Communications on Applied Nonlinear Analysis
ISSN: 1074-133X
Vol 32 No. 1s (2025)

A New Heuristic Hybrid Model for Cancer-Based P-Glycoprotein Drug Classification: Automated Prediction

Mylavarapu Kalyan Ram[1], S Kavitha [2]

[1] Research Scholar, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Vaddeswaram, Guntur, Andhra Pradesh, India.; Email: 193030002@kluniversity.in/kalyanram1985@gmail.com

[2] Associate Professor, Department of Computer Science & Engineering, Koneru Lakshmaiah Education Foundation, Greenfields, Vaddeswaram, Guntur, Andhra Pradesh, India; Emal: kavithabtech05@kluniversity.in

Article History:

Received: *10-08-2024*

Revised: *25-09-2024*

Accepted: *12-10-2024*

Abstract:
As the population improves in real-time applications, drug identification becomes an increasingly difficult task. Various biotechnological applications based on in vitro fertilization can benefit from better drug identification. For many nations, agriculture is the backbone of cancer medication development; increasing crop yields is the key to better drug identification in agricultural contexts. Computer vision applications suffer as a result of the efficient degradation of production quality and quantity caused by p-glycoprotein drugs used in cancer treatment. The identification of cancer-based p-glycoprotein drugs is an aggressive concept in the present day, due to the various symptoms present in cancer. Various authors use soft computing techniques and machine learning approaches to predict these drugs. For the purpose of automatically identifying based p-glycoprotein drugs in cancers, this paper proposes a Hybrid Novel Heuristic Model (HNHM) that combines Self-Organizing Maps (SOMs) with Convolution neural networks (CNNs). The primary challenge in identifying cancer drugs based on p-glycoprotein is feature extraction; hence, dimensionality reduction is crucial in filtering out noise from the cancer sub-cancer drug dataset and separating the subset of the cancer drug dataset affected by the p-glycoprotein drug using support vector machines (SOM). It is a promising approach to automatically identify based p-glycoprotein drugs in cancers by exploring protein features in cancer and predicting which patches are affected by virus-bacteria. CNN is used to match cancer sub-cancer drug data sets. In this study, we use sub-Cancer drug data sets to develop a hybrid heuristic approach to identify based p-glycoprotein drugs found in cancers. Clinical trials of the suggested hybrid model were conducted using sub-cancer drug data sets derived from peach cancers and the publicly available Cancer's drug datasets housed in the UCI repository. With less training required to explore based on p-glycoprotein drugs from sub-Cancer drug datasets, the suggested hybrid model achieves an accuracy of nearly 99.93% when tested on different testing datasets for sub cancer drugs. When compared to other state-of-the-art methods, this one significantly reduces running time while providing efficient performance in a number of p-glycoprotein drug prediction metrics, including recall, sensitivity, and specificity

Keywords: Prediction of p-glycoprotein drugs based on subsets of the population, feature extraction, histograms, convolutional neural networks, self-organizing maps, deep learning, and drugs derived from fruits.

1. Introduction

The enormous reduction in the physical or financial efficiency of the harvests and, in some cases, even a barrier to this action are both caused by cancer infections, which are considered one of the primary factors affecting cancer drug creation. According to [1], in order to keep production losses to a minimum and crop maintainability high, it is crucial to properly convey infection executives and control allocations, which should include consistent crop monitoring and the rapid and accurate diagnosis of illnesses. Psychopathologists recommend these methods most often [1].The correct identification of the symptoms of major diseases affecting crops is the major test in horticulture [2].Conventional establishing procedures, which rely on manual and motorised practices, are unable to cover vast areas of estate or provide dynamic cycles with basic early data [3]. Therefore, it is critical to develop automated systems that are practical, robust, and financially stable in order to monitor cancer patients' health and provide important data for dynamic interactions, such as the appropriate use and dosage of pesticides in the treatment of specific diseases [4]. Figure 1 displays various cancer subcancer drug data sets

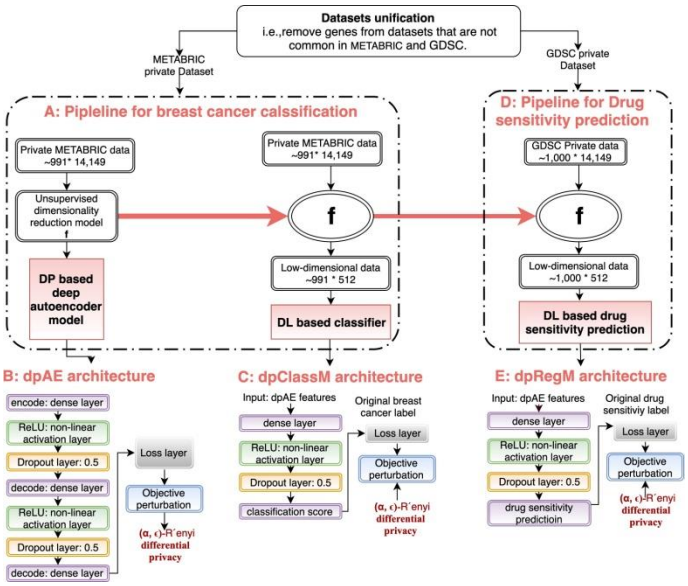

Figure 1 Cancer data sets with different drug parameters

Along with AI, computer vision has been making great strides in developing strategies and techniques for seeing and describing objects [5]. Using Convolution Neural Networks (CNNs) [6], these systems outperform humans in certain tests and have already been used for large-scale observational tasks. Using images is one of the most cutting-edge ways to detect and identify cancer disorders, as mentioned in the works of [7] and [8].A new approach to detecting content-based sub cancer medication data set information and understanding the data sets' recently debased and symbolically modified protests has been created (Rabbit et al., J. S. [4]). They also made use of a space vector approach to efficiently list any subset of cancer medication datasets. Afterwards, a two-venture grading method is used to choose the correct sub cancer medication data set. The computation is especially complicated when dealing with variations in queries and sub-drug databases inside the cancer drug dataset. For this method to run, Scale Invariant Feature Transform (SIFT) Descriptor13 is used. A vector representation of the images is generated from the images

using the Cancersubcancer drug data set's frequency distribution of "visual" terms. Data sets used in informational indexes include 850 cancer subtype drug datasets sourced from the National Cancer Subtype Drug Data Set assortments. The public historical center's photograph collections contained eight hundred fifty sets of cancer subtype drugs. The computer was validated using 200 randomly selected cancer subcancer drug data sets.

C. Szegedy and colleagues [7] I used the power of the DNNs in an effort to precisely separate the articles. In addition, they plotted out how to construct a double veil for the aerial vehicle. Relapse is possible for DNN. As an example, you show that the DNN relapse can pick up on order highlights and put together mathematical data. After several items in a set were covered with DNN relapse, a small group of large subwindows would be equipped with a DNN tracker. Using a single organisation, it makes educated guesses about four parts of the case in the item box veil. Using the Pascal Visual Object Challenge (VOC) 2007 dataset, the method was validated with 5,000 cancer subcancer drug data sets across 20 classes. Criteria for evaluation and testing in VOC2012 X. F. Hermida et al. [8] The suggested Braille text recognition system which makes use of Optical Character Recognition (OCR) made use of numerous methods for dealing with cancer subcancer drug datasets, such as flexible edge discovery and point identification. In high-contrast areas, a strategy has altered the areas with two flexible edges. We used the luminance histogram to find the limits. The computer also detected inflated scores and lost concentration.

Presently, numerous approaches make use of well-known designs such as LeNet[9], AlexNet[10], VGGNet[11], GoogleLeNet[12], InceptionV3[13], ResNet[14], and DenseNet[15], greatly improving the precision of cancer infection detection. The correct classification of physiopathology is obstructed by a number of factors, including the phenotypic and genetic diversity of yields, the abundance of pests and diseases, and so on. Included in this are the features and qualities of the data sets, the types of organisation models and designs, the complexity of the convolutional brain, and the methods used to streamline the results.

The engineering of neurons found in the human brain is the driving force behind Deep Learning techniques (Haykin, 1998). Artificial Neural Networks (ANNs) and its variants, such as Convolution Neural Networks (CNNs) and Recurrent Neural Networks (RNNs), are utilised in these approaches to identify hidden structures within data. Deep Learning methods have two obvious advantages over Machine Learning methods. The first is that they eliminate the need for an extra element extraction module since they naturally remove various highlights from crude data. As a second point, Deep Learning methods reduce the anticipated time required to manage massive, high-quality datasets. We then use Deep Learning procedures to build the hybrid model that you suggested. Two Deep Learning techniques that have shown success on data sets containing information about cancer subtypes and drugs are Convolution Neural Networks (CNNs) and Self-Organizing Maps (SOMs). These networks find widespread use in computer vision applications.Both of these methods extract distinct spatial and temporary highlights from the Cancersubcancer drug dataset by means of convolutional activity. SOMs effectively reduce the dimensionality of a Cancer subcancer drug dataset, while CNNs characterise the input datasets to their respective classes.

In light of the fact that there are a lot of training classes in the cancer subcancer drug dataset, this research proposes a novel heuristic hybrid model based on support vector machines (SOM) and

convolutional neural networks (CNN) to automatically identify cancer-based p-glycoprotein drugs. The model will be compared to various state-of-the-art methods. In order to process cancer subcancer drug data sets, SOM is employed for dimensionality reduction with noise removal, which in turn reduces the training samples. Convolutional neural networks (CNNs) are then employed to identify automatically based p-glycoprotein drug predictions from these datasets.

a) What follows is an explanation of the suggested method's contribution:
b) a) The suggested framework finds cancer drug-based p-glycoprotein drug classification using CNN and SOM, and it is based on data sets of cancer drugs.
c) b) In order to achieve reliable p-glycoprotein drug recognition, we employ feature extraction methods that are based on cancer subcancer drug data sets and p-glycoprotein drugs.
d) c) A high-resolution dataset on cancer subcancer drugs is used to assess the suggested framework.
e) d) In comparison to cutting-edge methods, the Implemented Approach provides higher-quality service..

2. Background Methods

In this section, we will go over the fundamentals of p-glycoprotein drug identification using sub cancer drug data sets related to cancer. When it comes to sub cancer drug data sets, SOM is great for reducing the dimensionality of certain sets, and CNN is great for predicting based p-glycoprotein drugs from these sets. Here you will find information regarding the SOM and CNN procedures.

i) Comprehensive Calculation of Self-Organizing Maps

In the field of artificial neural networks, Kohenen et al. presented an unsupervised learning calculation methodology called extensive self-organizing maps (ESOMs). When it comes to extracting patterns that are pertinent to based p-glycoprotein drugs, ESOM is an expressive and excellent tool. The m*n dimensional cancer subcancer drug data sets contain the relevant attributes that make up ESOM. For every attribute i in the subcancer drug data sets, there is a 2-dimensional weight factor with the values i=1.2...n.

The following are the definitions of extensive SOM:

The subcancer drug data set attributes are defined by a weight vector with n*m dimensions.

Subcancer medication data sets have a n*m dimensionality reduction characteristic that is defined as and the weight vector a(t) is likely to be identified according to the dimensions present in the Cancer subcancer drug data set.

$$w = \arg\left(\max_{1\le i\le n}\{\| w_i(t) - a(t) \|\}\right) \quad \text{..... (1)}$$

||.|| represents a euclidean distance measure; a(t) is the input weight vector with regard to attribute reading iterations, and w(t) is the output weight vector. Following the execution of the update, the weight attribute is defined as derived from the sub cancer drug data sets using equation (1).

$$w_i(t+1) = w_i(t) + h_{c,i}(t)[a(t) + w_i(t)] \quad \text{..... (2)}$$

$h_{c,i}(t)$ the color-based neighbourhood attribute identification using the distance between the weight vector in Equation (2), which is expressed as

$$h_{c,i}(t) = \alpha(t).\exp\left(-\frac{\| r_c - r_i \|^2}{2\sigma^2(t)}\right) \quad \ldots\ldots\ldots\ldots\ldots\ldots (3)$$

The dimensionality reduction from cancer-related subcancer medication datasets is achieved by using the weight function $\alpha(t)$ and $2\sigma^2$ from equation (3) and the associative attribute identification of map, denoted as r.

i) Convolution Neural Network Prediction

A Deep Learning technique known as a Convolution Neural Network (CNN) uses convolution operations instead of simple network augmentation. When it comes to managing images, CNN is the most effective Deep Learning method. It plays an important role in image characterization and other computer vision tasks by extracting various spatial and transient components from input images.

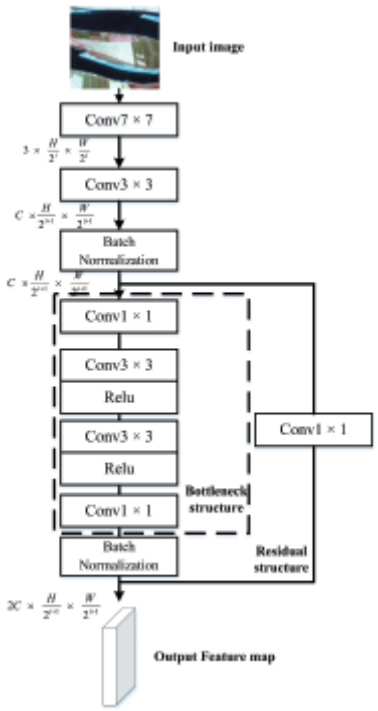

Figure 1. Read sub cancer drug datasets layer by layer using CNN's associative properties.

One way in which convolution matrix functions are defined is by

$$g(a) * f(a) = \int g(a).f(a-l) \quad \ldots\ldots\ldots\ldots\ldots (4)$$

$$g(a) * f(a) = \sum_{l=1,2,\ldots,\alpha}^{\alpha} g(a).f(a-l) \quad \ldots\ldots\ldots\ldots (5)$$

In Fig. 1, we can see the architecture of a typical convolutional neural network (CNN) with its many layers: input, output, convolution, pooling, and fully connected. Each layer in the network performs

an action. The CNN's convolution layer is responsible for performing the convolutional operations. Two types of convolutional neural networks (CNNs) are used in this setup: one at the base of the network separates sophisticated higher-level picture highlights, while the other at the top extracts simple lower-level picture highlights. The convolution activity is given by equations (4) and (5) and is a twofold activity between two real esteemed capacities, g(a) and f(a), in the cancer medication data set '∏'.

Both the information function (g(a)) and the channel-piece function (f(a)) from equations (4) and (5) are called "element maps" in convolutional neural networks (CNNs). Each component map, data set, and part-channel is stored in its own multi-faceted cluster. If the information framework is m × m and the channel is l × l (where l ≤ m), then the size of the resultant highlightmap, according to the definition of convolutional activity, is m−l+1×m−l+1. It is usually assumed that the size of the final highlight map is lowered after each convolution operation. To rephrase, after a particular amount of convolutions, the size of the info image becomes zero, as it reduces with each convolution operation. From this point on, it caps the amount of Convolution layers that a CNN can have, thus limiting the CNN's profundity. In addition, the components located on the edges and corners are used, rather than the exact components located in the centre of the input framework. The CNN's convolution layers use cushioning to deal with these two problems.

Integrate this strategy into your existing sub cancer drug data sets to automate the process of p-glycoprotein drug prediction.

ii) Multi CancerLabeled Generative Functional Approach

In order to assess the efficacy of a supervised learning procedure for analysing labelled cancer drug datasets, we will first look at the datasets as

$$\gamma = \sum_{(a,b)\in K_l} \zeta(a,b) + \sum_{a\in K_u} v(a) + \alpha \mathrm{E}_{(a,b)\in K_t}[-\log p_\phi(b \mid a)] \quad \text{…….. (6)}$$

K_t,K_u the distribution function for the multi-labeled Cancer drug dataset, α & are the points from the uncertain cancer drug dataset that range from single to multi-labeled, $p_\phi(b \mid a)$ and the labelled extra classification $\zeta(a,b)$ & $v(a)$ hyper-parameter is defined in equation (6). is defined $\zeta(a,b)$ in the following way:

$$\zeta(a,b) = \mathrm{H}_{CL}(p_\phi(z(a,b) \,||\, q(z)) - \log p_\phi(b) - \mathrm{E}_{p_\phi((z(a,b))}[\log q\theta(a \mid b,z) \quad \text{…….. (7)}$$

The first term in this context is the Leibler-Kullback divergence, which is defined as the function that describes the relationship between the posterior function (eq (7)) and the prior distribution function (q(z)), and the last term is the conditional expectation of the latent variable, which is similar to the hood function.

$$v(a) = \sum_b p_\phi(b \mid a)\zeta(a,b) - \mathrm{T}(p_\phi(b \mid a)) \quad \text{……. (8)}$$

Using the classifier entropy of input data (T) ($p_\phi(b|a)$) as a starting point, we may produce a new set of multi-labeled ($p_\phi(b|a)$) cancer medication data sets by applying generative functions to the input data set $p_\phi(b|a), p_\phi(z(a,b)\,\&\,q\theta(a|b,z)$ after removing noise (eq. (8)).

Systems that use recurrent neural networks

One of the many scientific challenges that RNNs have helped tackle with remarkable precision is the prediction of cancer prediction parameters. RNNs are time series generative functions. This state-of-the-art deep learning method is the best for learning different features in a sequential order. The prediction of p-glycoprotein medications based on cancer and the exchange of data sets containing cancer drugs are two of the many applications for RNNs. In Figure 1, we can see the fundamental architecture of the RNN.

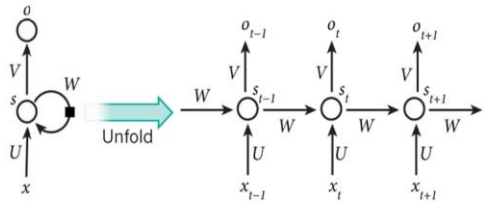

Figure 2: Recurrent neural network representation for p-glycoprotein drug prediction.

Figure 2 shows that the hidden state prediction follows each time step's output computation, with x, s, and 0 representing the input and hidden nodes, respectively. The weights of the shared matrices W, V, and U are used to perform inside-hidden processing on all time series, as well as input-to-hidden and output-to-hidden operations. Plat cancer drug data set data presented in real time processing uses three distinct models: a basic model, a long short-term memory (LSTM) model, and a GRU model. To classify multi-labeled data and reduce the number of tensor operations (update and reset), our strategy uses the GRU model instead of the LSTM model. Here are the characteristics and operations of tensor cells:

The GRU represents more cell equations, which include the following: update (zt), reset (rt), and hidden state (ht)

$$r_t = \sigma(W_r * [a_t, h_{t-1}] + y_r)$$
$$z_t = \sigma(W_r * [a_t, h_{t-1}] + y_z)$$
$$h_t = \tanh(r_t * [a_t, h_{t-1}] + y_h))$$
$$h_t = z_t * h_t + (1 - z_t) * h_{t-1} \quad \ldots\ldots\ldots\ldots\ldots\ldots (9)$$

According to what was said earlier, the input and output dependencies are distributed across different time stamps. Our implementation followed this procedure to measure multi label inputs and provides multi label outputs using data from various Cancer drug datasets.

3. Proposed Methodology

The procedure for creating an NHHM, or Novel Heuristic Hybrid Model, which combines SOM and RNN, will be covered in this section. The similarity matrix that maps the cancer drug to protein

sequences is formed using SOM. Here are the modules that make up the Implemented Approach: a) Defining the dataset b) Mapping the SOM c) Protein interactions d) Interactions between proteins and drugs the forecasting of drugs

Informational summary:

We take our cue from NPInter v2.055, a database that records confirmed interactions between ncRNAs and a wide range of biomolecules, such as genomic DNAs, proteins, and RNAs. In this database, we discovered the essential ncRNA-protein interactions. Based on our search through NONCODE, we were able to identify human G-glycoproteins that are known to interact with other G-glycoproteins.56; 57 We were able to compile 141,353 G-glycoprotein sequences according to NONCODE 4.0. A dataset of 4,158 G-glycoprotein-protein interactions with a high degree of certainty was obtained after 990 G-glycoproteins and 27 proteins were screened out, along with low-quality G-glycoprotein sequences that did not belong to humans.number of 58,59 When we used reliable datasets, our method performed better after we removed G-glycoproteins that were only linked to one protein and proteins that were only linked to one G-glycoprotein.

Mapping of Similarity Matrixes using SOM:

For every pair of G-glycoproteins, we used it to find out how similar they were. A pair of G-glycoproteins, denoted as LSM (li; lj), represent the similarity score between them. To determine the weights for each group in LSM (li; lj), we used the following formula.

$$LSM\left(l_i, l_j\right) = \frac{sw\left(l_i, l_j\right)}{\max\left(sw(l_i, l_i), sw\left(l_j, l_j\right)\right)},$$

where sw(li; lj) is the score based on SOM that compares the sequences of two G-glycoproteins, li and lj

Data set of cancer drugs derived from various cancer cells and p-glycoprotein sequences provided as input.

Results: Cancer cell protein-based medication target identification

S1: Assess dimensionality reduction using the aforementioned method
S2: Use relevant protein features to assess the explored sub-Cancerdrug data set's features.

$$R = \left\{h \mid \left(\widehat{a}(h) > \widehat{a}(h-1)\right) \& \left(\widehat{a}(h) > \widehat{a}(h+1)\right)\right\}$$

This is where the Cancer-based p-glycoprotein drug prediction region, denoted as R, exists.

Using the classified region's attribute locations, determine the threshold for all cancer subcancer drug datasets (S3)

$$K = \left\{h \mid \left(\widehat{a}(h) > \widehat{a}(h-1)\right) \& \& \left(\widehat{a}(h) > \widehat{a}(h+1)\right)\right\}$$

Based on the height and weight of Cancer subcancer drug dataset

S4 let K denote all attribute thresholds: Eliminate all of the background

properties.

$\widehat{a}(h+1) \neq \widehat{a}(h-1)$

S5: Check each foreground and background attributes values

If (h is high) &

Where R=R-h;

If (h is low) & $\widehat{a}(h+1) \neq \widehat{a}(h-1)$ then K=K-;

S6: Using the interaction values, we assessed and represented all attributes in the sub cancer medication dataset.

$\{[1,i_1],[2,i_2],[3,i_3],...,[n,i_n]\};$

S7: Cancer-based p-glycoprotein drug prediction using sub cancer drug data

data sets $P_1, P_2, P_3, ..., P_{|n|}$

Algorithm 1 Procedure to predict Cancerbased p-glycoprotein drug from sub-Cancersubcancer drug data sets.

Proteins Interact

To calculate protein-protein similarity scores, we followed the same procedure as for P-glycoproteins and used the SOM based calculation method. The entity PSM (pi; pj) represents the degree of similarity between proteins pi and pj, and it is used to denote the protein similarity matrix (PSM).The PSM (pi; pj) formula is as follows:

$$PSM(p_i, p_j) = \frac{sw(p_i, p_j)}{\max(sw(p_i, p_i), sw(p_j, p_j))}.$$

It is shown as sw(pi; pj) how similar the sequences of proteins pi and pj are according to the SOM algorithm.

Relationships between Protein-Based Drugs

We used the sequence similarity matrices to calculate the interaction scores between G-glycoproteins and proteins. It was thus believed that the adjacency matrix Y described the G-glycoprotein-protein relationships, with the entity Y (li; pj) being 1 if G-glycoprotein li is proven to be associated to the protein pj and 0 otherwise.

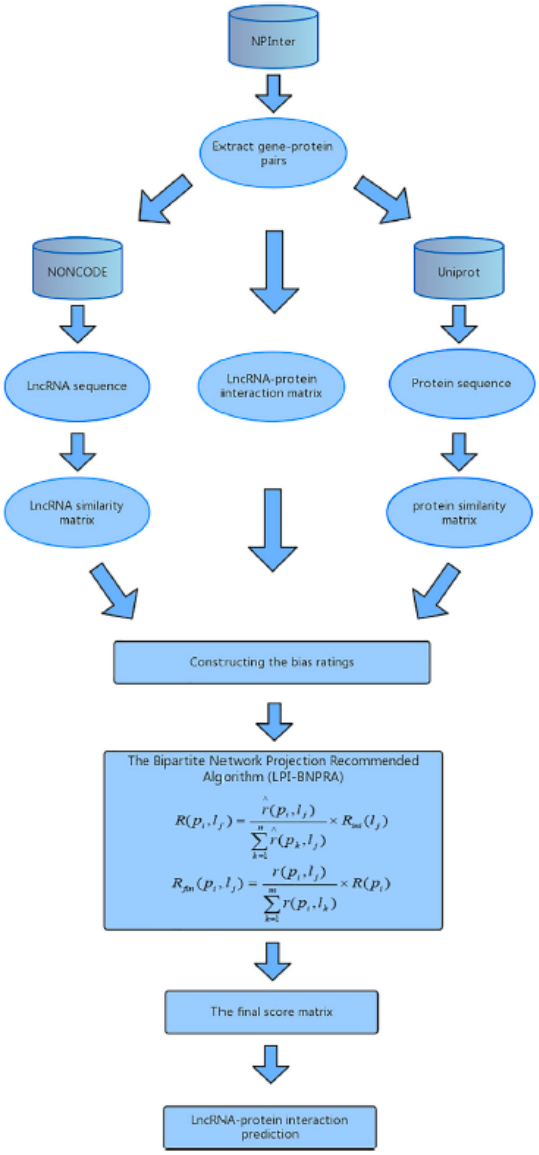

Figure 3. Process of Implemented Approach

Cancer Prognosis A P-glycoprotein Drug Derived from Interactions with Proteins

Figure 3 displays the G-glycoprotein interaction matrix, protein similarity matrix, and G-glycoprotein similarity matrix using LPI-BNPRA, which are used to compute G-glycoprotein interaction scores and establish potential G-glycoprotein-protein links.

Using agglomerative hierarchical clustering, we were able to obtain the bias ratings of all G-glycoproteins. For example, many proteins with identical sequence information find a particular G-glycoproteinli to be relevant. So, we can use this information to create bias ratings for the proteins that contain these G-glycoproteinli.

Agglomerative hierarchical clustering is a bottom-up approach; it starts by treating proteins (or G-glycoproteins) as independent clusters, and then uses the linkage criterion of minimum variance method to merge these individual clusters. The distance between two G-glycoproteins, denoted as LD(li; lj), and the distance between two proteins, denoted as PD(pi; pj), are both used in this context.

$$LD(l_i, l_j) = 1 - LSM(l_i, l_j)$$

$$PD(p_i, p_j) = 1 - PSM(p_i, p_j).$$

There are many reliable groups at the correct threshold when the hierarchical clustering tree is pruned. Repeatedly using LOOCV to conduct the same experiments yields the same prediction results regardless of the threshold. This led us to the following computation of protein pi's bias rating in comparison to P-glycoprotein lj::

$$r(p_i, l_j) = \frac{n_{cr}}{T(p_i)},$$

Here, ncr is the number of clusters that contain P-glycoprotein lj, and Tpi is the total number of P-glycoprotein that are somehow related to protein pi.The inclusion of protein-to-G-glycoprotein bias scores allowed us to complete the development of the bias-rating score matrix. So, we can propose the P-glycoproteins that correspond to a given protein by sorting the final resource ratings from highest to lowest.

4. Experimental Evaluation

Based on the system settings, we use a state-of-the-art deep learning framework to execute our suggested method. We first divide the original data into subsets called "cancer drug data sets," each of which has 256*256 attributes and semantic dimensions. We compare our suggested method with generalized learning abilities and use top-down and left-to-right flipping to randomly select transformation gradient factors. Sub cancer medication data set for cancer Storage facility: Experiments were conducted using the repository of the sub cancer drug data set to assess the performance of the proposed method. China, India, and Province are just a few of the countries that contributed to the regional cancer sub cancer drug data sets that were collected between 2018 and 2021. Data Every one of the 1500–2000 cancer sub drug data sets in the repository has an attribute region of 4950*4950. These sets are used for training and testing purposes. In contrast to the Arial sub drug data sets displayed in Figure 4, all the sub cancer drug data sets were collected using the sub cancer drug data sets repository.

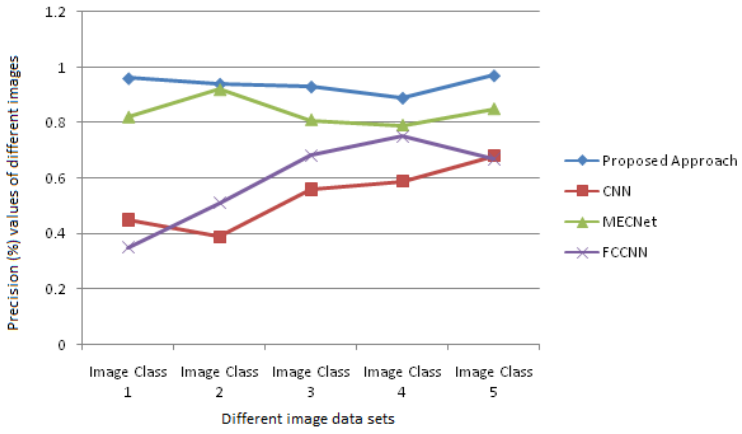

Figure 4. Accuracy testing using various sorts of cancer medication databases

Results of a p-glycoprotein medication study in Forecasting: We assess forecasting techniques and compare them with our own. We evaluate our new heuristic hybrid model (NHHM) against state-of-the-art approaches using real-time sub cancer medication data sets. These approaches include CNNs, FCCNNs, and others, and we measure their accuracy and other visual metrics. You can see the different precision values utilized to evaluate these datasets for prediction in Table 1, which provides several cancer sub cancer medication data sets**.** Databases of cancer drugs.

Table 1 Different Precision values.

Precision				
Cancer-Protein drug data set datasets	Implemented Approach	CNN	MECNet	FCCNN
Cancer-Protein drug data set	0.96	0.45	0.82	0.35
Cancer-Protein drug data set	0.94	0.39	0.92	0.51
Cancer-Protein drug data set	0.93	0.56	0.81	0.68
Cancer-Protien drug data set	0.89	0.59	0.79	0.75
Cancer-protein drug data set	0.97	0.68	0.85	0.67

The accuracy of the proposed method is compared to that of more traditional methods in Figure 4; we tested it on a variety of cancer subcancer drug datasets and found that it performed better overall. In contrast to other methods, the suggested one achieves a precision of nearly 97–99% when dealing with increasingly diverse classes of cancer subcancer drug data sets; this indicates that the layout contains p-glycoprotein drugs exactly as intended..

Table 2 Different recall values.

Recall				
Cancer-Protein drug data sets	Implemented Approach	CNN	MECNet	FCCNN
Cancer-Protein drug data set	0.96	0.52	0.81	0.51
Cancer-Protein drug data set	0.97	0.46	0.86	0.46
Cancer-Protein drug data set	0.98	0.61	0.86	0.56
Cancer-Protein drug data set	0.92	0.65	0.74	0.54
Cancer-Protein drug data set	0.99	0.59	0.96	0.46

Table 2 and Figure 5 compare the recall performance of traditional approaches with that of the proposed method utilizing different cancer drug datasets (subsets of cancer). The suggested method yields 99% accuracy from average sub cancer drug data sets whenever there is an improvement to the class of cancer drug data sets

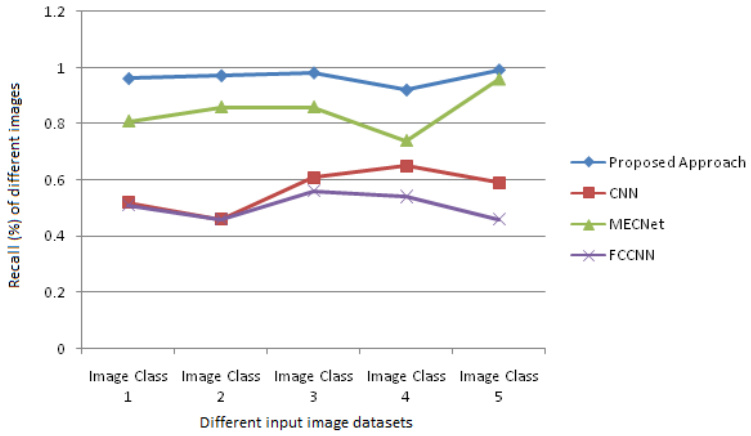

Figure 5: Evaluates the recall performance using various cancer subcancer drug datasets

While competing approaches provide inconsistent results when it comes to retrieving useful information from various satellite cancer sub cancer drug databases. The values associated with the f-measure as they pertain to the classes of cancer sub cancer drug data sets are detailed in Table 3

Table 3 F-measure values

	F-measure			
Cancer drug data set datasets	Implemented Approach	CNN	MECNet	FCCNN
Cancer-protein drug data set	0.96	0.72	0.91	0.84
Cancer-Protein drug data set	0.97	0.64	0.96	0.76
Cancer-Protein drug data set	0.97	0.81	0.89	0.79
Cancer-Protein drug data set	0.98	0.86	0.95	0.81
Cancer-Prtoein drug data set	0.99	0.79	0.91	0.78

Whenever we provide a number of cancer subcancer drug data sets, NHHM outperforms CNN, FCCNN, and MECNet, as demonstrated in figure 6, which compares the performance of the suggested approach with that of traditional approaches

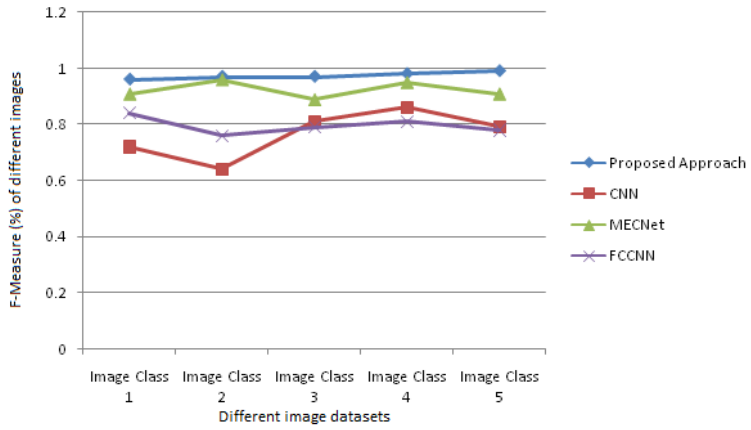

Figure 6. Application of the f-measure to the analysis of various cancer sub cancer drug databases
In comparison to state-of-the-art methods, NHHM processes matched content at a high rate—nearly 100%—when cancer subcancer drug data sets are increased. The proposed method, and all methods,

experience an automatic increase in the f-measure as the precision and recall parameters are raised. Results from subcancer drug datasets pertaining to the time required for the overall execution of the process and the prediction of a cancer-based p-glycoprotein drug are displayed in Table 4.

Table 4 Time efficiency values for different approaches.

	Time Efficiency			
Cancer drug data set datasets	Implemented Approach	CNN	MECNet	FCCNN
Cancer-Protein drug data set	5.1	9.2	5.1	7.2
Cancer-Protein drug data set	4.6	8.6	5.6	6.2
Cancer-Protein drug data set	5.6	9.1	6.4	6.8
Cancer-Protein drug data set	5.4	7.5	7.4	5.4
Cancer-protein drug data set	4.6	9.6	6.5	6.7

In Figure 7, the entire time length of the suggested strategy is compared to more conventional ones.

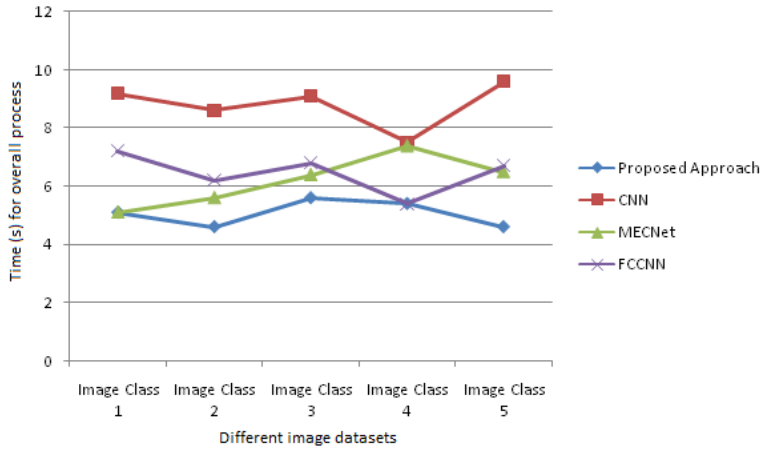

Figure 7: Time efficiency in processing data sets for cancer subcancer drugs.

The suggested method uses a fraction of the time required by CNN and FCCNN. Because of the iterations used to assess the water body prediction from the Cancer Subcancer drug datasets, these methods are very time-consuming, in contrast to MECNet, which takes about the same amount of time to process all of the datasets.

Values associated with subbased p-glycoprotein drug prediction from Cancer subcancer subcancer drug data sets are shown in Table 5. It explains the effective precision values in comparison to more conventional methods..

Table 5 Final prediction accuracy values.

	Accuracy			
Cancerdrug data set datasets	Implemented Approach	CNN	MECNet	FCCNN
Cancer-protein drug data set	84	71	68	69
Cancer-protein drug data set	96	61	81	62
Cancer-protein drug data set	92	76	76	75

Cancer-protein drug data set	89	60	84	63
Cancer-protein drug data set	93	67	88	71

The Implemented Approach outperforms conventional CNN, FCCNN, and MECNET with an overall accuracy close to 99%. MECNET performs nearly as well with the Implemented Approach as with NHHM, while the other approaches provide less accuracy..

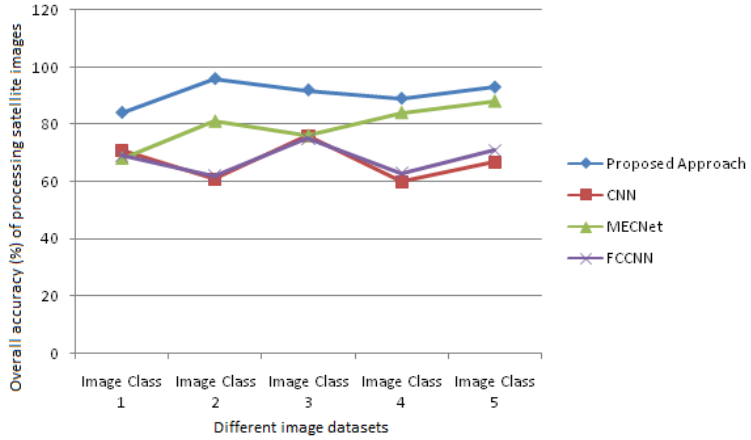

Figure 8 Accuracy of processing different data sets.

Using data sets on cancer drugs, Figure 8 shows the overall accuracy performance.

In comparison to more conventional methods for predicting based p-glycoprotein drugs from sub cancer drug datasets, the suggested method outperforms them in terms of efficiency and effectiveness (see Figure 8 for details).

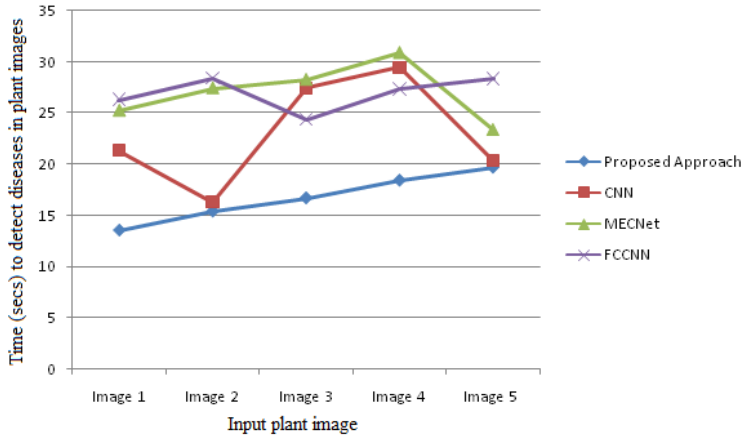

Figure 9. Performance evaluation of time with respect to Cancerbased p-glycoprotein drugprediction.

Figure 9 shows the overall performance of time about the prediction of cancer-based p-glycoprotein drugs, and it depicts how the performance of the implemented approach is better and more efficient than traditional ways as the number of cancer-subcancer drug data sets increases. When it comes to predicting cancer-based p-glycoprotein medications from cancer subcancer drug datasets, CNN and FCCN are nearly on par.

Table 6. Cancerbased p-glycoprotein drugprediction time values

	Time (Secs) CancerBased p-glycoprotein drugPrediction			
Cancer drug data sets	Implemented Approach	CNN	MECNet	FCCNN
Cancer-protein drug data set	14.56	22.42	26.34	27.42
Cancer-protein drug data set	16.52	17.41	28.45	29.51
Cancer-protein drug data set	17.82	28.49	29.38	25.48
Cancer-protein drug data set	19.55	30.54	31.82	38.15
Cancer-protein drug data set	19.72	20.33	23.35	28.35

The time values of the sub cancer Drug Data Set in relation to the based p-glycoprotein drug prediction of various Cancer Sub Drug Data Sets are displayed in Table 6. Table 8 displays the results of the cancer sub drug data sets' predictions for based p-glycoprotein drugs in relation to the enhancement of these databases

	Cancerbased p-glycoprotein drugprediction ratio			
Cancer drug data sets	Implemented Approach	CNN	MECNet	FCCNN
Cancer-protein drug data set	99.75	89.42	87.24	80.14
Cancer-protein drug data set	93.71	87.52	88.45	82.45
Cancer-protein drug data set	97.24	82.14	84.26	85.64
Cancer-protein drug data set	96.52	84.32	87.45	79.23
Cancer-protein drug data set	98.74	88.12	88.24	72.56

Table 7. Cancerbased p-glycoprotein drugprediction ratio values.

As the amount of cancer-related drug data sets grows, the prediction ratio improves in comparison to more conventional methods

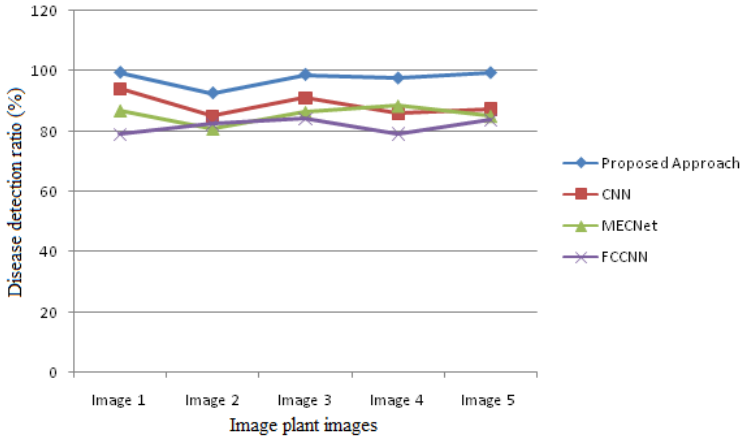

Figure 10. Performance of prediction of with respect to sub cancer drug data sets.

Figure 10 shows the results of the sub cancer drug data sets' performance in relation to the improvement of based p-glycoprotein drug prediction for cancer. In comparison to more conventional methods, the suggested method produces a more accurate and efficient prediction ratio when dealing with larger cancer sub cancer drug datasets. When it comes to the dimensionality of the

cancer drug dataset, CNN and MECNET provide a much better prediction ratio. Prediction failures occur across the board due to the high dimensionality of protein space and the difficulty of layer extraction from various cancer drug data sets that contain patched information. In conclusion, when compared to competing methods, our suggested method produces superior and more efficient results

5. Conclusions

Scientific computer vision applications face a formidable obstacle in the prediction of based p-glycoprotein drugs for cancer. Researchers in the field of computer vision have developed a variety of deep learning and machine learning techniques for the purpose of identifying p-glycoprotein drugs in subcancer drug databases. In order to automatically identify based p-glycoprotein drugs in cancers, we present a Novel Heuristic Hybrid Model that combines Self-Organizing Maps (SOMs) with Convolution neural networks (CNNs). When it comes to cancer subcancer drug data sets, feature extraction is the main challenge. Dimensionality reduction, on the other hand, helps filter out noise and separates the patches that are affected by based p-glycoprotein drugs in SOM. The purpose of this CNN study is to investigate cancer-related protein features and to identify the cancer-related patches that are associated with p-glycoprotein drugs. Biotechnological applications of the suggested method include the identification of p-glycoprotein drugs that originate in viruses and bacteria. When compared to more conventional methods, the suggested approach achieves a prediction accuracy of 99.98% for cancer-based p-glycoprotein drugs using subcancer drug datasets. Predicting the subconsistency of p-glycoprotein drugs in cancer subcancer drug datasets using relevant visual characteristics is an area where our research is constantly evolving and improving

References

[1] Chen, Z.;Wu, R.; Lin, Y.; Li, C.; Chen, S.; Yuan, Z.; Chen, S.; Zou, X. CancerBased p-glycoprotein drug Recognition ModelBased on Improved YOLOv5. Agronomy 2022, 12, 365. https:--doi.org-10.3390-agronomy12020365

[2] Joshi, R.C.; Kaushik, M.; Dutta, M.K.; Srivastava, A.; Choudhary, N. VirSubNet: Automatic analysis and viral based p-glycoprotein drug diagnosisusing deep-learning in Vigna mungoCancer. Ecol. Inform. 2021, 61, 101197. [CrossRef]

[3] Buja, I.; Sabella, E.; Monteduro, A.G.; Chiriacò, M.S.; De Bellis, L.; Luvisi, A.; Maruccio, G. Advances in CancerBased p-glycoprotein drugPredictionand Monitoring: From Traditional Assays to In-Field Diagnostics. Sensors 2021, 21, 2129. [CrossRef] [PubMed]

[4] Liu, S.; Liu, D.; Srivastava, G.; Połap, D.;Wo´zniak, M. Overview and methods of correlation filter algorithms in object tracking.Complex Intell. Syst. 2020, 7, 1895–1917. [CrossRef]

[5] Tang, Y.; Chen, M.; Wang, C.; Luo, L.; Li, J.; Lian, G.; Zou, X. Recognition and localization methods for vision-based fruit pickingrobots: A review. Front. Cancer Sci. 2020, 11, 510. [CrossRef]

[6] Li, J.; Tang, Y.; Zou, X.; Lin, G.;Wang, H. Prediction of fruit-bearing branches and localization of litchi clusters for vision-basedharvesting robots. IEEE Access 2020, 8, 117746–117758. [CrossRef]

[7] Wu, F.; Duan, J.; Chen, S.; Ye, Y.; Ai, P.; Yang, Z. Multi-Target Recognition of Bananas and Automatic Positioning for theInflorescence Axis Cutting Point. Front. Cancer Sci. 2021, 12, 705021. [CrossRef]

[8] Wang, C.; Tang, Y.; Zou, X.; Luo, L.; Chen, X. Recognition and matching of clustered mature litchi fruits using binocularcharge-coupled device (CCD) protein cameras. Sensors 2017, 17, 2564. [CrossRef]

[9] Lin, G.; Tang, Y.; Zou, X.; Xiong, J.; Fang, Y. Protein-, depth-, and shape-based 3D fruit prediction. Precis. Agric. 2020, 21, 1–17.[CrossRef]

[10] Luo, L.; Liu, W.; Lu, Q.; Wang, J.; Wen, W.; Yan, D.; Tang, Y. Grape Berry Prediction and Size Measurement Based on Edge Cancer drug data setProcessing and Geometric morphology. Machines 2021, 9, 233. [CrossRef]

[11] Luo, L.; Chang, Q.; Wang, Q.; Huang, Y. Identification and Severity Monitoring of Maize Dwarf Mosaic Virus Infection Based onHyperspectral Measurements. Remote Sens. 2021, 13, 4560. [CrossRef]
[12] Appeltans, S.; Pieters, J.G.; Mouazen, A.M. Prediction of leek white tip based p-glycoprotein drug under field conditions using hyperspectral proximalsensing and supervised machine learning. Comput. Electron. Agric. 2021, 190, 106453. [CrossRef]
[13] Fazari, A.; Pellicer-Valero, O.J.; Gómez-Sanchıs, J.; Bernardi, B.; Cubero, S.; Benalia, S.; Zimbalatti, G.; Blasco, J. Application ofdeep convolutional neural networks for the prediction of anthracnose in olives using VIS-NIR hyperspectral cancer drug data sets. Comput.Electron. Agric. 2021, 187, 106252. [CrossRef]
[14] Julian Stöttinger, Allan Hanbury, NicuSebe, Senior Member, IEEE, and Theo Gevers, Member, IEEE, "Sparse Protein Interest Points for Cancer drug data setRetrieval and Object Categorization", IEEE TRANSACTIONS ON CANCER DRUG DATA SET PROCESSING, VOL. 21, NO. 5, MAY 2012.
[15] Punam Bedi, Pushkar Gole , "Cancerbased p-glycoprotein drugprediction using hybrid model based on convolutionalautoencoder and convolutional neural network", Artificial Intelligence in Agriculture 5 (2021) 90–101
[16] Vikas Chaudhary a,*, R.S. Bhatia a, Anil K. Ahlawat,"A novel Self-Organizing Map (SOM) learningalgorithm with nearest and farthest neurons", Alexandria Engineering Journal (2014) 53, 827–831.
[17] Mostafa, A.M.; Kumar, S.A.; Meraj, T.; Rauf, H.T.; Alnuaim, A.A.; Alkhayyal, M.A. Guava Based p-glycoprotein drugPrediction Using Deep Convolutional Neural Networks: A Case Study of Cancer drugs. Appl. Sci. 2022, 12, 239.https:--doi.org-10.3390-app12010239.
[18] Ahmed, K., Shahidi, T.R., IrfanulAlam, S.M., Momen, S., 2019. Rice subbased p-glycoprotein drugpredictionusing machine learning techniques. 2019 International Conference on SustainableTechnologies for Industry 4.0 (STI). IEEE, Dhaka, Bangladesh, pp. 1–5.
[19] Ashraf, T., Khan, Y.N., 2020. Weed density classification in rice crop usingcomputer vision. Comput. Electron. Agric. 175, 105590. https:--doi.org-10.1016-j.compag.2020.105590.
[20] Bisong, E., 2019. Autoencoders. Building Machine Learning and Deep Learning Models onGoogle Cloud Platform: A Comprehensive Guide for Beginners. Apress, Berkeley, CA,pp. 475–482. https:--doi.org-10.1007-978-1-4842-4470-8_37.
[20] Zhang, W., Qu, Q., Zhang, Y., and Wang, W. (2018). The linear neighborhood propagation method for predicting long non-coding RNA–protein interactions.Neurocomputing 273, 526–534.
[21] Chen, X., and Yan, G.Y. (2013). Novel human lncRNA-disease association inferencebased on lncRNA expression profiles. Bioinformatics 29, 2617–2624.
[22] Chen, X., Yan, C.C., Zhang, X., Zhang, X., Dai, F., Yin, J., and Zhang, Y. (2016). Drugtarget interaction prediction: databases, web servers and computational models. Brief.Bioinform. 17, 696–712.
[23] Kumar, T. Rajasanthosh, et al. "Implementation of Intelligent CPS for Integrating the Industry and Manufacturing Process." AI-Driven IoT Systems for Industry 4.0. CRC Press, 2024. 273-288.
[24] Wang, E., Zou, J., Zaman, N., Beitel, L.K., Trifiro, M., and Paliouras, M. (2013).Cancer systems biology in the genome sequencing era: part 1, dissecting andmodeling of tumor clones and their networks. Semin. Cancer Biol. 23, 279–285.
[25] Wang, E., Zou, J., Zaman, N., Beitel, L.K., Trifiro, M., and Paliouras, M. (2013).Cancer systems biology in the genome sequencing era: part 2, evolutionary dynamicsof tumor clonal networks and drug resistance. Semin. Cancer Biol. 23, 286–292.
[26] Wang, E., Zaman, N., Mcgee, S., Milanese, J.S., Masoudi-Nejad, A., and O'Connor-McCourt, M. (2015). Predictive genomics: a cancer hallmark network frameworkfor predicting tumor clinical phenotypes using genome sequencing data. Semin.Cancer Biol. 30, 4–12.
[27] Chen, X., Xie, D.,Wang, L.,Zhao, Q.,You, Z.H., and Liu, H. (2018).BNPMDA: BipartiteNetwork Projection for MiRNA-Disease Association prediction. Bioinformatics 34,3178–3186.
[28] Bellucci, M., Agostini, F., Masin, M., and Tartaglia, G.G. (2011). Predicting proteinassociations with long noncoding RNAs. Nat. Methods 8, 444–445.
[29] Muppirala, U.K., Honavar, V.G., and Dobbs, D. (2011). Predicting RNA-protein interactions using only sequence information. BMC Bioinformatics 12, 489.

[30] Laxmaiah, G., et al. "Experimental and Simulation analysis of Monitoring a Industrial process by Adaptive transfer Learning." 2023 International Conference on New Frontiers in Communication, Automation, Management and Security (ICCAMS). Vol. 1. IEEE, 2023.
[31] Hearst, M.A. (1998). Support Vector Machines. IEEE Intell. Syst. 13, 18–28.
[32] Wang, Y., Chen, X., Liu, Z.-P., Huang, Q., Wang, Y., Xu, D., Zhang, X.-S., Chen, R.,and Chen, L. (2013). De novo prediction of RNA-protein interactions from sequenceinformation. Mol Biosyst. 9, 133–42.
[33] Lu, Q., Ren, S., Lu, M., Zhang, Y., Zhu, D., Zhang, X., and Li, T. (2013).Computational prediction of associations between long non-coding RNAs and proteins.BMC Genomics 14, 651.
[34] Suresh, V., Liu, L., Adjeroh, D., and Zhou, X. (2015). RPI-Pred: predicting ncRNAprotein interaction using sequence and structural information. Nucleic Acids Res.43, 1370–1379.
[35] Ge, M., Li, A., and Wang, M. (2016). A Bipartite Network-based Method forPrediction of Long Non-coding RNA-protein Interactions. Genomics ProteomicsBioinformatics 14, 62–71.
[36] Hu, H., Zhu, C., Ai, H., Zhang, L., Zhao, J., Zhao, Q., and Liu, H. (2017). LPI-ETSLP: lncRNA-protein interaction prediction using eigenvalue transformation-based semisupervised link prediction. Mol. Biosyst. 13, 1781–1787.
[37] Liu, H., Ren, G., Hu, H., Zhang, L., Ai, H., Zhang, W., and Zhao, Q. (2017). LPINRLMF: lncRNA-protein interaction prediction by neighborhood regularized logisticmatrix factorization. Oncotarget 8, 103975–103984.
[38] Hao, Y., Wu, W., Li, H., Yuan, J., Luo, J., Zhao, Y., and Chen, R. (2016). NPInter v3.0:an upgraded database of noncoding RNA-associated interactions. Database (Oxford)2016, baw057.