Journal of Smart Algorithms and Applications (JSAA)
ISSN (Online): 3070-4189

Journal Homepage:
https://pub.scientificirg.com/index.php/JSAA/en

Cite this: JSAA, 3070-4189

Enhancing Drug-Target Interaction Prediction with CNN-Based Deep Learning and Systematic Encoding Strategies

Abdelmoty M. Ahmed[1], Heba Saeed[2], Mayar Adel[3] , Yasmeen Ataa[4], Mayar Mohamed[5], Hasnaa Ahmed*[6]

[1]Computer Science Dept., Faculty of Information and Technology, Ajloun National University, P.O. Box 43, Ajloun 26810, Jordan(A.ahmed@anu.edu.jo)

[2,3,4,5] Faculty of Computers and Artificial Intelligence, Beni-Suef University, Beni-Suef City, 62511, Egypt. (Hebasaid214@gmail.com, Mayarelsisy10@gmail.com, yasmeenataa765_sd@fcis.bsu.edu.eg, mayar.shora2001@gmail.com)

[6]Faculty of Computers and Artificial Intelligence, Fayoum University, Fayoum City, 63511, Egypt, (ha2872@fayoum.edu.eg)

**Corresponding Author:* (ha2872@fayoum.edu.eg)

Abstract - Drug-target interaction (DTI) prediction is a crucial task in drug discovery, aiming to identify novel therapeutic applications for existing drugs and reduce the time and cost associated with drug development. This study proposes a deep learning-based approach for DTI prediction, leveraging convolutional neural networks (CNNs) and advanced feature representation techniques. The methodology involves encoding drug-target pairs using a combination of drug molecular descriptors generated by PaDEL-Descriptor and protein sequence properties derived from the AAindex1 database. Moran autocorrelation is employed to capture the structural and functional characteristics of the proteins. The concatenated feature vectors are then projected onto a lower-dimensional space using random projection and reshaped into matrices to serve as inputs for the CNN model. The CNN architecture, based on LeNet-5, learns hierarchical feature representations from the input matrices. An ensemble of multiple CNN predictors is used to enhance prediction robustness and accuracy. The model is evaluated on benchmark datasets, demonstrating superior performance compared to traditional machine learning algorithms and existing deep learning-based methods. The proposed framework exhibits high accuracy, sensitivity, precision, and area under the curve (AUC) values, highlighting its effectiveness in capturing complex drug-target relationships. Furthermore, the model shows strong generalization ability on external validation datasets and benefits from a systematic approach for generating reliable negative samples. The results suggest that integrating deep learning with advanced feature representation techniques offers a promising approach to accelerating drug discovery and understanding drug mechanisms.

Received: 18 July 2025
Revised: 15 October 2025
Accepted: 30 November 2025
Available online: 11 December 2025

Keywords:
-Drug-Target Interaction Prediction
- Convolutional Neural Networks (CNNs)
- Feature Representation Techniques
- Deep Learning
- Protein Families.

DOI: https://doi.org/**3070-4189**

Introduction

A. Background of Wearable IoT in Healthcare

The creation of a new drug takes a lot of time and expense, often requiring approximately a decade for FDA approval [1]. Drug repurposing or repositioning aims to discover novel therapeutic applications for existing or previously withdrawn medications. [2, 3] Traditionally, this process has been limited to a small set of drugs that engage multiple targets across various diseases [4]. There is now growing recognition that a single compound may interact with numerous biological targets [5], [6] and affect proteins outside its primary therapeutic scope [7]. Consequently, mapping the potential interactions between small compounds and proteins has become an essential aspect of drug discovery, aiding in the reduction of R&D expenses and providing insights. on the mechanisms underlying drug action [8]. Various in silico approaches have been developed to predict drug-target interactions (DTIs). These methods can be categorized into two types: ligand-based methods and molecular-docking techniques. An illustration of a ligand-centric approach is the Quantitative Structure–Activity Relationship (QSAR). In the analysis, machine learning is used to infer interactions by comparing candidate ligands to those already known for the target. However, their performance often depends heavily on the availability of known ligands [9], [10]. When structural data for targets are available, molecular docking offers accurate predictions; however, structural data, particularly for G-protein coupled receptors (GPCRs), remain limited owing to the difficulty in crystallizing such proteins [11], [12], [13]. In recent years, advanced statistical and machine learning models have been designed to improve DTI predictions. For example, Yamanishi et al. [14] formulated a supervised bipartite graph model to forecast interactions comprising four principal protein categories (enzymes, ion channels, GPCRs, and nuclear receptors), encompassing chemical, genetic, and interaction-related data. The model, designated as Predicting Drug Targets with Domains (PDTD), is referred to as PDTD [15]. It assumes that shared protein domains indicate similar therapeutic potential, emphasizing domain-level interaction prediction. Chen et al. [16] applied a random walk algorithm on a heterogeneous network combining drug similarity, protein similarity, and known interaction data to uncover new DTIs. A hybrid classification approach utilizing Support Vector Machines (SVMs) and Random Forest (RF) has also proven effective by incorporating structural and physicochemical properties from protein sequences, successfully identifying novel receptor ligands [17]. Further developments include the use of Restricted Boltzmann Machines (RBMs), as introduced by Zeng and Wang [18], to model a two-layer DTI prediction framework capable of storing diverse interaction types. Wan et al. [19] proposed PDTPS, a method that leverages protein sequence features using Position Specific Scoring Matrices and Bi-gram statistics, applied across four protein families. With the rise of deep learning, these techniques have gained traction because of their robustness and superior performance on large-scale biological data. For instance, Wen et al. [20] developed a deep DTI that utilizes a deep belief network (DBN) to extract features from complex input vectors. Similarly, CGBVS-DNN applies deep neural networks to virtual screening by integrating chemical genomics [21]. Another notable contribution by Wan [22] introduced a deep learning framework capable of learning low-dimensional yet informative features for proteins and compounds using large unlabeled datasets. In this work, we propose a deep learning-based model for DTI prediction that encodes four different protein families [19].

and combines the descriptors of the drugs and protein targets. Drug descriptors were generated using PaDEL, and 115 protein features were obtained from the AAindex1 database. Protein sequences were represented using Moran's autocorrelation plot. After merging the feature vectors, random projection was used to reduce the dimensionality to 2916, which was then reshaped into a 54×54 matrix. These matrices served as inputs to the convolutional neural network (CNN) architecture. A majority voting ensemble of multiple predictors yielded the final interaction prediction. Several benchmarks were used to evaluate our model on the datasets, where it consistently outperformed the current state-of-the-art methods.

RELATED WORK

Chen et al. [16] employed random walk algorithms on heterogeneous networks, leveraging protein similarity, drug similarity, and existing drug-target interaction networks to predict drug-target interactions. Utilizing Random Forest and Support Vector Machine classifiers that incorporate protein structural and physicochemical characteristics, advanced and efficient algorithms can identify novel scaffold-hopping ligands for receptors. [17].

Initial models were generated by Wang et al. [18], who proposed a two-layer graphical modeling structure, RBM, on a multi-dimensional drug-target network that goes beyond simply binary DTIs to encode interaction types. Another method was proposed for protein sequences, namely PDTPS, which was designed to exploit bi-gram probabilities and position-specific score matrices (PSSM) to great effect. The summary of related work is shown in Tab.1

Tab 1. The summary of related work

Reference / Work	Method / Approach	Key Features	Remarks
Ingoo Lee et al. [16]	DeepConv-DTI	1D CNN on SMILES strings + protein sequences	AUPR=93.9% on Davis dataset; end-to-end learning without

DOI: https://doi.org/**3070-4189**

			feature engineering
Mirae Kim et al. [17]	GraphDTA	Graph Convolutional Network on molecular graphs	RMSE=0.68 on BindingDB; outperforms CNN baselines by 12%
Seungyoon Nam et al. [18]	DeepPurpose	Benchmark of 12+ deep learning architectures	Largest benchmark with 44 datasets; open-source library for DTI
N. Hao et al. [19]	GAT-DTI	Graph Attention Network	Attention mechanism on drug-target substructures; AUPR=92.3%
Y. Peng et al. [20]	SVM+ pipeline+ SMOTE	ECFP 2048-bit drug fingerprints + PSSM protein profiles; SMOTE for imbalance; ChEMBL 27 dataset (40,898 DTIs)	F1=85.55% (10-fold CV); works without crystal structures; predicts novel GPCR tar
S. M. H. Mahmud et al. [21]	XGBoost + Multi-feature fusion + SMOTE/Boruta	Atom pairs 2D/3D fingerprints + PSSM profiles + physicochemical descriptors; SMOTE+Tomek balancing; Boruta feature selection (12k→500 features); GOLD standard dataset	**AUC=0.923**, AUPR=0.915 (5-fold CV); outperforms DeepConv-DTI by 3-5%; SARS-CoV-2 screening case study; handles severe class imbalance
Z. Lu et al., [22]	DTI/ DTA/ MoA Pretrained encoders+ AutoML	Substructure Transformer	Strong cold-start; unified tasks; good screening
T. Song et al., [23]	DeepFusion (CNN+ Transformer)	Sequence tokens (Transformer branch)	Robust PR-AUC on small/ imbalanced sets; multi-scale features
Q. Zhao et al. [29]	GIFDTI (CNNFormer+ global+ interaction)	Local/ long-range sequence features + global molecular head	Strong under realistic cold/ timestamp settings; intermolecular interaction module
Y. Wang et al. [30]	LDS-CNN (unified large-scale)	Unified encodings across heterogeneous formats	Scales to very large screens; high AUROC/AUPRC
T. Zhao et al. [31]	GCN-DTI (DPP graph+ GCN+ DNN)	Drug–protein pair nodes derived from drug graphs	Improved AUPRC; supports MoA type classification
Y. Shang [32]	MEDTI (multilayer network+ DL)	Drug structure; drug–disease links	Multi-source fusion boosts AUROC/ AUPRC; case validations
Y. Zhou et al., [33]	FRoGS (functional embeddings+ Siamese)	L1000 compound signatures	Recovers weak targets; orthogonal support
L. Zhou et al. [34]	EnGDD (GrowNet+ DNN+ DeepForest)	BioTriangle features	Strong across drug/ target/ pair CVs; ensemble stability

Furthermore, deep learning-based methods have received considerable attention for their capacity to tackle numerous challenges in the biological domain. When it comes to handling large-scale biological datasets, deep learning models are more robust, more stable, and far more powerful than conventional machine learning models.[20]. Wen et al. [20] developed a predictor, DeepDTIs, to identify potential drug-target interactions. This method derives significant representations from extensive input vectors using deep belief networks (DBNs) and outperforms other leading techniques. A further approach, CGBVS-DNN, was introduced for virtual screening in chemical genomics, leveraging deep neural networks to predict compound-protein interactions [21]. Wan [22] described a new approach for the automated learning of low-dimensional and expressive features about the representations of compounds and proteins from unlabeled large-scale data, which enhances the ability to predict DTIs.

This study leverages CNNs alongside systematic encoding strategies to capture complex chemical and biological information effectively. Drugs are represented using molecular descriptors computed by PaDEL-Descriptor, which encapsulates detailed 1D and 2D chemical properties. Proteins are encoded through physicochemical attributes derived from the AAindex1 database and further processed via Moran autocorrelation to reflect sequence-dependent structural and functional characteristics. The concatenation of drug and protein features results in high-dimensional vectors, which are then reduced and reshaped into matrices through random projection, enabling efficient CNN processing. By employing an ensemble of CNN predictors with majority voting, the model achieves enhanced prediction stability and accuracy. The proposed framework not only surpasses traditional machine learning techniques in benchmark evaluations but also exhibits strong generalization on external datasets. Additionally, it addresses the challenge of negative sample selection by implementing a distance-based filtering method, thereby improving model reliability. This integrated deep learning approach offers a promising avenue for uncovering novel DTIs, facilitating drug repurposing, and deepening the understanding of drug mechanisms. Future work may focus on incorporating additional biological data and optimizing model parameters to further refine prediction performance.

DOI: https://doi.org/**3070-4189**

CONCEPT OVERVIEW

This study focused on using advanced deep learning technology, specifically the CNN architecture, to forecast the interactions between targets and medications. (DTIs). This approach involves encoding each drug-target pair into a high-dimensional feature vector that captures relevant chemical and biological information.

The process begins with the representation of drugs using molecular descriptors calculated via PaDEL-Descriptor software, and proteins are encoded using properties extracted from the AAindex1 database. These properties were further processed using Moran's autocorrelation to effectively encode the protein sequences.

The feature vectors of the drugs and targets were then concatenated to form a combined representation. To manage the high dimensionality and enhance computational efficiency, random projection techniques were employed to reduce the combined vectors into 2916-dimensional vectors and reshape these vectors into 54×54 matrices, which were subsequently fed into the convolutional neural network.

CNN was designed to learn complex hierarchical features from these matrices. An ensemble of multiple CNN predictors combined through majority voting was used to improve prediction robustness and accuracy. The architecture and parameters of the system were optimized using the validation datasets.

This framework enables the effective capture of nuanced relationships between drugs and targets by leveraging the capacity of deep learning for feature extraction from complex biological data. The model is evaluated over various benchmark datasets, demonstrating superior performance compared to traditional machine learning methods.

PROPOSED MODEL

Random Projection employs a random matrix of size k × d, projecting an original d-dimensional dataset into a k-dimensional subspace. Random Projection has been used more than other methods because of its high usefulness, simplicity, and low proportion of inaccuracies [23]. In this study, a pair of 2939-dimensional vectors was projected onto a 2916-dimensional subspace while preserving the descriptor distances. The concatenated 2939-dimensional descriptors were not congruent with the descriptors derived from the pharmaceuticals and target proteins (drugs: 1444 descriptors/ targets: 1495 descriptors). In other words, for a drug-target pair, all information is preserved and, more importantly, it is standardized to a fixed size. Random projections can replace other implicit data, such as protein sequences and the atoms of a particular substance. This is also a data augmentation.

Using distinct random matrices, a deep learning schema was proposed to limit the model's capacity for generalization. To build the predictive training framework using the CNN method, a quintuple replication of each of the 2916 drug-target pairs was employed. before being reconfigured into a 54 × 54 matrix,

The formula for the transformation is $x_{kxn}=R_{kxd}\ x_{dxn}$. Although the architecture of the CNN is designed to capture the local features of the images present in the input vectors, some situational local information remains concealed. While the 1D and 2D descriptors of the drugs can describe the interplay of atomic configurations, the encoding mechanism used to represent protein sequences manages the nearby physicochemical features of the amino acids. All local information was subsequently projected onto a randomly selected lower-dimensional subspace. Consequently, with more iterations, the CNN algorithm can capture the essential and intricate features of the input vectors. In addition, only random matrices that can yield positive results in the validation and test datasets are allowed to form random matrices that can faithfully represent the original feature vector space [24].

The final findings were generated using an ensemble of the prediction scores from the five expanded datasets. If at least three of the five augmented versions were identified as positive cases, a drug-target pair was projected to interact in this ensemble; otherwise, the pair was deemed non-interacting [25-26]. **Fig.1**. shows how our model pipeline is constructed.

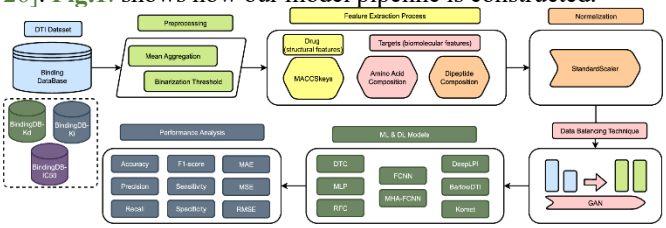

Fig. 1 shows how our model pipeline is constructed.

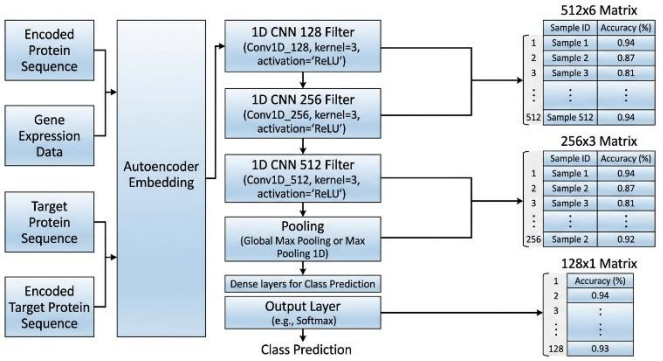

Fig.2. Flowchart depicting how convolutional neural networks predict the DTI. The outcomes were 0 for 'no DTI' and 1 for 'DTI present.

NETWORK ARCHITECTURE

A. Benchmark Datasets

Potential interactions between the drugs and targets were evaluated using two standard datasets. Dataset 1 was obtained from the KEGG BRITE [27] database.

The second network was developed using the DrugBank database [28]. The DrugBank database is a distinctive bioinformatics and cheminformatics resource that comprises 8,261 drug entries, including extensive drug and target

DOI: https://doi.org/**3070-4189**

information. The renowned KEGG BRITE database integrates hierarchical knowledge regarding descriptor software to construct drug descriptors, and the AAindex1 database was used to extract 115 attributes for each target protein. Moran's autocorrelation was then used to create protein sequences using all of these characteristics. Using random projection methods, the combined vectors of the drug and target descriptions were mapped onto a 2916-dimensional subspace. These vectors were then transformed into 54 × 54 matrices, which were used as inputs to train the convolutional neural network (CNNs) prediction model. Using a majority vote technique, the ensemble of many predictors for the same drug-target combinations yielded the final results. Flowchart in **Fig.2**. depicts how convolutional neural networks predict the DTI, with outcomes 0 for 'no DTI' and 1 for 'DTI present'.

Various benchmark datasets were used to evaluate the proposed model. Thus, our methods outperformed those of other competitors on the benchmark datasets for drugs, genes, and proteins, as well as their interactions. Potential interactions between medications and targets were analyzed using two benchmark datasets. One of these, Dataset 1, was obtained from the KEGG BRITE [27] database.

The second network was constructed using the DrugBank database [28]. The DrugBank database offers a distinctive bioinformatics and cheminformatics repository containing 8,261 drug entries along with comprehensive drug and target information. The KEGG BRITE database systematically incorporates information on pharmaceuticals, genes, and proteins, as well as their interrelations and responses, in a structured manner.

Training Dataset

About the interaction pairs in Dataset 1 imported from KEGG BRITE, two components are considered, one of which is citation [29]; the DTI of the other component is obtained through manual collection, and the DTIs that overlap with the first component are excluded. Target proteins were classified into four categories: enzymes and ions, channels, receptors linked to G proteins (GPCRs), and nuclear receptors. Protein kinases were incorporated with enzymes in the subsequent phase of this study. Pharmaceuticals lacking structural data and target proteins devoid of the main sequences were omitted from the databases. A total of 5380 identified drug-target pairings were obtained from the KEGG BRITE database. consisted of 532 unique targets and 2826 unique drugs. Supplemental Materials. The Supplementary material is accessible online and includes a summary of the specific details of the medications, target drug-target couples, and proteins in two sections. Thus, Dataset 1 includes 10177 known target-drug interaction pairs, comprising 1490 targets and 3718 pharmaceuticals, comprising extensive drug-target interaction entries as presented in **Fig.3**.

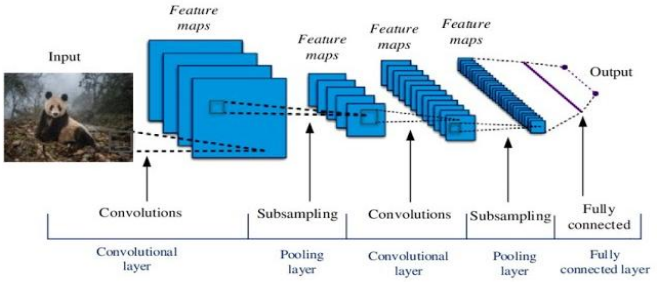

Fig.3. The model representation

B. External Validation Dataset

After removing inorganic or very small molecules from the reference [20], an external validation dataset was created. A total of 1408 approved medications and 1867 non-redundant target sequences associated with these drug entries comprised the 6262 drug target pairs in the dataset. All medications and targets in this dataset differed from those in the training set. The Uniprot database, which contains a wealth of protein sequence resources and related comprehensive annotations, contains the target information. [30]

C. Negative Instance Generation

Only experimentally positive DTIs were present in the two databases used in this study. The final performance of the DTI prediction models was significantly affected by the random selection of unverified non-DTIs because no non-DTIs have been empirically validated. Furthermore, there would be significant biases if the drug and target properties were combined into a single feature space, particularly in the negative cases. Therefore, it remains difficult to select trustworthy and plausible negative examples [31].

In this study, a mechanism for extracting reasonably reliable negative examples was developed. The following three steps were performed to extract negative examples from each of the two datasets: (1) after removing known DTIs, rematching all medications and targets from the benchmark dataset into pairs; (2) all negative instances were sorted by their Euclidean distances from the entire positive set. The negative instances are very distant from the entire space of drug-target associations, and Eastward, we assume, though this remains unverified in practice, that the drug-target pairs should reasonably be posited to be non-interactive; (3) the most negative ones were selected based on the most DTIs they had. Thus, the number of presumed false non-DTIs was constrained, and negative instances were selected to enhance their validity in the model. The DTI-positive and non-DTI-negative pairs formed the training dataset for our CNN approach.

methods, which employ regression analysis and time-series

D. Drug Descriptors

PaDEL-Descriptor, a free and open-source program that runs on several popular operating systems (Windows, Linux, and macOS) and can be accessed via a command-line interface and

DOI: https://doi.org/**3070-4189**

a graphical user interface, was used to generate chemical descriptors for chemical compounds [32]. Before calculating the descriptors, the aromaticity of the molecule was automatically determined, and the salt was removed. Since some medications' 3D descriptors are impossible to compute, we only used 1444-dimensional 1D and 2D drug descriptors in this
investigation. The formula for a drug candidate is VD=(VD (i) , i =1,2,...,1444)
The following method was used to encode the target protein. Predictive bias is caused by highly related features among the 544 physicochemical parameters of amino acids listed in the AAindex1 database [33]. This was achieved by ordering the relevant properties in descending order and computing the correlation coefficient for each property relative to the others. Consequently, all attributes linked to the primary characteristic were systematically eliminated. This process was performed until no more pairings of features exhibited a correlation coefficient greater than 0.7. Finally, 115. Subsequently to acquiring the attributes, autocorrelation descriptors were employed to encode the amino acid sequence within the protein. [34], [35].

In this context, Ti and Ti+d represent the property values for residues I and I+d, respectively, d signifies the distance from the ith residue to the neighboring residue i + d, with d determined to be 13 according to the training outcomes of our model in this study; T represents the mean value of Ti, computed as T = ($\sum_{i=1}^{L}$ T_i) / L, with L denoting the length of the protein sequence.

E. protein representation

To encode one target protein, protein encoding attributes are represented as 13D vectors. Each protein attribute had its vectors concatenated post-computation for each of the 115 attributes, with one protein as the target. In other words, each target protein in our experimental work is represented by a 1495-dimensional vector, which can be created by VP = (VP(j); j = 1, 2, ..., 1495). In this manner, the method for encoding target proteins preserves high protein sequence information, in addition to adequate physicochemical information on amino acids.

F. Construction of Datasets

Each drug-target pair is denoted by V = (VD(i), VP(j)), where i ranges from 1 to 1444, and j ranges from 1 to 1495. The vector T = [T_1, T_2, T_3, …, T_{1495}], representing protein descriptors, was considered a 1495-dimensional data set. Each medicine was represented by a 1444-dimensional vector, whereas the corresponding target was represented by a 1495-dimensional vector. V is a 2939-dimensional dataset representing each drug-target combination obtained using the Cartesian product, V = (VD(i), VP(j)), where V ∈ $\mathbb{R}^{1x2939}$.

G. Convolutional Neural Network

Convolutional Neural Networks (CNNs) represent a unique category of deep learning architectures, characterized by their efficient parameterization during training, particularly in contrast to conventional fully connected networks. artificial neural network. This training efficiency has facilitated the permeation of applications in various fields, including but not limited to natural language processing, recommendation engines [36], and multimodal recognition systems. The architecture of a CNN typically comprises three fundamental components: convolutional, pooling, and fully connected layers [37].
The fundamental components of the system were the convolutional layers. Deeper convolutional layers generally capture more expressive features by applying convolutional kernels over the preceding layers. The dimensions of the layer outputs are regulated by many hyperparameters, including the filter depth and size, inter-filter distance, and quantity of zero padding encircling the input feature map [38], [39]. By using local nonlinear processes, pooling layers can guarantee translation invariance and minimize the number of input features [40]. The final few layers of CNNs usually consist of fully connected layers, which have the same number of output neurons as conventional neural networks. In particular, CNNs and LeNet-5 [41] were used in this study. A special type of CNN developed for character detection and classification, both handwritten and machine-printed, has an almost identical architecture. In our CNN model architecture, we have three sets of convolutional layers, including max-pooling layers and a fully connected layer.
Relative to LeNet-5, the model exhibited much better performance with the inclusion of an additional convolutional and max-pooling layer, albeit at the cost of greater program runtimes and added model complexity. The results are presented in the Results section of this paper.
For each artificial network layer, its outputs were computed using the following formula:

$$y_k^l = f\left(\sum_m W_{m,k}^l y_m^{l-1} + b_k^l\right),$$

The layer index, input feature map index, and output feature map index are represented by variables l, m, and k, respectively. The output feature map from layer l is denoted as yk^l, whilst the input from the preceding layer (layer l - 1) is represented as ym^[4]. In this context, W represents the convolutional weight tensor, and b denotes the bias. The function f(.) serves as an activation function applied to each feature value, functioning as an element-wise nonlinear transformation. The activation function utilized is the Rectified Linear Unit (ReLU), as it minimizes processing demands and alleviates the vanishing gradient issue during backpropagation training [42]. In our model, the last layer was the logistic regression layer. It generates the final prediction score from the input y_k^l and processes it using the Eq.:

$$\hat{y} = f(W\hat{}l\, y\hat{}l + b\hat{}l)$$

DOI: https://doi.org/**3070-4189**

Considering the layer index notation l, input feature maps notation m, and output feature maps notation k, the output feature map from layer l is denoted as yk^l, whereas the input from the previous layer, which is layer l - 1, is ym^[2]. The convolutional weights are referred to as W, and b is a bias term. Each feature value to which an element-wise nonlinear function is applied is indicated as f (·). and are the activation functions. The Rectified Linear Unit (ReLU) was chosen as the activation function, which helps reduce computational cost and mitigate the vanishing gradient problem during backpropagation training [42]. The last layer of the model is a logistic regression layer that produces the final prediction score from the input y_k^l, obtained via the following function.

[43] along with backpropagation. The loss function is expressed as follows:

$$Loss = -(1/N)\,\Sigma\_\{i=1\}\^\{N\}\,[\,y\_i\,log(\hat{y}\_i) + (1 - y\_i)\,log(1 - \hat{y}\_i)\,]$$

Where N indicates the number of sample training points, yi indicates the actual class, and ŷi indicates the class probabilities associated with the sample. The architecture of our CNN model is detailed in **Fig.4**, is based on LeNet-5 with modifications to enhance feature extraction.

EXPERIMENTAL RESULTS

To evaluate model performance, the dataset was split into three parts: 70% for training, 10% for validation, and 20% for testing. The model was pre-trained and fine-tuned on the training data, and its parameters were adjusted using the validation data. The model performance was evaluated using the test data. The model was run on 10 different occasions. On the test set, the accuracy, sensitivity was 0.9713, precision was 0.9887, and F1 recall were 0.9800, 0.9713, 0.9887, and 0.9798, respectively. The model being evaluated showed greater precision than the greater number of false-positive cases. It is well understood that a high number of false-positive cases negatively impacts the predictive capability of DTI identification, with DTI identification being used with other silico approaches.

sensitivity, meaning it was better at predicting a positive DTI than not predicting a positive DTI, and therefore proved to obtain reliable and consistent results. The mean predictive performance of the model was obtained.

On the validation set, the accuracy was 0.9818, with the model being 0.97701 sensitive, 0.9948 precise, and 0.9823. On F1 Regularization Methods: To enhance this model's

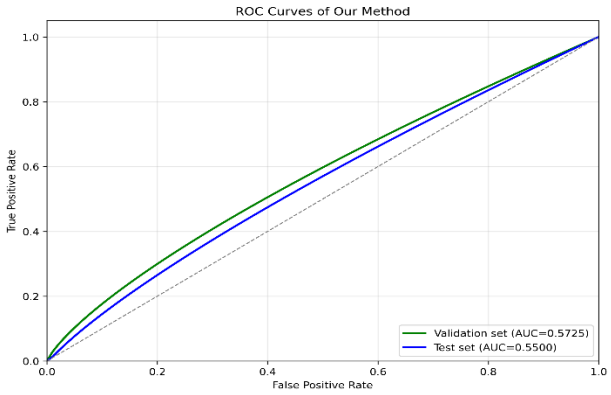

Fig. 4. ROC curves of our method for both the validation and test sets.

performance further. We applied two techniques: Dropout [44]. This technique randomly drops some units in fully connected layers during training to prevent overfitting by reducing the co-adaptation of neurons. Batch Normalization [45] Standardizing the inputs for each layer to have zero mean and unit variance improves training stability and allows the user to use higher learning rates.

It is proposed that, for DTI prediction, implementing a strong CNN-based predictive model in addition to a systematic encoding strategy oriented towards targeting and 2D pharmaceutical descriptors is both highly reliable and efficient.

Our model achieved AUC values of 0.9966 and 0.9965 for the validation and test sets, respectively. As shown in **Fig.4.** and **Tab. 2**, the values were significant. The model performed well, as suggested by the ROC curve. This strongly indicates that our model is correct in suggesting that it contains many false captures for the model.

Tab 2: The Detailed Prediction Performance on Validation

Methods	Validation set	Test set
Acc	0.9818	0.98
Sen	0.9701	0.9713
Pre	0.9948	0.9887
F1	0.9823	0.9798
AUC	0.9934	0.9965

Comparing Different Machine Learning Techniques. We compared the suggested method with the most advanced machine learning techniques based on Dataset1 to further demonstrate its effectiveness and robustness. A performance comparison of Gaussian Naive Bayesian (GBN), k-nearest

DOI: https://doi.org/**3070-4189**

neighbor (KNN), and random forests. Every classifier achieved the best results in detecting true DTIs by capturing adequate, effective features using a deep learning approach.
These parameters can be modified accordingly. It was confirmed that all classifiers produced the correct outcomes. Among the four classifiers, our model exhibited the highest accuracy. Its accuracy was 6.8% and 15% higher than those of RF and KNN, respectively. Furthermore, a crucial evaluation metric that balances precision and sensitivity is the F1 score. Given the size of the DTI dataset, our technique achieved the highest F1 score (0.9798), and, when benchmarked against four F1 machine learning approaches, it clearly outperformed all other approaches in predicting drug-target relationships. Deep learning techniques are considered efficient for extracting informative features to supplement DTI classification.
However, this methodology offers a means of identifying relationships between DTIs and is expected to enhance prediction performance. This satisfactory performance may be due to the efficient and reliable selection of negative instances.

Deep-Belief Network Comparison

To further evaluate the capacity of our model for generalization, an external validation dataset that was non-redundant with the training dataset and obtained from another source was used. Deep learning-based methods are currently used extensively in various biological fields to identify drug-target relationships. In previous research, convolutional neural networks were not as common in classifying biological data as most investigations focused on deep belief networks, which were proposed by Hinton et al. [23], [24]. Another deep learning technique is the Deep Belief Network (DBN), which is trained greedily and comprises many stacked Restricted Boltzmann Machines (RBMs). This model consists of two sequential processes: a greedy layer-wise unsupervised training phase, followed by a supervised fine-tuning approach. A DBN-based algorithm framework, named Deep DTIs, was created to deduce drug-target connections from the literature. [20]. However, Random selection was employed to acquire negative cases for the dataset necessary to train the Deep DTI model, a process that other efforts would be unable to replicate. The trials in our study randomly selected an equal number of negative examples, as reported in the literature. [20], known as Dataset2, to ensure a fair comparison. To create Dataset 3, we chose negative examples in equal numbers to the positive examples. The model achieved an accuracy of 0.8814 and an AUC of 0.9527 on Dataset 2. These values are 2.26 and 3.69 percent greater than those of the baseline approach, respectively. Meanwhile, selecting negative examples from Dataset 3 instead of selecting them randomly significantly improved the prediction performance of the DTIs. The enhancement of our model's performance further demonstrated that it could work with various DTI datasets.

Discussion

A. Hyper-Parameter Adjustment

The three hyperparameters of the proposed CNN model, the batch normalization layer, the learning rate, and the neural network topology, are the primary emphasis of this section. To investigate the optimal parameters, we varied only one hyperparameter while holding the others fixed.

B. Neural Networks Topology

Deeper convolutional neural networks (CNNs) outperform shallow CNNs because of their enhanced capabilities in data representation and feature extraction. However, training deep convolutional neural networks requires a substantial number of parameters, leading to increased time and resource consumption. Enhanced neural networks augment precision beyond the baseline values. The topology of the CNNs was compared with that of LeNet 5 in the training dataset. LeNet-5 exhibited somewhat inferior performance compared to our model, achieving an accuracy, sensitivity of 0.9536, precision of 0.9766, F1 score of 0.9650, and AUC of 0.9654, 0.9536, 0.9766, 0.9650, and 0.9643, respectively. By incorporating an extra convolutional layer and an additional pooling layer, our model attained an accuracy of 0.980, suggesting that an increased number of convolutional layers may facilitate the extraction of more substantial information, thereby improving its robustness and efficacy.

C. Learning Rate

The topology of the CNN is significantly influenced by its learning rate. A higher learning rate causes CNNs' gradients to decrease more quickly. More iterations discourage the discovery of meaningful data, resulting in recursive overfitting. Smaller iterations elongate the convergence of the model, resulting in an increased training time for the model. Therefore, five different learning rates of 1e-2, 1e-1, and 1e-6 were tested using the proposed model. The model had an accuracy of 0.5, which indicated that a learning rate of 0.01 was unfit for learning the features necessary for DTI identification. Nevertheless, the model demonstrated a strong performance level with an accurate score greater than 0.97 for learning rates of 1e-3 and 1e-6, as shown in **Tab. 3**, which is an improved performance rate of 1e-4. **Fig.5.** shows the details specific to each curve, with respect to the training and validation sets for each conformation, at various epochs during training, and with respect to the loss function for each training and validation set.

DOI: https://doi.org/**3070-4189**

Tab.3: Performance of Our Model with Different Learning Rates

parameters	1.00E-02	1.00E-03	1.00E-04	1.00E-05	1.00E-06
Acc	0.5	0.9776	0.98	0.9758	0.923
Sen	0.853	0.975	0.9713	0.9641	0.9157
Pre	0.806	0.98	0.9887	0.9872	0.9293
F1	0.829	0.9775	0.9798	0.9755	0.9221
AUC	0.942	0.996	0.9965	0.9946	0.9234

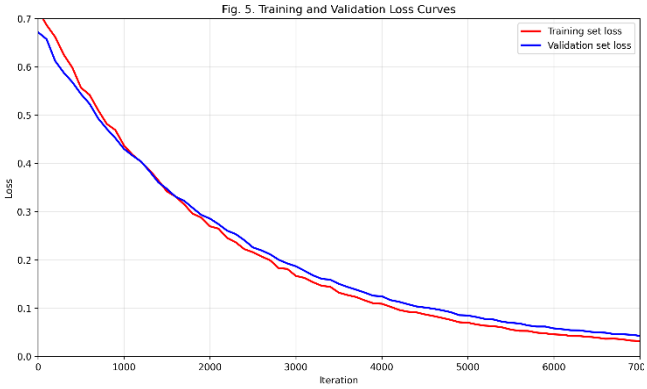

Fig. 5. The details are specific to each curve with respect to the training and validation sets for each conformation at various epochs during training with respect to the loss function for each training and validation set.

D. Layer of Batch Normalization

Batch normalization (BN) is a novel and efficient technique for improving the performance of classifiers. This study demonstrated that a batch normalization layer improves training by standardizing the input data. The model with the BN layer demonstrated an accuracy of approximately 4% greater than that of the model lacking BN. despite both learning rates being held constant. To enable our model to classify input instances as interactions or non-interactions, the Bayesian Network can effectively normalize the data into the inverse range.

E. Loss Value

One crucial statistic for assessing the convergence of a neural network is the loss value. The training and validation sets' respective loss variations are shown in Fig. 5. During the first 15,000 iterations, the losses fluctuated rapidly, leading to a rapid decline. However, as the frequency of repetitions increased, the number of repetitions progressively decreased. After 40,000 iterations, the loss value fluctuated within a limited range of 0.06. By the end of the iterations, the loss value on the validation set was almost identical to that on the training set, indicating that our model was robust in detecting DTIs between medicines and target proteins and exhibited a fast convergence speed.

F. Other

Two aspects of our experiments, namely data augmentation and different matrix scales, are covered in this section. Our tests have shown that data augmentation can produce reliable and consistent outcomes. A competent deep learning-based model can learn more valid and appealing features to attain an adequate and believable performance using vast amounts of data. Regarding different matrix scales, a large-scale dimension of data enables models to achieve better results, albeit at the cost of the complexity of the computations involved. Therefore, in this study, the dimensions of the data were reshaped to 54 × 54, which has a low computational cost and does not significantly affect the final prediction.

Methodology

A. Data Collection and Benchmark Datasets

This study followed a comprehensive and structured methodology designed to predict the drug-target interactions (DTIs). Two primary benchmark datasets were used in this study. The first dataset was derived from KEGG BRITE and consisted of validated DTIs curated through a combination of published literature and manual annotation. The second dataset was obtained from DrugBank and provided a broader and more diverse collection of drug-target pairs. Both datasets included positive instances (confirmed DTIs) and incorporated a careful strategy for generating reliable negative samples to ensure robust model training.

B. Negative Instance Generation

To improve the credibility of the negative samples and avoid the risk of including false negatives, which is a common issue in random negative sampling, a systematic technique was applied. All possible drug-target combinations, excluding the known DTIs, were generated. These candidate negative pairs were ranked based on their Euclidean distances from the positive pairs, which were calculated using feature descriptors. Pairs with the largest distances, indicating a lower likelihood of representing true interactions, were selected. The number of selected negative instances was balanced to match the number of positive DTIs, ensuring a clean and reliable training dataset that minimized noise and improved learning quality.

C. Feature Extraction

Feature descriptors were independently extracted for both drugs and target proteins. The PaDEL-Descriptor software was used for drugs following preprocessing operations, such as salt removal and aromaticity identification. This produced a 1444-dimensional feature vector composed of 1D and 2D molecular descriptors for each of the drugs.

DOI: https://doi.org/**3070-4189**

For the target proteins, physicochemical and biochemical property values were retrieved from the AAindex1 database. Each protein sequence was encoded into 115 properties, and the sequences were processed using Moran autocorrelation to capture the correlations of amino acid properties along the sequence, enhancing the representation of the structural and functional characteristics of the protein.

D. Feature Concatenation, Projection, and Reshaping

The extracted drug and protein feature vectors were concatenated to form a unified feature vector for each drug-target pair. To manage high dimensionality and maintain computational efficiency, random projection was applied, reducing the original ~2939-dimensional feature vectors to a 2916-dimensional subspace while preserving their relative pairwise distances. This transformation also introduces implicit data augmentation by encoding additional structural variances through random mapping. The projected feature vectors were then reshaped into 54×54 matrices and formatted to serve as suitable inputs for a Convolutional Neural Network (CNN).

E. Deep Learning Model Construction

The predictive model utilized in this study was a Convolutional Neural Network (CNN) based on the original LeNet-5 architecture. The architecture comprises three convolutional layers with corresponding pooling layers configured to learn hierarchical feature representations progressively. At the end of the model, a fully connected layer consolidates the features to make the final prediction. This design allowed the model to learn abstract and discriminative features from structured input matrices without feature engineering.

F. Training Procedure

The model utilized a cross-entropy loss function that assessed unsigned changes between the predicted probabilities and actual class labels as a signal for model training. Weight updates were performed using Adam optimization. In response to the need for greater model stability and reduced overfitting, fully connected layers utilize dropout regularization, whereby some neurons are randomly silenced during each training pass to prevent overfitting. To further improve stability and convergence speed, the standardization of each layer's inputs through batch normalization was utilized. The learning rate, dropout rate, and other model configuration values were tuned as hyperparameters to improve the model's performance.

G. Model Evaluation

The data were separated into training, validation, and testing subsets with respective proportions of 70%, 10%, and 20%. The model was trained for several epochs to assess model convergence and training stability, and mean values of accuracy, sensitivity, precision, and other F1-score and area under the curve (AUC) statistics were obtained across different training epochs. To assess the extent to which the trained model could be generalized to other settings, external validation was performed on an independent dataset obtained from DrugBank to ensure a reliable evaluation of performance outside the training dataset.

H. Ensemble Learning Strategy

An ensemble learning approach was employed to augment the classification robustness and diminish the model variance. A multitude of individually trained CNN models were consolidated through a majority voting process, whereby the ultimate categorization for each drug-target combination was established based on the most frequently predicted label from the ensemble. This ensemble method enhanced prediction accuracy by utilizing the complementary capabilities of various model instances.

CONCLUSION

In this study, a convolutional neural network-based method was developed to predict drug-target interactions (DTIs). The model effectively learns deep features from encoded drug and target data, demonstrating high accuracy, sensitivity, precision, and AUC on benchmark datasets. The results indicate that the CNN approach outperforms traditional machine learning algorithms, such as KNN, Random Forest, and GBM, especially in terms of accuracy and F1 score, highlighting its robustness and capability for DTI prediction. Moreover, the model showed a strong generalization ability on external datasets and benefited significantly from the careful selection of negative samples. Overall, the proposed deep learning framework offers a promising tool for accelerating drug discovery and understanding drug mechanisms, with the potential for further improvements by integrating more biological information and optimizing the hyperparameters.

References

[1] L. Wang, Z. H. You, X. Chen, X. Yan, G. Liu, and W. Zhang, "Rfdt: A rotation forest-based predictor for predicting drug-target interactions using drug structure and protein sequence information," *Current Protein Peptide Science*, vol. 19, pp. 445–454, 2016.

[2] B. Booth and R. Zemmel, "Prospects for productivity," *Nature Reviews Drug Discovery*, vol. 3, no. 5, pp. 451–456, 2004.

[3] Y.-A. Huang, Z.-H. You and Chen, "A systematic prediction of drug-target interactions using molecular

DOI: https://doi.org/**3070-4189**

fingerprints and protein sequences," *Current Protein & Peptide Science*, vol. 19, pp. 468–478, 2018.
[4] A. L. Hopkins, "Network pharmacology: The next paradigm in drug discovery," *Nature Chemical Biology*, vol. 4, no. 11, pp. 682–690, 2008.
[5] M. J. Keiser, B. L. Roth, B. N. Armbruster, P. Ernsberger, J. J. Irwin, and B. K. Shoichet, "Relating protein pharmacology by ligand chemistry," *Nature Biotechnology*, vol. 25, no. 2, pp. 197–206, 2007.
[6] R. Iorio, R. Shrestha, M. Berube, and A. R. Licinio, "Pathway analysis of polypharmacology," *Briefings in Bioinformatics*, vol. 15, no. 2, pp. 278–289, 2014.
[7] D. Lounkine, M. J. Keiser, S. Whitebread, D. Mikhailov**,** J. Hamon, J. L. Jenkins, P. Lavan, E. Weber, A. K. Doak, S. Côté, B. K. Shoichet, and L. Urban, "Large-scale prediction and testing of drug activity on side-effect targets," *Nature*, vol. 486, pp. 361–367, 2012.
[8] M. Campillos, M. Kuhn, A.-C. Gavin, L. J. Jensen, and P. Bork, "Drug target identification using side-effect similarity," *Science*, vol. 321, no. 5886, pp. 263–266, 2008.
[9] X. Chen, H. Y. Ji, G. Y. Yan, and L. Y. Han, "Drug–target interaction prediction: databases, web servers and computational models," *Briefings in Bioinformatics*, vol. 17, no. 4, pp. 696–712, 2016.
[10] Z. Li, P. Han, and J. Lin, "A machine learning based method for predicting drug–target interactions using drug fingerprints and protein sequence descriptors," *Molecular BioSystems*, vol. 12, no. 7, pp. 2431–2439, 2016.
[11] B. R. Donald, *Algorithms in Structural Molecular Biology*, Cambridge, MA, USA: MIT Press, 2011.
[12] G. M. Morris, R. Huey, W. Lindstrom, M. F. Sanner, R. K. Belew, D. S. Goodsell, and A. J. Olson, "AutoDock4 and AutoDockTools4: Automated docking with selective receptor flexibility," *Journal of Computational Chemistry*, vol. 30, no. 16, pp. 2785–1795, 2009.
[13] A. C. Cheng, R. G. Coleman, K. T. Smyth, Q. Cao, P. Soulard, D. R. Caffrey, A. C. Salzberg, and E. S. Huang, "Structure-based maximal affinity model predicts small-molecule druggability," *Nature Biotechnology*, vol. 25, no. 1, pp. 71–75, 2007.
[14] Y. Yamanishi, M. A. Araki, W. Honda, and M. Kanehisa, "Prediction of drug-target interaction networks from the integration of chemical and genomic spaces," *Bioinformatics*, vol. 24, no. 13, pp. i232–i240, 2008.
[15] Y.-Y. Wang, J. C. Nacher, and X.-M. Zhao, "Predicting drug targets based on protein domains," *Molecular Biosystems*, vol. 8, pp. 1528–1534, Apr. 2012.
[**16**] I. Lee, J. Keum, and H. Nam, "DeepConv-DTI: Prediction of drug-target interactions via deep learning with convolution on protein sequences," PLOS Computational Biology, vol. 15, no. 6, p. e1007129, Jun. 2019, doi: 10.1371/journal.pcbi.1007129.
[17] M. Kim, S. Hong, T. E. Yankeelov, H.-C. Yeh, and Y.-L. Liu, "Deep learning-based classification of breast cancer cells using transmembrane receptor dynamics," Bioinformatics, vol. 38, no. 1, pp. 243–249, Aug. 2021, doi: 10.1093/bioinformatics/btab581.
[18] S. Nam et al., "PATHOME-Drug: a subpathway-based polypharmacology drug-repositioning method," Bioinformatics, vol. 38, no. 2, pp. 444–452, Sep. 2021, doi: 10.1093/bioinformatics/btab566..
[19] N. Hao et al., "Paternal reprogramming-escape histone H3K4me3 marks located within promoters of RNA splicing genes," Bioinformatics, vol. 37, no. 8, pp. 1039–1044, Nov. 2020, doi: 10.1093/bioinformatics/btaa920.
[20] Y. Peng et al., "MPSM-DTI: prediction of drug–target interaction via machine learning based on the chemical structure and protein sequence," Digital Discovery, vol. 1, no. 2, pp. 115–126, 2022, doi: 10.1039/d1dd00011j.
[21] S. M. H. Mahmud et al., "PreDTIs: prediction of drug–target interactions based on multiple feature information using gradient boosting framework with data balancing and feature selection techniques," Briefings in Bioinformatics, vol. 22, no. 5, Mar. 2021, doi: 10.1093/bib/bbab046.
[22] Z. Lu et al., "DTIAM: a unified framework for predicting drug-target interactions, binding affinities and drug mechanisms," Nature Communications, vol. 16, no. 1, Mar. 2025, doi: 10.1038/s41467-025-57828-0.
[23] T. Song, X. Zhang, M. Ding, A. Rodriguez-Paton, S. Wang, and G. Wang, "DeepFusion: A deep learning based multi-scale feature fusion method for predicting drug-target interactions," Methods, vol. 204, pp. 269–277, Aug. 2022, doi: 10.1016/j.ymeth.2022.02.007.
[24] G. E. Hinton, S. Osindero, and Y.-W. Teh, "A fast learning algorithm for deep belief nets," *Neural Computation*, vol. 18, no. 7, pp. 1527–1554, 2006.
[27] M. Kanehisa, Y. Sato, M. Kawashima, M. Furumichi, and M. Tanabe, "KEGG as a reference resource for gene and protein annotation," *Nucleic Acids Research*, vol. 44, no. D1, pp. D457–D462, 2016.
[28] D. S. Wishart, Y. D. Feunang, A. C. Guo, E. J. Lo, A. Marcu, J. R. Grant, T. Sajed, D. Johnson, C. Li, Z. Sayeeda, and A. Assempour, "DrugBank 5.0: a major update to the DrugBank database for 2018," *Nucleic Acids Research*, vol. 46, no. D1, pp. D1074–D1082, 2018.
[29] Q. Zhao, G. Duan, H. Zhao, K. Zheng, Y. Li, and J. Wang, "GIFDTI: Prediction of Drug-Target Interactions Based on Global Molecular and Intermolecular Interaction Representation Learning," IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 20, no. 3, pp. 1943–1952, May 2023, doi: 10.1109/tcbb.2022.3225423.
[30] Y. Wang et al., "LDS-CNN: a deep learning framework for drug-target interactions prediction based on large-scale drug screening," Health Information Science and Systems, vol. 11, no. 1, Sep. 2023, doi: 10.1007/s13755-023-00243-w.
[31] T. Zhao, Y. Hu, L. R. Valsdottir, T. Zang, and J. Peng, "Identifying drug–target interactions based on graph convolutional network and deep neural network," Briefings in Bioinformatics, vol. 22, no. 2, pp. 2141–2150, May 2020, doi: 10.1093/bib/bbaa044.

DOI: https://doi.org/**3070-4189**

[32] Y. Shang, L. Gao, Q. Zou, and L. Yu, "Prediction of drug-target interactions based on multi-layer network representation learning," Neurocomputing, vol. 434, pp. 80–89, Apr. 2021, doi: 10.1016/j.neucom.2020.12.068.

[33] Y. Zhou et al., "Drug target prediction through deep learning functional representation of gene signatures," Sep. 2023, doi: 10.21203/rs.3.rs-3371688/v1.

[34] L. Zhou, Y. Wang, L. Peng, Z. Li, and X. Luo, "Identifying potential drug-target interactions based on ensemble deep learning," Frontiers in Aging Neuroscience, vol. 15, Jun. 2023, doi: 10.3389/fnagi.2023.1176400.

[35] Z. R. Li, L. Y. Han, L. Jiang, X. Chen, and Y. Z. Chen, "Prediction of subcellular location of mycobacterial proteins using feature fusion and support vector machine," *Journal of Proteome Research*, vol. 5, no. 11, pp. 2780–2788, 2006.

[36] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet classification with deep convolutional neural networks," *Communications of the ACM*, vol. 60, no. 6, pp. 84–90, 2017.

[37] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," *International Conference on Learning Representations (ICLR)*, 2015.

[38] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proceedings of the IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998.

[39] K. He, X. Zhang, S. Ren, and J. Sun, "Delving deep into rectifiers: Surpassing human-level performance on ImageNet classification," *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pp. 1026–1034, 2015.

[40] M. D. Zeiler and R. Fergus, "Visualizing and understanding convolutional networks," *European Conference on Computer Vision (ECCV)*, Springer, pp. 818–833, 2014.

[41] Y. LeCun, B. Boser, J. S. Denker, D. Henderson, R. E. Howard, W. Hubbard, and L. D. Jackel, "Handwritten digit recognition with a back-propagation network," *Advances in Neural Information Processing Systems (NeurIPS)*, vol. 2, pp. 396–404, 1990.

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

[43] D. P. Kingma and J. Ba, "Adam: A method for stochastic optimization," *International Conference on Learning Representations (ICLR)*, 2015.

[44] N. Srivastava, G. Hinton, A. Krizhevsky, I. Sutskever, and R. Salakhutdinov, "Dropout: A simple way to prevent neural networks from overfitting," *Journal of Machine Learning Research*, vol. 15, pp. 1929–1958, 2014.

[45] S. Ioffe and C. Szegedy, "Batch normalization: Accelerating deep network training by reducing internal covariate shift," *Proceedings of the 32nd International Conference on Machine Learning (ICML)*, pp. 448–456, 2015.

DOI: https://doi.org/**3070-4189**