19] Another CNN model reported by Ragoza et al. used a 3D grid point of a protein-ligand complex, where the atom densities are stored at each grid point,[20] to classify correct and incorrect protein-ligand binding poses and scores for binding and nonbinding pairs. The multichannel topological neural network (TopologyNet) by Cang et al. uses a topological approach to represent the 3D biomolecular geometry[21] while maintaining important biological information to predict binding affinities. Other published models to forecast drug-receptor binding affinity based on drug SMILES, drug maximum common substructure, amino acid sequence, information about protein domains and motifs, and more are DeepDTA,[22] WideDTA,[23] DeepAffinity,[24] and PADME[25].

The potential use of CNN in drug development and material chemistry was made possible by DeepChem, a Python program. It has the potential to streamline many steps in the drug development process, including virtual screening, molecular property prediction, and molecule production. The platform is intuitive and flexible, allowing researchers to fine-tune their models for usage with a wide variety of datasets and queries.[26] Several parallel CNN architectures were trained to identify chemicals based on their toxicity and to predict the activity of certain drugs utilizing molecular descriptors, images, and genomic data. Some models reported attaining accuracies as high as 99%, albeit this ranged widely depending on the dataset and the model employed. Overall, CNNs have shown great promise in drug design and discovery and have the potential to revolutionize the field by enabling the design of new drugs more quickly and cost-effectively than traditional methods.

Pharmacokinetics and Pharmacodynamics

Pharmacokinetics (PK) and pharmacodynamics (PD) are important aspects of drug development that involve the study of the absorption, distribution, metabolism, and excretion (ADMET) of drugs, as well as their effects on the body. CNNs have been used to predict the PK and PD properties of drugs such as solubility, toxicity, and bioactivity, based on their chemical structure and physicochemical properties. Chemception, a deep CNN model, predicts periodicity, molecular descriptors and fingerprints, toxicity, activity, and solvation characteristics using 2D molecular images.[27] Several groups have applied CNNs using molecular properties to predict various properties, such as aqueous solubility,[28] blood-brain permeability,[29] etc. Recently, graph convolutional neural networks (GCNN),[30] a variant of CNN has been successfully applied for predicting molecular properties.[31] GCNN first turns the topology of a molecule into a graph, with bonds as the edges and atoms as the nodes. It then creates hierarchical representations of molecules based on how far apart the bonds are around the atomic centers. This encoding gives "neural fingerprints'' that are rich, multiscale vectorial representations of molecules that are fed into additional layers of the neural network to classify activity.[32]

Additionally, CNNs have been applied in the development of personalized dosing regimens based on patient-specific factors. Deep-dose, a novel CNN-based method for estimating the distribution of personalized internal radiation dosage, leverages information from medical imaging scans to construct a personalized 3D model of a patient's anatomy. It effectively predicted the radiation dose distribution with an average error of less than 10%.[33] Similarly, another study reported the development of a CNN model using computed tomography (CT) images and contour masks as input to predict the dose distribution for nasopharyngeal carcinoma patients (NPC) treated with tomotherapy with high precision. The model provides a therapeutically viable result

and a training strategy for a dose prediction model experimentally.[34] Another study explored the application of the CNN model (Tox_R) to predict toxicity from DNA-specific fluorescent probe (DAPI)-stained cell images using several drugs with different toxicity mechanisms. Tox_R was able to automatically generate feature maps that categorized drugs by mechanism of action and further extrapolated the results to categorize nuclei and predict per-cell toxicity from raw screening images using fully automated region-based CNNs (RCNN).[35] Future studies may incorporate genomic and proteome data into the CNNs framework to improve PK and PD prediction and enable personalized treatment.

Disease Diagnosis and progression

CNNs have performed exceptionally well in diagnosing diseases using data from various imaging techniques such as radiology, magnetic resonance (MR), CT, X-ray, and microscope imaging.[36] CNN training traditionally began with image pre-processing, including scaling and augmentation. Data augmentation randomly selects a few images to flip vertically and horizontally, adjust height and breadth, and zoom up to a certain percentage. CNN models have demonstrated promising applications in the diagnosis of diabetic retinopathy,[37] skin lesion classification,[38] identification of lymph node metastasis,[39] lung nodule classification,[40-42] gastrointestinal disease classification,[43-45] tuberculosis,[46] and cancers such as breast cancer,[47] brain tumors,[48] lung cancer,[49,50] etc. Several researchers have reported the successful implementation of CNN algorithms, with high accuracy, to diagnose infections from chest X-rays, which could aid in the early identification and care of COVID-19 patients.[51-54] One such system is COVID-Net, an open-source diagnosis system, based on a deep CNN for detecting COVID-19 patients using X-ray images.[55] Mahmud et al. designed a deep CNN architecture, CovXNet, that makes use of depthwise convolution with different dilation rates to extract various features from chest X-rays.[56] Another intelligent diagnostic system using switch-controllable nanocatcher and CNN has been developed by Feng et al. to analyze pathological images for Cryptococcus infections (a common cause of illness and death in HIV/AIDS patients).[57]

In addition to classifying medical images, CNNs have also been used to identify or annotate the regions of abnormality within the images, for example, segmentation of tumor region in the uterus,[58,59] polyps in the colon,[60] liver,[61] etc. For this, a probability map of the organ or anatomical structure is first constructed using CNN and image patches, and then the segmentation is refined utilizing both the probability map and the overall context of the images. Unlike MR, ultrasound, and CT, microscope imaging has complex characteristics. Stains, background clutter, inhomogeneous intensity, contacting or overlapping nuclei/cells, etc., make it difficult to manually interpret pathology images. Due to their ability to learn complex features and patterns, CNNs have been successfully implemented to identify the cell nucleus, cell count, cell area, and mitosis in the microscopic images of different cancers (brain, lung, cervical, and brain tumors) without prior knowledge.[62-70] In a nutshell, various studies have underlined the promising role of CNN-based intelligent diagnostic systems for early and fast diagnosis of diseases with high sensitivity and specificity.

Drug-Drug Interactions

Interactions between drugs are referred to as drug-drug interactions (DDIs) which are classified as synergistic, antagonistic, or neutral. DDIs are crucial to drug development and disease detection, yet they require substantial investments of time, money, and resources.[71,72] The earlier DDI prediction models focused on the biological interactions between drugs only rather than the intricate atomic interactions. In contrast, the CNN-based DDI prediction models built are more robust and account for atomic interactions. The CNN model described by Liu et al. was the first to apply the CNN algorithm for DDI prediction based on position embeddings and word embeddings, which capture the relative distances between words and semantic information of words for two drugs of interest.[73] Quan et al. resolved the three major issues associated with extracting biological relations from medical records (namely vocabulary gap, incorporation of semantic information, and manual feature selection) by integrating the CNN framework with multichannel word embedding (MCCNN) to predict DDI.[74] MCNN consists of three components: improving word representation by increasing vocabulary and decreasing unfamiliar words, application of attention mechanism, and using DL models for prediction. DeepCNN, a 10-layer CNN architecture, showed improvement in extracting DDI information by building a high-quality learning representation of long input sequences using multi-channel word embedding.[75]

CNN-DDI, a semi-supervised CNN system (comprising of five convolutional, two fully connected, and a CNN-based softmax layer) by Zhang et. al. predicts DDI and associated events using input vectors comprising feature interactions from pharmacological categories, targets, pathways, and enzymes.[76] To forecast DDI incidents, Yang et al. suggested the CNN-Siam algorithm, which makes use of the drug multimodal information and a Siamese neural network (SNN) architecture. SNNs have two or more identical subnetworks with the same configuration, specifications, and weights. All subnetworks undergo identical parameter updates, and their output feature vectors are compared to identify input-output pairs with shared characteristics. CNN-Siam learns a representation for an individual drug by feeding its chemical substructure, target, and enzyme information into two CNNs with shared parameters; the resulting drug pair representations are then fused and further sent to a multilayer perceptron for classification.[77] Another end-to-end model, an attention-convolutional neural network (ACNN), for predicting DDI from solely drug sequence information (feature matrix) was given by Wang et al. By assigning a different attention vector to each atom in the drug feature matrix, ACNN was able to simulate the intricate interaction between the drug atoms.[78] A unique neural technique to extract DDIs from texts by utilizing drug molecular structure was proposed by Asada et al. where textual drug pairs are encoded using CNN, while their molecular pairs are encoded using graph convolutional networks (GCNs). The outputs of these two networks are then combined for final prediction.[79] CNN-based models have proven their usefulness in extracting and learning different drug representations to predict DDIs, allowing for more accurate prediction of drug efficacy and safety.

Drug Safety and Adverse Event Prediction:

Drug safety is a critical aspect of drug development and involves the identification and mitigation of potential adverse drug reactions (ADRs). While ML algorithms have been widely utilized for ADR prediction, only a small number of CNN-based models currently exist; nonetheless, research into this area has shown that CNNs can be useful for predicting and detecting ADRs, which can lead to their early discovery and elimination. Current CNN models predict the safety profile of drugs based on their chemical structure and physicochemical properties. Chen et al. presented a novel CNN model to predict drug ADRs with 88% accuracy from chemical structures by using bioactivity data of FDA-approved drugs.80 Yao et al. successfully trained a CNN model with semantic embedding to predict the likelihood of ADRs for a given drug using a dataset of drug labels and their associated ADRS from the US Food and Drug Administration (FDA).[81] In another study Dey et al. used SIDER database drug-ADR information, to develop a neural fingerprint technique that investigates all possible substructures present in drug molecules up to a certain radius, where the radius is defined as half of the maximum path length between any two atoms of that substructure, and trained a CNN framework with an attention mechanism to identify which drug-molecule substructures linked to a given ADR and also determines if the substructures could be used to anticipate ADRs in novel drugs. This study is helpful for drug developers to discover problematic substructures and may improve pipeline drug safety reviews.[82]

Additionally, CNNs could be applied to predict adverse events based on patient-specific factors, such as genetic information, electronic medical/health records (EMR/HER), and medical history.

Pharmaceutical Manufacturing:

Pharmaceutical manufacturing involves the production of drugs in large quantities while ensuring consistent quality and purity. CNNs have been used to optimize various aspects of pharmaceutical manufacturing, such as the optimization of production processes and the detection of defects in drug products. Additionally, CNNs have been applied in the development of automated systems for quality control and assurance. Optical coherence tomography (OCT) is a real-time and contactless process analytical technology (PAT) for solid dosage form coating operations in the pharmaceutical industry. Researchers have successfully employed CNNs to evaluate OCT images for pharmaceutical solid dosage forms using image data from both in- and at-line OCT implementations and for monitoring film-coated tablets and single- and multi-layer pellets.[83] Researchers have also used CNNs to detect defects in tablet production and to monitor the quality of drug products during the manufacturing process.[84]

Drug Formulation and Delivery:

Drug formulation and delivery is a crucial aspect of drug development that involves the development of

safe and effective drug delivery systems. CNNs have been used to optimize drug formulations by predicting the stability, solubility, and bioavailability of drugs based on their physicochemical properties. Additionally, CNNs have been applied in the development of drug delivery systems, such as liposomes and nanoparticles, to improve drug targeting and efficacy. For example, researchers have used CNNs to predict the optimal size and shape of nanoparticles for drug delivery.[85] Recently, using the position of the drug injection and the geometry of the blood vessel as inputs, CNN has been successfully used to develop a data-driven reduced-order model (ROM) for real-time prediction of the spatial-temporal drug trajectory and concentration field in trans-arterial chemoembolization therapy.[86]

Personalized Medicine:

To capture the unique characteristics of each patient, personalized prediction models draw data from cohorts of patients with comparable characteristics. CNNs have demonstrated remarkable performance in tasks such as tumor detection, segmentation, and classification, aiding in personalized medicine by providing precise and individualized insights.[87] These networks can capture subtle visual cues and intricate spatial relationships in medical images, enabling the identification of biomarkers, early disease detection, and assisting in treatment planning. The article highlights the potential of CNNs in revolutionizing radiology and emphasizes their contribution to personalized medicine through the accurate and efficient analysis of medical images. Suo et al. introduced a novel time fusion CNN framework that learns patient representations and measures pairwise similarity, considering the temporal relationships and contributions from different time intervals.[88] The framework aims to accurately identify similar patients and utilizes the similarity scores for personalized disease predictions, evaluating the effects of different vector representations and similarity learning metrics.

Other studies

The CNN application is not just limited to drug research, but it has also proven to be useful in other studies. Beck et al. utilized a pre-trained hybrid CNN and RNN model called molecule transformer-drug target interaction (MT-DTI) to predict whether any commercial antiviral medicines would function in SARS-CoV-2. Remdesivir, ritonavir, atazanavir, efavirenz, and dolutegravir were computationally identified for SARS-CoV-2 treatment.[89] DeeplyTough, a 3D CNN model, quantifies pocket similarity by analysing binding sites of proteins without alignment. It encodes 3D protein pockets into descriptor vectors to compute pairwise Euclidean distances using a positive and negative selection of proteins that bind chemically similar and dissimilar ligands.[90] CNN has also exhibited success predicting protein structures in the absence of a template structure.[91] Tong et al. employed a CNN model for analyzing protein structures to forecast how mutations may alter protein structure, and study amino acid microenvironments without any prior information or feature assumptions. The method was reported to double the accuracy of predictions compared to models that necessitated the selection of features by hand.[92] Images of cells pre-treated with a variety of drugs have been used by CNNs to make predictions about their toxicity. The model was able to effectively predict a broad variety of toxicity pathways from different medicines, nuclear stains, and cell lines.[93] Finally, to extract biomedical relationships from PubMed literature, Nourani et al. created a hybrid transfer learning framework (Deep-GDAE) using attrition-based BiLSTM and a CNN based on textual feature extracted using BERT (Bidirectional Encoder Representations from Transformers, a pre-trained natural language processing model) and BioBERT base models.[94]

Future prospects and challenges

CNNs have demonstrated exponential development across all fields as they excel at finding patterns and features in various types of data (textual, image, audio, video, time-series) with minimal pre-processing, handling enormous volumes of data with end-to-end training, attaining excellent accuracy while being robust to scaling, rotation, and translation invariance. It correctly performs the task (classification, segmentation) by reducing the image down to its fundamental features via convolutions and pooling mechanisms. Compared to conventional neural networks, they are easier to train owing to relatively few initial parameters and the capacity of convolutions to manage all hidden layer discoveries. CNNs have also shown great promise in various applications of pharmaceutical science, such as drug discovery, molecular property prediction, and image-based-disease diagnosis. CNNs could speed up drug discovery by evaluating enormous databases of

chemical compounds and identifying promising drug candidates or molecular targets (receptors) based on their molecular features or structure, thereby reducing the time and expense of drug development. CNNs could be trained to design drug formulation strategies, predict solubility (a critical factor in drug formulation), and identify manufacturing defects. Another area where CNNs can be employed directly with pharmaceutical applications is analysing microscopic images of cells and tissues to identify disease states or drug effects. CNNs can also be used in personalized medicine by analysing genetic data to predict drug responses or identify potential adverse reactions. This could lead to more effective and safer treatment options for patients. Overall, the future of CNNs in pharmaceutical science is very promising. As technology advances, we can expect to see even more applications in drug discovery, personalized medicine, and other areas of pharmaceutical research.

Despite the promising results obtained and bright future, there are some challenges associated with CNNs that need to be addressed. First, training a CNN model is computationally taxing and requires considerable memory, necessitating the use of graphics processing units (GPUs). Second, the lack of large, diverse, and well-labeled datasets or appropriate regularization limits CNN performance, and they are susceptible to overfitting. Third, as the CNN's depth increases, more and more input-related data may be lost or "washed out" before it reaches the output layer.2, 3,[95] Short pathways between layers and feed-forward networks, in which each layer is connected to all other layers, have been offered as potential solutions to these and related issues.[96] The application of CNNs in pharmaceuticals is still in its early stages, and more research is required to fully explore its potential. Despite the challenges, CNNs offer a promising approach to improving drug development and patient outcomes.

References

1. Mouchlis V D, Afantitis A, Serra A, Fratello M, Papadiamantis A G, Aidinis V, Lynch I, Greco D and Melagraki G 2021 Advances in de Novo Drug Design: From Conventional to Machine Learning Methods. Int. J. Mol. Sci. 22(4) 1676
2. LeCun Y and Bengio Y. Convolutional Networks for Images, Speech, and Time Series. In The Handbook of Brain Theory and Neural Networks, MIT Press, Cambridge, 1995.
3. LeCun Y, Kavukcuoglu K and Farabet C 2010 Convolutional Networks and Applications in Vision. In

Abbreviations

ACNN	Attention Neural Network
ADMET	Absorption, Distribution, Metabolism, Excretion, And Toxicity
ADR	Adverse Drug Reaction
ANN	Artificial Neural Network
BERT	Bidirectional Encoder Representation
CL	Convolutional Layer
CNN	Convolutional Neural Network
CYCLEGAN	Cycle-Consistent Adversarial Network
DAPI	DNA-Specific Fluorescent Probe
DDI	Drug-Drug Interactions
DL	Deep Learning
EMR	Electronic Medical Record
FCL	Fully Connected Layer
FDA	Food And Drug Administration
GCNN	Graph Convolutional Neural Network
GPU	General Processing Unit
HER	Health Electronic Record
LSTM	Long Short-Term Model
MCCNN	Multichannel Word Embedding CNN
MGT-DTI	Molecule Transformer-Drug Target Interaction
ML	Machine Learning
MR	Magnetic Resonance
NPC	Nasopharyngeal Carcinoma
OCT	Optical Coherence Tomography
PAT	Process Analytical Technology
PD	Pharmacodynamics
PK	Pharmacokinetics
PL	Pooling Layer
RCNN	Region-Based Convolutional Neural Network
RELU	Rectified Linear Unit
RGB	Red, Green, Blue
RNN	Recurrent Neural Network
ROM	Reduced Order Model
SAR	Structure-Activity Relationship
SNN	Siamese Neural Network

Proceedings of 2010 IEEE International Symposium on Circuits and Systems. 253
4. Currie G, Hawk K E, Rohren E, Vial A and Klein R 2019 Machine Learning and Deep Learning in Medical Imaging: Intelligent Imaging. J. Med. Imaging Radiat. Sci. 50(4) 477
5. Ghosh A, Sufian A, Sultana F, Chakrabarti A and De D 2020 Fundamental Concepts of Convolutional Neural Network. In: Balas V, Kumar R, Srivastava R. (eds) Recent Trends and Advances in Artificial Intelligence and Internet of Things. Intelligent Systems Reference Library. vol 172. Springer, Cham.
6. Sakib S, Ahmed N, Kabir A J and Ahmed H 2018 An Overview of Convolutional Neural Network: Its Architecture and Applications. Preprints.org 2018110546.
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 (ICML'10). Omnipress, Madison, WI, USA, 807
8. Jiménez-Luna G, Grisoni J F and Schneider G 2020 Drug discovery with explainable artificial intelligence. Nat. Mach. Intell. 2 573
9. Goumiri S, Benboudjema D and Pieczynski W 2023 A new hybrid model of convolutional neural networks and hidden Markov chains for image classification. Neural. Comput. Appl. 31 1
10. Velmurugan D, Pachaiappan R and Ramakrishnan C 2020 Recent Trends in Drug Design and Discovery. Curr. Top. Med. Chem. 20(19) 1761
11. Lindsay G W 2021 Convolutional Neural Networks as a Model of the Visual System: Past, Present, and Future. J. Cogn. Neurosci. 33(10) 2017
12. Segler M H S, Kogej T, Tyrchan C and Waller M P 2018 Generating Focused Molecule Libraries for Drug Discovery with Recurrent Neural Networks. ACS Cent. Sci. 4 120
13. Maziarka L, Pocha A, Kaczmarczyk J, Rataj K, Danel T and Warchol M 2020 Mol-CycleGAN: a generative model for molecular optimization. J. Cheminform. 12 2
14. Harel S and Radinsky K 2018 Prototype-Based Compound Discovery Using Deep Generative Models. Mol. Pharm. 15 4406
15. Maziarka L, Pocha A, Kaczmarczyk J, Rataj K, Danel T and Warchol M 2020 Mol-CycleGAN: a generative model for molecular optimization. J. Cheminform. 12 2
16. Wallach I, Dzamba M and Heifets A 2015 AtomNet: a deep, convolutional neural network for bioactivity prediction in structure-based drug discovery. arXiv:1510.02855v1
17. Imrie F, Bradley A R, van der Schaar M and Deane C M 2018 Protein Family Specific Models Using Deep Neural Networks and Transfer Learning Improve Virtual Screening and Highlight the Need for More Data. J. Chem. Inf. Model 58 2319
18. Francoeur P G, Masuda T, Sunseri J, Jia A, Iovanisci R B, Snyder I and Koes D R 2020 Three-dimensional convolutional Neural Networks and a Cross-Docked Data Set for Structure-Based Drug Design. J. Chem. Inf. Model. 60 4200.
19. Shayakhmetov R, Kuznetsov M, Zhebrak A, Kadurin A, Nikolenko S, Aliper A and Polykovskiy D 2020 Molecular Generation for Desired Transcriptome Changes with Adversarial Autoencoders. Front. Pharmacol. 11 269
20. 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
21. Cang Z and Wei G 2017 TopologyNet: Topology Based Deep Convolutional and Multi-Task Neural Networks for Biomolecular Property Predictions. PLoS Comput. Biol. 13 e1005690
22. Mahmud S H, Chen W, Jahan H, Liu Y, Sujan N I and Ahmed S 2019 iDTi-CSsmoteB: identification of drug–target interaction based on drug chemical structure and protein sequence using XGBoost with over-sampling technique SMOTE. IEEE Access. 7 48699
23. Gao K Y, Fokoue A, Luo H, Iyengar A, Dey S and Zhang P 2018 Interpretable drug target prediction using deep neural representation. In Proceedings of the Twenty-Seventh International Joint Conference on Artificial Intelligence (IJCAI) 3371
24. Karimi M, Wu D, Wang Z and Shen Y 2019 DeepAffinity: interpretable deep learning of compound-protein affinity through unified recurrent and convolutional neural networks. Bioinformatics. 35 3329
25. Feng Q, Dueva E, Cherkasov A and Ester M 2018 Padme: a deep learning-based framework for drug–target interaction prediction. arXiv arXiv:1807.09741
26. Bharath R, Peter E, Patrick W, and Vijay P 2019 Deep Learning for the Life Sciences: Applying Deep Learning to Genomics, Microscopy, Drug Discovery, and More. Sebastopol, CA: O'Reilly Media
27. Goh G, Siegel C, Vishnu A, Hodas N and Baker N 2017 Chemception: A Deep Neural Network with Minimal Chemistry Knowledge Matches the Performance of Expert-developed QSAR/QSPR Models. arXiv:1706.06689
28. Lusci A, Pollastri G and Baldi P 2013 Deep architectures and deep learning in chemoinformatics: the prediction of aqueous solubility for drug-like molecules. J. Chem. Inf. Model. 53 1563
29. Alsenan S, Al-Turaiki I and Hafez A 2021 A deep learning approach to predict blood-brain barrier permeability. J. Comput. Sci. 7 e515
30. Weber J K, Morrone J A, Bagchi S, Pabon J D E, Kang S G, Zhang L and Cornell W D 2022 Simplified, interpretable graph convolutional neural networks for small molecule activity prediction. J. Comput. Aided Mol. Des. 36 391
31. Ståhl N, Falkman G, Karlsson A, Mathiason G and Boström J 2018 Deep Convolutional Neural Networks for the Prediction of Molecular Properties: Challenges and Opportunities Connected to the Data. J. Integr. Bioinform. 16 20180065
32. 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
33. Lee M S, Hwang D, Kim J H and Lee J S 2019 Deep-dose: a voxel dose estimation method using deep convolutional neural network for personalized internal dosimetry. Sci. Rep. 9 10308
34. Liu Y, Chen Z, Wang J, Wang X, Qu B, Ma L, Zhao W, Zhang G and Xu S 2021 Dose Prediction Using a Three-Dimensional Convolutional Neural Network for Nasopharyngeal Carcinoma with Tomotherapy. Front. Oncol. 11 752007
35. Jimenez-Carretero D, Abrishami V, Fernández-de-Manuel L, Palacios I, Quílez-Álvarez A, Díez-Sánchez A, Del Pozo M A and Montoya M C 2018 Tox_(R)CNN: Deep learning-based nuclei profiling tool for drug toxicity screening. PLoS Comput. Biol. 14 e1006238
36. Yamashita R, Nishio M, Do R K G and Togashi K 2018 Convolutional neural networks: an overview and application in radiology. Insights Imaging. 9 611
37. Gulshan V, Peng L, Coram M, Stumpe M C, Wu D, Narayanaswamy A, Venugopalan S, Widner K, Madams

T, Cuadros J, Kim R, Raman R, Nelson P C, Mega J L andWebster D R 2016 Development and Validation of a Deep Learning Algorithm for Detection of Diabetic Retinopathy in Retinal Fundus Photographs. JAMA. 316 2402

38. Esteva A, Kuprel B, Novoa R A, Ko J, Swetter S M, Blau H M and Thrun S 2017 Dermatologist-level classification of skin cancer with deep neural networks. Nature. 542 115
39. Ehteshami Bejnordi B, Veta M, Johannes van Diest P, van Ginneken B, Karssemeijer N, Litjens G, van der Laak JAWM and the CAMELYON16 Consortium 2017 Diagnostic Assessment of Deep Learning Algorithms for Detection of Lymph Node Metastases in Women with Breast Cancer. JAMA. 318 2199
40. Liang M, Tang W, Xu D M, Jirapatnakul A C, Reeves A P, Henschke C I and Yankelevitz D 2016 Low-Dose CT Screening for Lung Cancer: Computer-aided Detection of Missed Lung Cancers. Radiology. 281 279
41. van Ginneken B, Armato SG 3rd, de Hoop B, van Amelsvoort-van de Vorst S, Duindam T, Niemeijer M, Murphy K, Schilham A, Retico A, Fantacci ME, Camarlinghi N, Bagagli F, Gori I, Hara T, Fujita H, Gargano G, Bellotti R, Tangaro S, Bolaños L, De Carlo F, Cerello P, Cristian Cheran S, Lopez Torres E and Prokop M 2010 Comparing and combining algorithms for computer-aided detection of pulmonary nodules in computed tomography scans: The ANODE09 study. Med. Image Anal. 14 707
42. Pedersen J H, Ashraf H, Dirksen A, Bach K, Hansen H, Toennesen P, Thorsen H, Brodersen J, Skov B G, Døssing M, Mortensen J, Richter K, Clementsen P and Seersholm N 2009 The Danish randomized lung cancer CT screening trial--overall design and results of the prevalence round. J. Thorac. Oncol. 4 608
43. Fan S, Xu L, Fan Y, Wei K and Li L 2018 Computer-aided detection of small intestinal ulcer and erosion in wireless capsule endoscopy images. Phys. Med. Biol. 63 165001
44. Liaqat A, Attique M K, Jamal H S, Muhammad S, Yasmin M and Fernandes S L 2018 Automated ulcer and bleeding classification from WCE images using multiple features fusion and selection. J. Mech. Med. Bio. 18 1850038
45. Sharif M, Md. Attique K, Rashid M, Yasmin M, Afza F and Tanik U J 2021 Deep CNN and geometric features-based gastrointestinal tract diseases detection and classification from wireless capsule endoscopy images. J. Exp. Theo. Artificial Intelli. 33 577-599
46. Lakhani P and Sundaram B 2017 Deep Learning at Chest Radiography: Automated Classification of Pulmonary Tuberculosis by Using Convolutional Neural Networks. Radiology. 284 574
47. Kavitha T, Mathai P P, Karthikeyan C, Ashok M, Kohar R, Avanija J and Neelakandan S 2022 Deep Learning Based Capsule Neural Network Model for Breast Cancer Diagnosis Using Mammogram Images. Interdiscip Sci. 14 113
48. Lotlikar V S, Satpute N and Gupta A 2022 Brain Tumor Detection Using Machine Learning and Deep Learning: A Review. Curr. Med. Imaging. 18 604
49. Aberle D R, Adams A M, Berg C D, Black W C, Clapp J D, Fagerstrom R M, Gareen I F, Gatsonis C, Marcus P M and Sicks J D 2011 Reduced lung-cancer mortality with low-dose computed tomographic screening. N. Engl. J. Med. 365 395
50. Nishio M, Nagashima C, Hirabayashi S, Ohnishi A, Sasaki K, Sagawa T, Hamada M and Yamashita T 2017 Convolutional auto-encoder for image denoising of ultra-low-dose CT. Heliyon. 3 e00393
51. Kugunavar S and Prabhakar C J 2021 Convolutional neural networks for the diagnosis and prognosis of the coronavirus disease pandemic. Vis. Comput. Ind. Biomed. Art 4, 12
52. Reshi A A, Rustam F, Mehmood A, Alhossan A, Alrabiah Z, Ahmad A and Alsuwailem H 2021 An Efficient CNN Model for COVID-19 Disease Detection Based on X-Ray Image Classification. Complexity. 1076
53. Mukherjee H, Ghosh S, Dhar A, Obaidullah S M, Santosh KC and Roy K 2021 Shallow Convolutional Neural Network for COVID-19 Outbreak Screening Using Chest X-rays. Cognit. Comput. 1 14
54. Narin A, Kaya C and Pamuk Z 2021 Automatic detection of coronavirus disease (COVID-19) using X-ray images and deep convolutional neural networks. Pattern Anal. Appl. 24 1207
55. Wang L, Lin Z Q and Wong A 2020 COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-ray images. Sci. Rep. 10 19549
56. Mahmud T, Rahman M A and Fattah S A 2020 CovXNet: A multi-dilation convolutional neural network for automatic COVID-19 and other pneumonia detection from chest X-ray images with transferable multi-receptive feature optimization. Comput. Biol. Med. 122 103869
57. Feng G, Shan H, Benshun T, Tianyi M, Yu X, Ying Y and Bing G 2023 Intelligent diagnostic system for Cryptococcus: Switch-controllable nanocatcher and CNN-based artificial intelligence. Chem. Eng. J. 465, 142674
58. Lucchesi F R and Aredes N D 2016 Radiology data from The Cancer Genome Atlas Cervical Squamous Cell Carcinoma and Endocervical Adenocarcinoma (TCGA-CESC) collection. The Cancer Imaging Archive. https://doi.org/10.7937/K9/TCIA.2016.SQ4M8YP4
59. Kurata Y, Nishio M, Kido A, Fujimoto K, Yakami M, Isoda H and Togashi K 2019 Automatic segmentation of the uterus on MRI using a convolutional neural network. Comput. Biol. Med. 114 103438
60. Mehshan A K, Muhammad A K, Fawad A, Mamta M, Lalit M, Goyal D, Jude H and Suresh CS 2020 Gastrointestinal diseases segmentation and classification based on duo-deep architectures. Patt. Recog. Letters. 131 193
61. Lu F, Wu F, Hu P, Peng Z and Kong D 2017 Automatic 3D liver location and segmentation via convolutional neural network and graph cut. Int. J. Comput. Assist Radiol. Surg. 12 171
62. M Veta, van Diest P J and Pluim J P W 2016 Cutting out the middleman: Measuring nuclear area in histopathology slides without segmentation. In Proc. 19th Int. Conf. Med. Image Comput. Comput. Assist. Intervent. 632
63. Xu Z and Huang J 2016 Detecting 10,000 cells in one second. In: Ourselin S, Joskowicz L, Sabuncu M, Unal G, Wells W (eds) Medical Image Computing and Computer-Assisted Intervention – MICCAI. Lecture Notes in Computer Science, Springer, Cham 9901
64. Ciresan D C, Giusti A, Gambardella L M and Schmidhuber J 2013 Mitosis detection in breast cancer histology images with deep neural networks. Med. Image Comput. Comput. Assist Interv. 16 411
65. Xing F, Xie Y and Yang L 2016 An Automatic Learning-Based Framework for Robust Nucleus Segmentation. IEEE Trans. Med. Imaging. 35 550
66. Liu F and Yang L 2015 A novel cell detection method using deep convolutional neural network and maximum-weight independent set. In Proc. 18th Int. Conf. Med. Image Comput. Comput. Assist. Intervent. 9351 349
67. Lapierre-Landry M, Liu Z, Ling S, Bayat M, Wilson D L

and Jenkins M W 2021 Nuclei Detection for 3D Microscopy with a Fully Convolutional Regression Network. IEEE Access. 9 60396
68. Xie Y, Kong X, Xing F, Liu F, Su H and Yang L 2015 Deep Voting: A Robust Approach Toward Nucleus Localization in Microscopy Images. Med. Image Comput. Comput. Assist. Interv. 9351 374
69. Albarqouni S, Baur C, Achilles F, Belagiannis V, Demirci S and Navab N 2016 AggNet: Deep Learning from Crowds for Mitosis Detection in Breast Cancer Histology Images. IEEE Trans. Med. Imaging. 35 1313
70. Xu J, Xiang L, Liu Q, Gilmore H, Wu J, Tang J and Madabhushi A 2016 Stacked Sparse Autoencoder (SSAE) for Nuclei Detection on Breast Cancer Histopathology Images. IEEE Trans. Med. Imaging. 35 119
71. Liu S, Tang B, Chen Q and Wang X 2016 Drug-Drug Interaction Extraction via Convolutional Neural Networks. Comput. Math Methods Med. 6918381
72. Hiroyuki K 2014 How far should we go? Perspective of drug-drug interaction studies in drug development. Drug Metab. Pharmacokinet. 29 227
73. Liu S, Tang B, Chen Q and Wang X 2016 Drug-Drug Interaction Extraction via Convolutional Neural Networks. Comput. Math Methods Med. 6918381
74. Quan C, Hua L, Sun X and Bai W 2016 Multichannel Convolutional Neural Network for Biological Relation Extraction. Biomed Res. Int. 1850404
75. Dewi I N, Dong S and Hu J 2017 Drug-drug interaction relation extraction with deep convolutional neural networks. In IEEE Int. Conf. on Bioinfo. Biomed. (BIBM), Kansas City, USA. 795
76. Zhang C, Lu Y and Zang T 2022 CNN-DDI: a learning-based method for predicting drug-drug interactions using convolution neural networks. BMC Bioinformatics. 23 88
77. Yang Z, Tong K, Jin S, Wang S, Yang C and Jiang F 2023 CNN-Siam: multimodal siamese CNN-based deep learning approach for drug-drug interaction prediction. BMC Bioinformatics. 24 110
78. Wang W and Liu H 2022 ACNN: Drug-Drug Interaction Prediction Through CNN and Attention Mechanism. In: Huang D S, Jo KH, Jing J, Premaratne P, Bevilacqua V, Hussain A. (eds) Intelligent Computing Theories and Application. ICIC Lecture Notes in Computer Science. Springer, Cham. 13394
79. Asada M, Miwa M and Sasaki Y 2018 Enhancing drug–drug interaction extraction from texts by molecular structure information. arXiv:1805.05593
80. Chen H, Engkvist O, Wang Y, Olivecrona M and Blaschke T 2018 Using Convolutional Neural Networks to Predict Adverse Drug Reactions from Bioactivity Data. J. Chem. Info. Mod. 58 2101
81. Yao L, Mao C, Luo Y and Zhang J 2018 Predicting adverse drug reactions with convolutional neural networks and semantic embedding. IEEE/ACM Trans. Comput. Biol. Bioinform. 15 1865
82. Dey S, Luo H, Fokoue A, Hu J and Zhang P 2018 Predicting adverse drug reactions through interpretable deep learning framework. BMC Bioinfor. 19 476
83. Wolfgang M, Weißensteiner M, Clarke P, Hsiao W K and Khinast J G 2020 Deep convolutional neural networks: Outperforming established algorithms in the evaluation of industrial optical coherence tomography (OCT) images of pharmaceutical coatings. Int. J. Pharm. X. 2 100058
84. Junlin Z J H, Guoli L and Yongbin L 2020 Identifying Capsule Defect Based on an Improved Convolutional Neural Network. Shock and Vibration. 88877232020
85. Harrison P J, Wieslander H, Sabirsh A, Karlsson J, Malmsjö V, Hellander A, Wählby C and Spjuth O 2021 Deep-learning models for lipid nanoparticle-based drug delivery. Nanomedicine. 16 1097
86. Yuan X Y, Hua Y, Aubry N, Zhussupbekov M, Antaki J F, Zhou Z F and Peng J Z 2022 Real-Time Prediction of Transarterial Drug Delivery Based on a Deep Convolutional Neural Network. Applied Sciences. 12 10554
87. Mazurowski M A, Buda M, Saha A and Bashir M R (2019) Deep learning in radiology: An overview of the concepts and a survey of the state of the art with focus on MRI. J Magn. Reson. Imaging. 49(4) 939–954.
88. Suo Q, Ma F, Yuan Y, Huai M, Zhong W, Zhang A and Gao J 2017 Personalized disease prediction using a CNN-based similarity learning method. IEEE International Conference on Bioinformatics and Biomedicine (BIBM), Kansas City, MO, USA 811-816.
89. Beck B R, Shin B, Choi Y, Park S and Kang K 2020 Predicting commercially available antiviral drugs that may act on the novel coronavirus (SARS-CoV-2) through a drug-target interaction deep learning model. Comput. Struct. Biotechnol J. 18 784
90. Simonovsky M and Meyers J 2020 DeeplyTough: learning structural comparison of protein binding sites. J. Chem. Inf. Model. 60 2356
91. Kinch L N, Kryshtafovych A, Monastyrskyy B and Grishin N V 2019 CASP13 target classification into tertiary structure prediction categories. Proteins. 87 1021
92. Torng W and Altman R B 2017 3D deep convolutional neural networks for amino acid environment similarity analysis. BMC Bioinformatics. 18 302
93. Jimenez-Carretero D, Abrishami V, Fernández-de-Manuel L, Palacios I, Quílez-Álvarez A, Díez-Sánchez A, Del Pozo M A and Montoya M C 2018 Tox_(R)CNN: Deep learning-based nuclei profiling tool for drug toxicity screening. PLoS Comput. Biol. 14 e1006238
94. Nourani E and Reshadat V 2020 Association extraction from biomedical literature based on representation and transfer learning. J. Theor. Biol. 488 110112
95. He K, Zhang X, Ren S and Sun J 2016 Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 770
96. Huang G, Liu Z, Van Der Maaten L and Weinberger KQ 2017 Densely Connected Convolutional Networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition. 4700