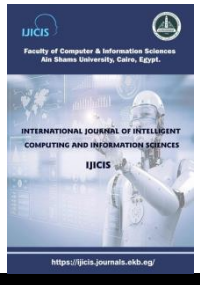
IJICIS, Vol.23, No.1, 98-105
DOI: 10.21608/ijicis.2023.180508.1239

International Journal of Intelligent Computing and Information Sciences

https://ijicis.journals.ekb.eg/



CONVOLUTIONAL NEURAL NETWORK MODELS FOR CANCER TREATMENT RESPONSE PREDICTION

Hanan Ahmed*
Scientific Computing, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
hanan.ahmed20100@cis.asu.edu.eg

Howida A. Shedeed
Scientific Computing, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
dr_howida@cis.asu.edu.eg

Safwat Hamad
Scientific Computing, Faculty of Computer and Information Sciences, Ain Shams University, Cairo, Egypt
shamad@cis.asu.edu.eg

Ashraf S. Hussein
Scientific Computing, Faculty of Computer and Information Sciences, Ain Shams University, *Cairo, Egypt*
ashussein@cis.asu.edu.eg

Received 2022-12-12; Revised 2023-02-06; Accepted 2023-02-09

Abstract: *Recently, efforts are exerted on cancer treatment prediction based on the biomarkers related to the tumor. Gene expression and mutation profiles are the most used biomarkers for cancer prediction. Machine learning and deep learning algorithms have been used to predict drug response. The recent research show that the performance of deep learning models is better than the performance of machine learning based one. In this paper, Convolutional Neural Network (CNN) models use are introduced to predict different drugs response. DeepInsight algorithm used to convert the input data to images to be more suitable as input to the CNN. Three different pretrained CNNs-models (InceptionV3, Xception, EfficientNetB7) are introduced with alternatives in their settings of the training process and modification in their architectures to be able to predict the drug response using IC_{50} regression values. Those models are selected due to their efficiency for ImageNet applications.the proposed modified Xception model achieves the best accuracy over the 2 others. At first, the whole data input passes through DeepInsight which converts the gene expression data and mutation data to images. Dimension reduction is then applied using the helper technique inside the DeepIsignt. Comparative analysis with other Deep models, shows that the proposed approach improve the prediction accuracy in a range between 14% and 22% as a reduction in mean squared error (MSE).*

Keywords: *Artificial Intelligence, Artificial neural networks, Biomedical, Convolutional neural networks, Personalized Medicine, Drug Response Prediction.*

1. Introduction

According to [1] cancer is the second cause of death worldwide. The expectation says that by 2040 the number of cancer cases will be 29.5 million and the number of cancer deaths will be 16.4 million [2]. So huge efforts are exerted to find an effective cancer treatment.

***Corresponding Author**: Hanan Ahmed

Scientific Computing Department, Faculty of Computer and Information Science, Ain Shams University, Cairo, Egypt

Email address: hanan.ahmed20100@cis.asu.edu.eg

Cancer is caused by abnormal cell growth. There are more than 100 types of cancer most of them called by their organs or tissues. Most of the cancer types form some solid tumors but some others do not as leukemia. Cancer has a huge variation in its characteristics, scientists found that the abnormal cell growth is related to mutations in genes. The same tumor may contain thousands of gene mutations some are the causes of cancer (drivers) and others are the result of cancer development (passengers). With the huge development in genetics research, they find that some drugs give good results for patients and some others do not due to the difference in the genetic profiles mutations related to the cases [3], [4]. That makes the prediction task of drug response based on genetic profile and its mutation. Studying the genetic profiles and deciding the best treatment is a very time-consuming process and requires a huge human effort. According to [5] the use of Artificial Intelligent (traditional machine learning models) in cancer field begins in 2010 to predict the cancer development, likelihood of redevelopment after recovery, life expectancy and the survivability. In this stage only non-genetic features are use as medical records and history. The start of genetic features was in 2015, the target of prediction is prognosis, recurrence and finding the cancer type. The use of genetic features is considered the foundation of individual treatment.

2. Related Work

As stated in [5] machine learning (ML) models are used starting from 2010 to predict the probability of developing, redeveloping and, survivability of cancer. Since 2017, both ML and deep learning (DL) algorithms have been used for predicting cancer treatment responses.

Different techniques are used for predicting the treatment response. Deep Neural networks (DNN) are used in [6] [7] [8]. As shown in [7], a comparison is made among DNN, Random Forest, and Elastic Net models using different datasets. The DNN model achieves better performance than other models on different datasets. In [8], a DL model is proposed (consDeepSignaling) that uses gene expression and copy number variation for genes as features. The CNN-models are used in drug response prediction such as in [9] [10]. Cancer Drug Response profile scan (CDRscan) is introduced in [9]. CDRscan consists of five different CNN-based models; each model predicts the drug response for the same cell line. The final result is the average of the five CNN-models. CDRscan uses data collects from four different datasets and predicts the response of 244 different drugs. A Twin Convolutional Neural Network (tCNN) is proposed in [10]. tCNN uses two CNNs; one for extracting the features of drugs from their simplified molecular input line entry specification (SMILES) format. The SMILES is converted to a one-hot matrix format to be suitable for CNN. The other one extracts the features from the cell lines. The results of the previous CNNs are input to a fully connected Network which does the regression task between the out of previous CNNs and the response values. In [11] DNN is introduced for continuous drug response prediction called PM. PM has a feature selection phase before it is using Recursive feature elimination (RFE) and a discretization phase after it.

In this paper, the use of Convolutional Neural Network (CNN) models is proposed to predict the drug response of 265 different drugs. The CNN predicts based on both the gene expression profile and the mutation profile. Both gene expression data and mutation data are converted to 2D images using a recent algorithm called “DeepInsight”. DeepInsight is an algorithm to convert non-image data to images to suit the use of CNN-models [12]. The used datasets are Cancer Genome Atlas (TCGA), Cancer Cell Line Encyclopedia (CCLE), and Genomics of Drug Sensitivity in Cancer (GDSC). The rest of the paper is

organized as follows: Section 3 introduces our proposed work and its results in Section 4, The comparative study in section 5. Finally, the conclusion is introduced in section 5.

3. Proposed Methodology

The proposed models are implemented to predict the half-maximal inhibitory concentration (IC_{50}) values. IC_{50} is a measure of the quantity is needed from a drug to inhibit a biological process by 50%. The response data in GDSC is stored in form of IC_{50}. The proposed models predict IC_{50} based on the genomic profiles of a cell or a tumor. The methodology is divided into the following steps:

1) The first step is the preprocessing done on the input datasets to enhance the data quality.
2) The second step is converting the genetic input data (gene expression and mutations) to a 2D image using DeepInsight algorithm
3) The third step is the using of Tsne feature reduction technique inside the DeepInsight.
4) The fourth and last step is the prediction network which predicts the IC_{50} values of the target 265 drugs.

3.1 Datasets

In the proposed work, the used datasets are Cancer Genome Atlas (TCGA), Cancer Cell Line Encyclopedia (CCLE), and Genomics of Drug Sensitivity in Cancer (GDSC). CCLE project, which targets to accurately identify the cancer cell lines characteristics. CCLE consists of 25 oncogenes mutation across 486 cancer cell lines, DNA copy number variations of 1,043 cancer cell lines with 23,316 genes, and mRNA expressions of 127 cancer cell lines [13] with 54,675 mRNAs. In 2019, the CCLE database received a major update. The update is adding new data [14] that can be summarized as follow: 1) DNA methylation data. 2) Whole genome sequencing data. 3)RNA-seq data. The TCGA dataset is comprised of 2.5 petabytes of data including the complete genetic profiles of tumors and their normal tissues from more than 11,000 patients which representing 33 cancer types [15]. The GDSC [16] project includes drug responses of 265 drugs in means of IC_{50} values for 1,001 cancer-cell lines. The number of transcripts per million (tpm) for genes are used from CTD^2 Portal [17]. For TCGA, the Pan-Cancer dataset was used and tpm values were downloaded from UCSC Xena [18] [19] for 10536 TCGA samples. The data federation is used as in [6] to integrate the data and improves its quality. The used samples from both CCLE and GDSC are 619 samples while TCGA is 9158 samples. The number of genes for both CCLE and TCGA is 19702 genes.

3.2 DeepInsight Algorithm

As it is explained, the second step in the proposed methodology is the use of the DeepInsignt algorithm to convert input data to 2D images to suit the CNN models.

DeepInsignt [12] appears in 2019. It aims overcome the CNN limitation to be used in different applications Simply DeepInsignt is a method to convert dataset to images to be suit for the CNN models..The algorithm can be summarized as follows:

1) Converting samples from row-wise to column-

2) For each feature apply Normalization so the values of the features range within [0,1]
3) If the number of features is too large, so it will be hard to deal with. So, dimension reduction technique or feature selection can be applied.
4) Constructing similarity matrix which measures the similarities among features.
5) Mapping features to Cartesian coordinates see Figure. 1
6) The convex hull algorithm applied to find the smallest rectangle contains all points. See Figure 2.
7) A rotation is performed to the previous rectangle to make it aligned horizontally or vertically.
8) The feature values are mapped according to locations determined in the previous step.

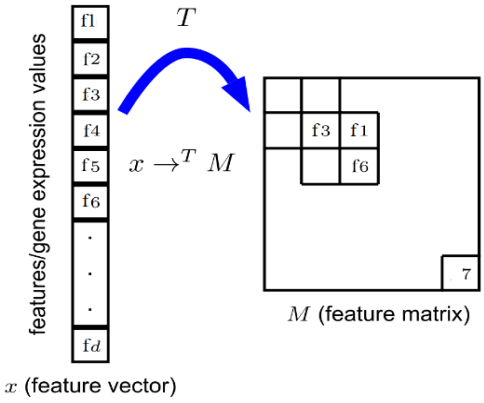


Figure 1. Mapping feature to Cartesian coordinates

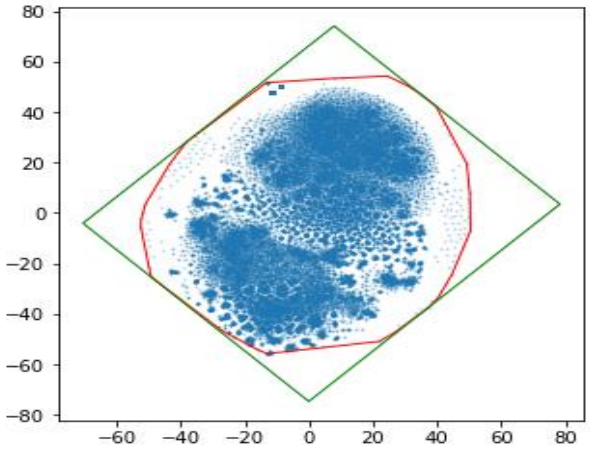


Figure 2. Applying convex hull algorithm

3.3 T-SNE Dimension reduction algorithm

DeepInsight, introduced in the previous section, has a helper for dimension reduction and similarity measurement. It uses one of the following three algorithms: T-distributed Stochastic Neighbor Embedding (t-SNE) [26], kernel principal component analysis (kPCA), or principal component analysis (PCA). The t-SNE is used here in the proposed model.

t-SNE is a nonlinear dimensionality technique that can be used with very high dimensional data. It not only used as dimension reduction technique is also used for visualizing high dimensional data. The main idea is converting data points to joint probabilities and the use those probabilities to represent points in lower dimensions using Kullback-Leibler divergence.

3.4 Convolutional Neural Network (CNN)

Convolutional Neural Network (CNN) is neural network model being used for image classification problem, see Figure 3. The big idea behind CNNs is reduces the images into a form that is easier to process, without losing features. CNN mainly consists of 3-layer types: a convolutional layer, a pooling layer, and a fully connected layer.

The Convolution Layer is the main layer and one with the highest computational load.

The pooling Layer replaces the output of the network at certain locations. The pooling operation is performed on every slice individually. The pooling functions that are commonly used are the average of the rectangular neighborhood, the L2 norm, etc. The fully connected layer is the layer in which all neurons are fully connected with all neurons in the preceding and succeeding layers. It helps to map the representation between the input and the output.

CNN has more than 20 [20] different models which are different in their architectures. In this paper, the use of 3 different CNN-models to predict the response of 265 anti-cancer drugs is introduced. The models are modified with alternatives in their settings in the training process and modification in their architectures to be able to predict the drug response using IC_{50} regression values. The recommended models are from the top 5 trained on ImageNet: InceptionV3, Xception, and EfficientNetB7 which proves their efficiency.

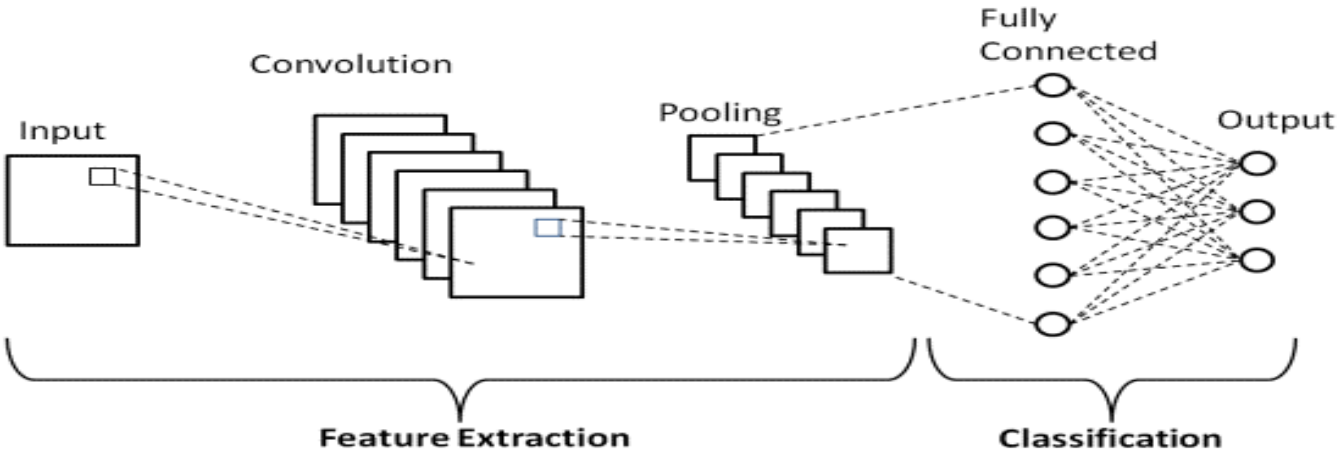


Figure 3. CNN network Architure

3.4.1 InceptionV3

InceptionV3 is introduced in [21], see Figure 4. It is used mainly for image classification and object detection. It is the third edition of Google's Inception Convolutional Neural Network, originally introduced during the ImageNet Recognition Challenge. The IncetionV3 model is efficient due to its low number of parameters compared to other models like VGGNet. So, it is suitable for big data where a huge amount of data needed to be processed at a reasonable cost.

3.4.2 Xception

Xception [22] is an Inception architecture that uses the concept of depthwise separable convolution, See Figure 5. To control the complexity, Xception use bigger block than the original inception block and replaced the multiple spatial dimensions with a single dimension (3x3) followed by a 1x1 convolution. Xception separates spatial and feature map, so it is computationally better than Inception. This architecture outperforms Inception V3, ResNet-50, ResNet-101, ResNet 152, and VGGNet on ImageNet. Xception simplifies computation by individually convolving each feature map across spatial axes, then performing cross-channel correlation via pointwise convolution (1x1 convolutions). Xception's

transformation use the same parameter number. Xception achieve better performance by enhance the learning process. Xception's reduces the number of connections.

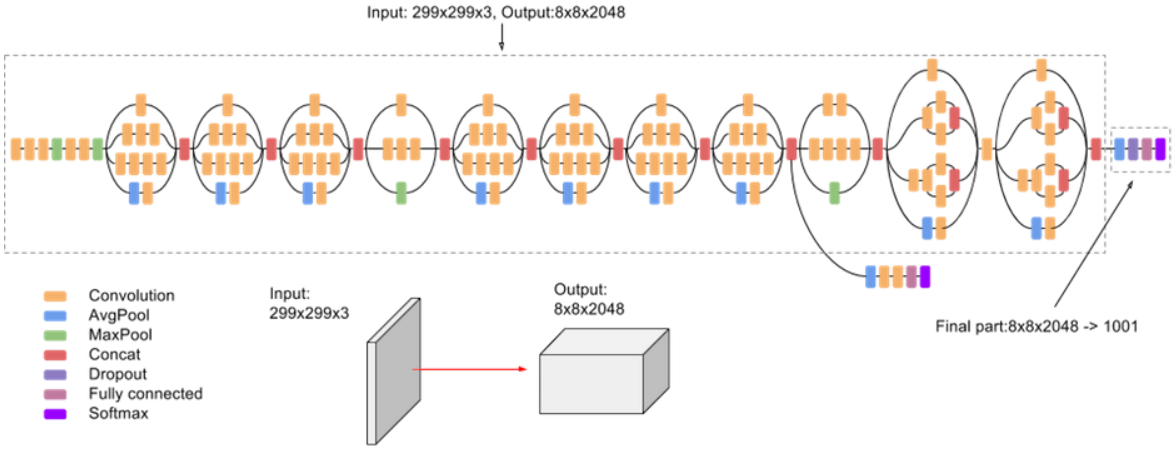


Figure 4. Inception V3 architecture [22]

3.4.3 EfficientNetB7

EfficientNetB7, one of the most recent models to CNN, is introduced in 2019 [23]. It is simply a scaling method that scales all dimensions (width, depth, resolution) with a set of scaling coefficients. It achieved better accuracy than InceptionV3, Xception, DenseNet, and ResNet models. The main unit in its architecture is a mobile inverted bottleneck (MBCon) [24] [25] See Figure 6.

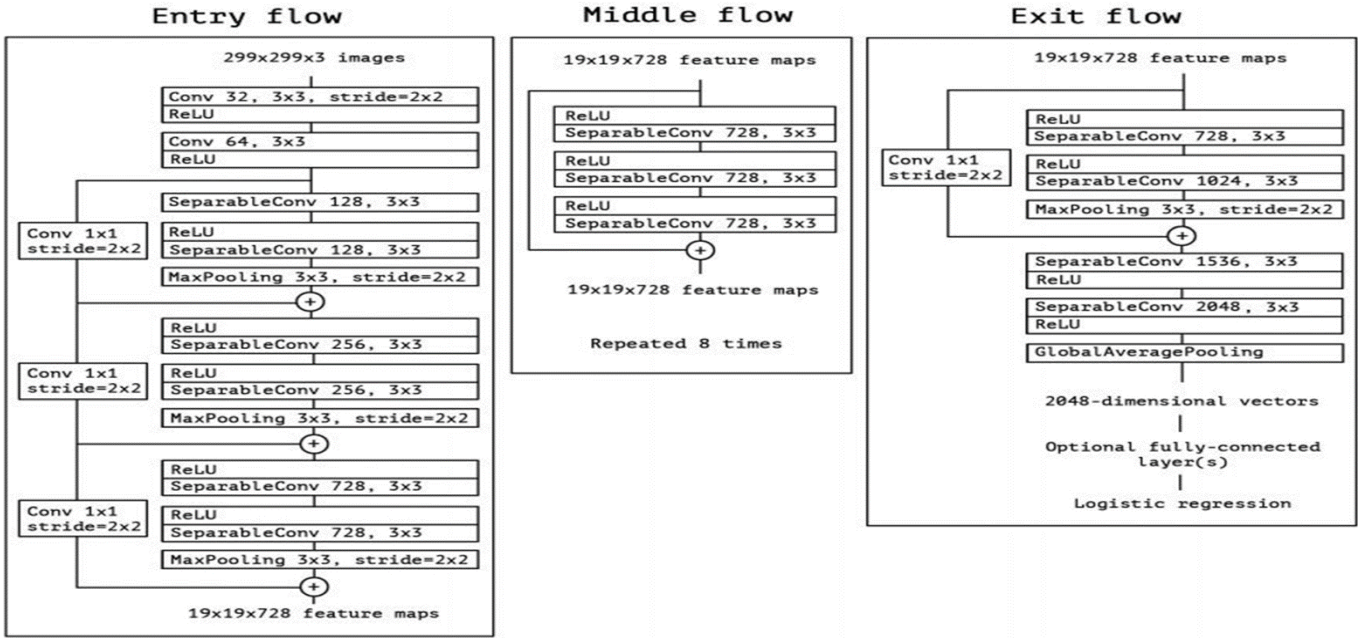


Figure 5. Xception Model Architecture [20]

Stage i	Operator $\hat{\mathcal{F}}_i$	Resolution $\hat{H}_i \times \hat{W}_i$	#Channels $\hat{C}_i$	#Layers $\hat{L}_i$
1	Conv3x3	224×224	32	1
2	MBConv1, k3x3	112×112	16	1
3	MBConv6, k3x3	112×112	24	2
4	MBConv6, k5x5	56×56	40	2
5	MBConv6, k3x3	28×28	80	3
6	MBConv6, k5x5	14×14	112	3
7	MBConv6, k5x5	14×14	192	4
8	MBConv6, k3x3	7×7	320	1
9	Conv1x1 & Pooling & FC	7×7	1280	1

Figure 6. EfficientNet baseline network

3.5 Proposed CNN-based Models

In this approach, different CNN models are used and investigated their performance, see Figure 7. The first step is how to make the data suitable for the CNN-based models, well known that it deals with images. To overcome this point, the DeepInsight [12] algorithm is used to convert the non-image input data to an image to suit the CNNs models. T SNE algorithm inside the Deepinsight is used as a dimension reduction algorithm Different image sizes also are used. Each sample is converted to a 2D image, so the number of images equals the number of cell line samples from CCLE (619 images as there are 619 cell lines from CCLE after applying data integration and federation introduced in [6]). The images are divided into training, validation, and testing datasets with the following percentages 80%, 10%, and 10% respectively.

For CNN-models: InceptionV3, Xception, and EfficientNetB7. For InceptionV3, Xception, and EfficientNetB7 Keras implementation is used with accuracy for Top-5 on ImgeNet 94.5%, 93.7% and 97.0% respectively. Those models are followed by the Average pooling 2D layers. Table 1 explained the used architecture of the dense layers. The output layer for all architectures is 265 nodes with a linear activation function. For the proposed CNN-models image size of 75 x 75 are used as input.

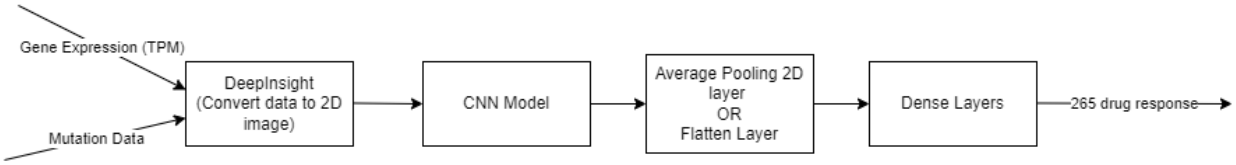


Figure 7. CNN-based model Architecture

Table 1. Dense Architecture Summary

Number of Layers	Number of Nodes	Activation function
1	64	ReLU
1	128	ReLU
1	1024	ReLU

4. Experimental Results

The proposed approach aims to predict the drug response by predicting the IC_{50} using the genetic and mutation profiles for CCLE and GDSC datasets. CCLE has 619 cell lines for 25 tissues with 19702 genes for each gene expression profile and mutation profile. First of all, the CNN-based models used with the minimum size of images can be used so for InceptionV3, Xception, and EfficientNetB7 the used image size is 75 x 75 images. The CNN-based models are run for 25 different datasets and each dataset runs 25 times. CNN-based models used the CCLE dataset as 80% training set, 10% validation set and 10% testing. The t-SNE is used as dimension reduction techniques for DeepInsight. For the InceptionV3, Xception, and EfficientNetB7 models, the used weights are the pre-trained weights on ImageNet or random weights. For models using pre-trained weights on ImageNet, two approaches are used one using the weights as it is and do not update them and the other one is allowing the update of the weights for those models while training Dense layers that use, He's uniform distribution [27] to initialize weight.

In Table 2, InceptionV3, Xception, and EfficientNetB7 are presented with minimum image sizes for each one. DeepInsignt deals with input image of size (75x75) pixels. DeepInsignt uses t-SNE as dimension reduction. All the models use a single dense layer after the flatten one with 64 nodes. For each algorithm, the table shows the best results of the best run according to the MSE value of the test dataset and the average results for all runs. The result shown in table 3, The EfficientNetB7 algorithm with randomly initialized weights (EfficientNetB7-Random), the Xception algorithm with random weights (Xception – Random), and the Xception algorithm with ImageNet weights (Xception–ImageNet) respectively achieved the best accuracy. But to ensure the model's stability, the average will be considered. The Xception model with all its variations shows the best stability performance. It also achieved the best convergence speed with an average between 7.33 and 8.87 epochs.

Table 3 shows the same models with the same image size and dimension reduction method as in Table 2. The dense layer is 128 nodes. The results show that Xception with both ImageNet and Random weights achieved the best performance with minimum MSE values. The best average performance is achieved by Xception models regardless of the method of initializing the weights and with the best convergence rate.

Table 2. The Results of CNN-based models using t-SNE and a Single 64 Dense Layer with a minimum size of images

Algorithm	Best – MSE				Average – MSE			
	Training	Validation	Test	#Epoch	Training	Validation	Test	#Epoch
InceptionV3-ImageNet	0.57	1.99	1.41	4	0.75	2.82	1.75	11.36
InceptionV3 - ImageNet – Trainable	0.62	1.67	1.57	17	0.724	1.74	2.5	24.6
InceptionV3 – Random	1.2	1.57	1.46	12	0.63	1.67	1.59	8.8
Xception – ImageNet	0.51	1.48	1.35	13	0.53	1.57	1.53	7.33
Xception - ImageNet – Trainable	0.45	1.628	1.4375	18	0.47	1.605	1.5683	8.2
Xception – Random	0.48	1.64	1.34	4	0.5	1.57	1.57	8.87
EfficientNetB7 –ImageNet	0.27416	1.94759	1.74905	4	0.29868	1.87756	1.90336	12.27
EfficientNetB7 –ImageNet – Trainable	0.274	1.948	1.749	4	0.2987	1.878	1.903	12.26

EfficientNetB7 – Random	1.17404	1.55189	1.33400	7.00000	1.20345	1.62474	1.61095	8.50000

Table 3. The Results of CNN-based models using t-SNE and a Single 128 Dense Layer with a minimum size of images

Algorithm	Best - MSE				Average - MSE			
	Training	Validation	Test	# Epoch	Training	Validation	Test	# Epoch
InceptionV3-ImageNet	1.03	1.65	1.55	13	0.68	1.87	1.79	10.86
InceptionV3 - ImageNet –Trainable	1.13	1.76	1.56	47	0.66	1.84	1.77	15.47
InceptionV3 - Random	0.53	1.9	1.57	6	0.55	1.79	1.77	9.28
Xception - ImageNet	0.28	1.6	1.49	19	0.3	1.6	1.64	8.4
Xception - ImageNet -Trainable	0.25	1.65	1.47	6	0.25	1.59	1.6	10.47
Xception - Random	0.36	1.67	1.49	13	0.38	1.55	1.56	7.2

Table 4 shows the same models with the same setting as in table 4 except that the dense layer is 1024 nodes. Table 4 shows that Xception models achieved the best results in all its cases.
From Table 2, 3, and 4, Xception models show the best accuracy with minimum MSE and best convergence rate. The t-SNE with a single dense layer has 64 nodes with the best setting which achieved the best accuracy.

Form all results for the proposed 3 CNN architectures; Xception–ImageNet achieves better overall results and stability. Individuals runs EfficientNetB7-Random achieves best.

Table 4. The Results of CNN-based models using t-SNE and a Single 1024 Dense Layer with a minimum size of images

Model	Best - MSE				Average - MSE			
	Training	Validation	Test	# Epoch	Training	Validation	Test	# Epoch
InceptionV3-ImageNet	0.37	1.64	1.61	6	0.57	2.0	1.74	12.72
InceptionV3 - ImageNet –Trainable	1.3	1.61	1.51	11	1.28	1.99	1.72	13.9
InceptionV3 – Random	0.3	1.8	1.54	5	0.37	1.71	1.66	8.48
Xception – ImageNet	0.008	1.7	1.437	12	0.2	1.66	1.62	8.93
Xception - ImageNet –Trainable	0.01	1.7	1.44	12	0.02	1.66	1.62	8.93
Xception – Random	0.02	1.59	1.44	7	0.04	1.61	1.58	7.47

5. Comparative Study

The Comparative study is used to prove the efficiency of the proposed model in terms of the accuracy. Some of the recent models based on DNN and CNNs are selected such as Enhanced Deep-DR [6], and Deep-DR [28] to be in consideration. Table 5 summarizes the datasets, data features, number of drugs, and prediction method that are used by the considered models. All models work on GDSC as a reference for drug response data. Table 6 shows the comparison among the different algorithms's performance using MSE values. The proposed Xception–ImageNet with image size (50 x50) model achieved the best accuracy by achieving the minimum MSE over both Deep-DR and Enhanced Deep-DR.

Table 5. Compare the setting of different algorithms

Algorithm	Datasets	Features	Number of genes	Number of drugs	Model Type
Enhanced Deep-DR [6]	CCLE, GDSC, TCGA	Gene expression and mutation	19702	265	FFN
Deep-DR [7]	CCLE, GDSC, TCGA	Gene expression and mutation	15363	265	FFN
Xception–ImageNet	CCLE, GDSC, TCGA	Gene expression and mutation	19702	265	CNN

Table 6. Compare different models in terms of MSE

Algorithm	MSE
Enhanced Deep-DR [6]	1.78
Deep-DR [7]	1.98
Xception–ImageNet	1.53

6. Conclusion

In this paper, we introduce the use of Convolutional Neural Network (CNN) models to predict 265 different drugs response. DeepInsight algorithm used to convert the input data to images to be more suitable as input to the CNN. the use of 3 different pretrained CNNs-models (InceptionV3, Xception, EfficientNetB7) with alternatives in their settings in the training are introduced. The 3 models used with their pre-trained weight on ImageNet in two ways fixed and updatable weights during the training process and also models used with completely random weights. The modification on their architectures is done to predict the drug response using IC_{50} regression values. Those models are selected due to their efficiency for ImageNet applications. The proposed modified Xception model achieves the best accuracy over the 2 others. At first, the whole data input passes through DeepInsight which converts the gene expression data and mutations data to images. T-SNE is used as dimension reduction inside the DeepIsignt. The Comparative analysis with other Deep models, shows that the proposed approach improves the prediction accuracy in a range between 14% and 22% as a reduction in mean squared error (MSE) compared to Enhanced Deep-DR and Deep-DR.

References

1. "World Health Organization," 12 September 2018. [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/cancer.
2. "Cancer Statistics," NCI, [Online]. Available: https://www.cancer.gov/about-cancer/understanding/statistics. [Accessed May 2022].
3. J. Liu, Y. Wu, I. Ong, D. Page, P. Peissig, C. McCarty, A. A. Onitilo, and E. Burnside, "Leveraging Interaction between Genetic Variants and Mammographic Findings for Personalized Breast Cancer Diagnosis," AMIA American Medical Informatics Association, 2015.

4. M. Verma, "Personalized Medicine and Cancer," Journal of Personalized Medicine , vol. 2, no. 1, p. 1–14, 2012.
5. H. Ahmed, S. Hamad, H. A. Shedeed and A. Saad, "Review of Personalized Cancer Treatment with Machine Learning," in ICCI, 2022.
6. H. Ahmed, S. Hamad, H. A. Shedeed and A. Saad, "Enhanced Deep Learning Model for Personalized Cancer Treatment," IEEE Access, 2022.
7. T. Sakellaropoulos, K. Vougas, S. Narang, R. Petty, A. Tsirigos and V. G. Gorgoulis, "A Deep Learning Framework for Predicting Response to Therapy in Cancer," Cell Reports, 2019.
8. H. Zhang, Y. Chen and F. Li, "Predicting Anticancer Drug Response With Deep Learning Constrained by Signaling Pathways," Front. Bioinform, vol. 1, 2021.
9. Y. Chang, H. Park, H. Yang, S. Lee, K. Lee, T. S. Kim, J. Jung and J. Shin, "Cancer Drug Response Profle scan (CDRscan): A Deep Learning Model That Predicts Drug Efectiveness from Cancer Genomic Signature," Scientific Reports, pp. 1-11, 2018.
10. P. Liu, H. Li, S. Li and K. Leung, "Improving prediction of phenotypic drug response on cancer cell lines using deep convolutional network," BMC Bioinformatics, no. 20, 2019.
11. Z. Zhao, K. Li, C. Toumazou and M. Kalofonou, "A computational model for anti-cancer drug sensitivity prediction," IEEE Biomedical Circuits and Systems Conference (BioCAS),, pp. 1-4, 2019.
12. A. Sharma, E. Vans, D. Shigemizu, K. A. Boroevich and T. Tsunoda, "DeepInsight: A methodology to transform a non-image data to an image for convolution neural network architecture," Scientific Reports, 2019.
13. J. Barretina, G. Caponigro, N. Stransky, K. Venkatesan,, A. A. Margolin, S. Kim, C. J. Wilson, J. Lehár, G. V. Kryukov, D. Sonkin, A. Reddy and et al, "The Cancer Cell Line Encyclopedia enables predictive modelling of anticancer drug sensitivity," Nature, vol. 483, p. 603–607, 2012.
14. M. Ghandi, F. W. Huang, J. Jané-Valbuena, G. V. Kryukov, C. C. Lo and E. R. McDonald, "Next-generation characterization of the Cancer Cell Line Encyclopedia," Nature, vol. 569, p. 503–508, 2019.
15. "TCGA Research Network:," [Online]. Available: https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga.
16. W. Yang, J. Soares, P. Greninger, E. J. Edelman and et al, "Genomics of drug sensitivity in cancer (GDSC): a resource for therapeutic biomarker discovery in cancer cells," Nucleic Acids Research, vol. 41, 2013.
17. R. Patro, G. Duggal, M. I. Love, R. A. Irizarry and C. Kingsford, "Accurate, fast, and model-aware transcript expression quantification with Salmon," bioRxiv, 2016.
18. "UCSC Xena datasets," University of California, Santa Cruz, [Online]. Available: https://xenabrowser.net/datapages/.

19. M. J. Goldman, B. Craft, J. Zhu and D. Haussler, "Abstract 250: UCSC Xena for the visualization and analysis of cancer genomics data," in the American Association for Cancer Research Annual Meeting, Philadelphia, 2021.
20. A. Dhillon and G. K. Verma , "Convolutional neural network: a review of models, methodologies and applications to object detection," Artif Intell, vol. 9, p. 85–112, 2019.
21. C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016.
22. F. Chollet, "Xception: Deep Learning with Depthwise Separable Convolutions," in IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017.
23. M. Tan and Q. V. Le, "EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks," in International Conference on Machine, California, 2019.
24. M. Sandler, A. Howard, M. Zhu, A. Zhmoginov and L. C. Chen, "Mobilenetv2: Inverted residuals and linear," in CVPR, 2018.
25. M. Tan, B. Chen, R. Pang, V. Vasudevan, M. Sandler, A. Howard and Q. V. Le, "MnasNet: Platform-aware neural architecture search for mobile," in CVPR, 2019.
26. L. J. P. V. D. M. Maaten and G. Hinton, "Visualizing High-Dimensional Data using t-SNE," Journal of Machine Learning Research, vol. 9, p. 2579–2605, 2008.
27. K. He, X. Zhang, S. Ren and J. Sun, "Delving deep into rectifiers: surpassing human level performance on imagenet classification," in Proceedings of the IEEE international conference on computer vision, 2015.
28. Y. Chiu, H. H. Chen, T. Zhang, S. Zhang, A. Gorthi, L. Wang, Y. Huang and Y. Chen, "Predicting drug response of tumors from integrated genomic profiles by deep neural networks," BMC Medical Genomics, vol. 12, pp. 143-189, 2019.