

Feasibility Study of Deep Learning Based Radiation Sensitivity Prediction Model Using Gene Expression Profiling

Euidam Kim

Hanyang University

Yoonsun Chung (✉ ychung@hanyang.ac.kr)

Hanyang University

Research Article

Keywords: radiation sensitivity prediction, vitro radiation sensitivity prediction, gene expression

Posted Date: July 21st, 2021

DOI: <https://doi.org/10.21203/rs.3.rs-719143/v1>

License:   This work is licensed under a Creative Commons Attribution 4.0 International License.

[Read Full License](#)

Feasibility Study of Deep Learning Based Radiation Sensitivity Prediction Model Using Gene Expression Profiling

Euidam Kim¹ and Yoonsun Chung, Ph.D.^{1,*}

*Department of Nuclear Engineering, Hanyang University College of Engineering, 222 Wangsimni-ro,
Seongdong-gu, Seoul 04763, Republic of Korea*

***Corresponding author:**

Yoonsun Chung, Ph.D.

Department of Nuclear Engineering, Hanyang University,

222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

Tel: 82-2-2220-0463; Fax: 82-2-2220-0463, E-mail: ychung@hanyang.ac.kr

Abstract

Since radiation sensitivity prediction can be used in various field, we investigate the feasibility of an *in vitro* radiation sensitivity prediction model using a deep neural network. A microarray of the National Cancer Institute-60 tumor cell lines and clonogenic surviving fraction at an absorbed dose of 2 Gy values are used to predict radiation sensitivity. The prediction model is based on convolutional neural network and 6-fold cross-validation approach is applied to validate the model. Of the 174 samples, 170 (97.7%) samples show less than 10% and 4 (2.30%) show more than 10% of relative error, respectively. Through an additional validation, model accurately predict 172 out of 174 samples, representing a prediction accuracy of 98.85% under the criteria of absolute error < 0.01 or the relative error $< 10\%$. This results demonstrate that *in vitro* radiation sensitivity prediction from gene expression can be carried out with the deep learning technology.

Introduction

Quantification for response prediction of normal tissue and tumor to radiation has been considered to be necessary in radiation risk assessment, radiation protection, or even radiotherapy. In radiation protection, it is basically assumed that members of the population subject to protection are equally sensitive to adverse health effects related to radiation exposure, which is the limitation of existing radiological protection practices. However, in order to improve this weakness under consideration of the various radiation sensitivity differences among members in a protection group, an accurate and robust method to evaluate the radiation sensitivity of individuals or subgroups is needed¹. Likewise, in radiotherapy, accurate prediction of radiation sensitivity is critical for determining patient-specific treatment methods, doses, fractionation schedules, corresponding clinical outcomes, and reducing possible side effects of radiotherapy^{2,3}.

From this perspective, several researchers have investigated that the sensitivity of cancer cells to radiation damage depends on the type, characteristics, and the gene expression level of the cancer cells⁴. Simultaneously, advances in gene expression profiling technology have allowed the analysis of a growing variety of genetic factors that influence gene expression in cancer cells⁵. Based on this, several recent studies have shown that *in vitro* radiation sensitivity can be quantitatively analyzed based on gene expression profiling data and have suggested models that can predict radiation sensitivity from gene expression data^{3,4,6,7,8,9}. These studies have improved our understanding of the relationship between gene expression and radiation sensitivity. However, further discussion and research are still needed to establish a robust paradigm for predicting radiation sensitivity^{4,10}.

Meanwhile, as a novel prediction and decision-making methodology, deep learning has recently emerged as a major tool for decision-making, classification, and prediction. The deep learning model updates itself using the hidden relationships between numerous data, which clearly exists but hard to represent numerically. With this characteristics, it seems reasonable to expect that deep learning can improve the performance of the prediction models when applied to *in vitro* radiation sensitivity prediction.

Therefore, in this study, we aimed to investigate the feasibility of a prediction model that predicts *in*

vitro radiation sensitivity from gene expression profiling data, based on several previously established deep learning modalities. Moreover, by comparing the performance of the resulting model with that of models from previous studies, we demonstrate the applicability and potential power of using deep learning algorithms to predict radiation sensitivity from gene expression data.

Results

Development of a deep learning-based radiation sensitivity prediction model

Fig. 1 shows the overall flowchart of the radiation sensitivity prediction model. A total of 174 samples from 59 NCI-60 cell lines and corresponding SF2 values were split into test and training sets by a 6-fold cross-validation. For each round of cross-validation, the training sets were fed to the model, and the parameters were trained based on a gradient descent algorithm. After training, the test data were fed to the model, and the SF2 value of the test data was predicted. Evaluation metrics including absolute error, relative error, and prediction accuracy were calculated for the entire test set. If a sample failed to be classified as “correct prediction”, such samples were subjected to additional prediction validation. If the error was still larger than the criteria after the prediction validation, these samples were classified as “prediction hard” cases. These processes were performed over the entire cross-validation test set, and the performance of the model was measured through prediction accuracy for the entire dataset obtained through this process.

Performance of model and validation of prediction

Table 1 shows the average of the predicted radiation sensitivity in five rounds of the 6-fold cross-validation experiment. As shown in fig.2, of the 174 triplicated samples, 142 (81.61%), 28 (16.09%), and 4 (2.30%) samples were included in groups with relative errors of less than 2%, 2 to 10%, and 10% or more, respectively. The model correctly predicted 171 of the 174 samples, indicating that the initial prediction accuracy of the model was 98.28%. Three samples (red points in the Fig. 3; one each from the cell lines MOLT-4, MDA-MB-435, and HL-60) with abnormally larger error (527.59%, 129.11%, and 72.88% of relative error, respectively) were subjected to prediction validation and the predicted SF2s were 0.302, 0.362, and 0.301 (relative error of 504.27%, 102.34%, and 4.54%), respectively. Therefore, one sample (from HL-60) was changed to “correct prediction”, and only two samples (one each from MOLT-4 and MDA-MB-435) that produced a relative error larger than 10% were classified as final “prediction-hard” cases. Overall average, standard deviation, absolute and relative error, and the other detailed information can be found in table S1.

As shown in Fig. 3, the SF2 value predicted by our model and the true (measured) SF2 value had a distinct linear correlation, indicating that the model successfully predicted the radiation sensitivity of the cell lines from their gene expression data (95% CI: 0.9834 to 0.9909, Pearson's $r = 0.9877$).

The average relative error and absolute error of the "correct prediction" samples were $1.351 \pm 1.875\%$ and 0.00596 ± 0.00638 , respectively ($n=172$). In contrast, the relative errors of the "prediction-hard" cases were 102.34% (MDA-MB-435) and 504.27% (MOLT-4), and the absolute errors were 0.1832 and 0.2521, respectively. The overall prediction accuracy after the validation of prediction was 98.85% (172 out of 174 were correct), and the RMSD was 0.0252 with prediction-hard cases and 0.00867 without prediction-hard cases.

Discussion

Deep learning is a recently emerging research field that has gained prominence as hardware has advanced. It is widely used for decision-making, prediction, and classification. With this perspective, we proposed the feasibility of deep learning as a novel methodology by developing a deep learning-based *in vitro* radiation sensitivity prediction model from gene expression with an accuracy of 98.85%. This is the first study to attempt to use deep learning in *in vitro* radiation sensitivity prediction.

In their analyses of model accuracy, Torres-Roca et al. and Zhang et al. who similarly tried to predict the radiation sensitivity of the NCI-60 cancer cell lines both set the criteria of the “correct prediction” when the predicted SF2s were within 10% of the true (measured) values^{4,9}. With these criteria, they proposed models with an accuracy of 62% (22 out of 35) and 91% (54 out of 59), respectively.

Comparably, in our study, 172 samples out of 174 samples were correctly classified using the similar but more reasonable criteria described in material section, representing a 98.85% accuracy. Moreover, the RMSD of our deep learning model was 0.0251 with the prediction-hard cases and 0.00867 without the prediction-hard cases, compared to 0.2 described by Torres-Roca et al., or 0.011 of Zhang et al.^{4,9}. These results indicate that biological nonlinear complex interactions that influence the radiation sensitivity of a cell are likely to be well represented by a deep learning model.

In case of the three samples with large errors in our study, one each from the cell lines MOLT-4, MDA-MB-435, and HL-60, these were subjected to prediction validation because it might not be due to merely variance and bias over trials. This was supported by the fact that the fluctuation, represented by the standard deviation of the prediction of each round of the experiment of these data, was not significantly different compared to the other samples, and the other samples in the same cell line showed a relatively low error and the prediction.

As a result of this prediction validation, the sample from HL-60 cell line showed a significantly improved prediction error. Hence, it can be inferred that the large error of the HL-60 sample from the initial prediction appears to be due to a lack of training data in the training fold of initial prediction.

For the remaining “prediction-hard” cases, MOLT-4 and MDA-MB-435, we were unable to determine

whether there was an insufficient amount of data to predict their SF2s correctly or if there were other possible problems that could not be investigated in this study, such as mislabeling issues. From this perspective, further research is needed to investigate these problems. What is noteworthy, even if these “prediction-hard” cases are due to a lack of training samples, the model still predicted these samples as radiosensitive. It could be considered that the model has a resistance to these “prediction-hard” cases, such that the model is still able to predict whether the cell is radiosensitive or not, which is fundamentally important.

There are two major limitations of this study. First, it should be noted that in general, deep learning algorithms are fed enormous amounts of data to train the model and thereby enable the model to provide general decision making as AlphaGo does ¹¹. However, in this study, the limited number of cell lines sample with survival data available for training may not have fully demonstrated the overall potential of deep learning. Thus, it seems necessary to further boost the performance of this deep learning-based radiation sensitivity prediction model by additional training using a large amount of radiation sensitivity datasets from not only NCI-60 cell lines but also from other types of cancer cell lines. Second, it may be considered as a limitation of using classical microarray analysis rather than the latest gene expression profiling methodology, RNA sequencing. Though microarray is a little outdated method and is constantly being replaced to the RNA sequencing, we used microarray data to demonstrate the feasibility of deep learning aided radiation sensitivity prediction through comparison with previous studies. In this perspective, further research is needed regarding the prediction model using RNA sequencing data.

Despite these limitations, several improvements in radiation sensitivity prediction analyses are expected from our study. First, since deep learning aims to “let the data speak” without any additional step to extract the feature that represents the characteristics of the input data (as is the case in existing statistical methods), we can expect the model to learn to represent a direct and transparent relationship between the input genes since data with large dimensions are fed as an input variable ¹². Second, the deep learning model can further learn (trained) from additional data presented after training, which enables deep learning to self-correct and absorb huge amount of data to make itself more robust ¹³.

Third, a characteristic of the literally “deep” model enables high-level feature learning, especially effective when it comes to handling complexly combined data such as genetic information. Therefore, the deep learning based methodology can provide better model performance compared to the conventional statistical shallow machine learning-based model, which leads more valid and accurate prediction result.

In deep learning, causes and results are the only information provided. One of their characteristics is that they maintain “black boxes” with respect to their internal processes even though they provide good results. Deep learning used in this study is also very useful for its ability to predict radiation sensitivity with high accuracy, but it cannot provide any scientific explanation for how such predictions are made. Therefore, further research is needed to reduce the non-explainability of deep learning. With advanced research that attempts to understand and explain the inner world of deep learning, it will help to identify the biological and medical mechanisms of how organisms react to radiation exposure.

In conclusion, this study successfully demonstrated the feasibility of a deep learning-based *in vitro* radiation sensitivity prediction model using gene expression profiling data. We established a CNN-based feature extractor and residual block-added prediction part of the model with previously established deep learning methodologies. With additional research and external validation, this model and its methodology can be expanded to *in vivo* radiation sensitivity prediction.

Methods

Radiation response

Since the clonogenic surviving fraction of cells at an absorbed dose of 2 Gy (SF2) is widely used as a measurement of *in vitro* radiation sensitivity, we also selected SF2 as an indicator of radiation sensitivity in this study. The true (measured) SF2 values used in this study were obtained from previous publications ^{9, 14}.

National Cancer Institute-60 (NCI-60) cell lines

The NCI-60 panel contains 60 cell lines representing nine types of tumors. It was established by the US National Cancer Institute in the 1980s for *in vitro* drug screening ⁶. This NCI-60 panel is now a valuable research resource, considering the continuous use of this panel for investigations of radiation response analysis ^{9, 15, 16, 17}. With this perspective, this panel was used as a platform representing multiple cancer cell lines to evaluate the performance of the radiation sensitivity prediction model of this study.

Gene expression profiling data

Gene expression profiling data of NCI-60 cancer cell lines were obtained from the Gene Expression Omnibus (GEO; available at <https://www.ncbi.nlm.nih.gov/sites/GDSbrowser>; series accession number GSE32474 ¹⁸) database, generated from microarray analysis performed with Affymetrix Human Genome U133 Plus 2.0 chips (54,675 probe sets). The entire transcript/gene set from the Affymetrix array was used to predict radiation sensitivity. Excluding the melanoma cell line MDA-N, which was shown to be “not available” from the NCI-60, duplicated or triplicated 174 samples of remaining 59 tumor cell lines were used as inputs in the radiation sensitivity prediction model.

Radiation sensitivity prediction modeling

The deep learning-based radiation sensitivity prediction model is based on the architecture of convolutional neural network (CNN). It comprises two distinct components: a feature vector extractor

based on a convolutional layer and a fully connected (FC) layer.

First, a feature extractor based on a convolutional layer was designed¹⁹. A convolutional layer is a type of layer consisting neural network that only connects nodes within a certain range, which leads to two distinct advantages: inherently prevent overfitting and can be trained with a relatively small amount of data. In this study, high-level feature vectors were extracted from the input gene expression vector using a one-dimensional convolutional layer with pooling and no padding.

After convolutional layers, radiation sensitivity is predicted via the FC layer with residual skip-connection²⁰. This FC layer utilizes a skip connection designed to make calculated gradients propagate over several hidden layers along the gradient descent algorithm, allowing the deep learning model to be constructed more deeply^{20, 21}. The residual block is applied by skipping each layer one by one. The overall structure of the prediction model is presented in Table 2.

For both the convolutional layer-based feature extractor and the residual block-added FC layer, a leaky rectified linear unit activation was applied²². L2 regularization and dropout while training were also used at the end of every convolution and FC layer while training to prevent overfitting to specific data or feature parts and to let the model learn from all interactions within the entire dataset^{19, 23, 24, 25}.

Training and testing of the prediction model

To train and test the developed model, five rounds of 6-fold cross-validation were applied. The 6-fold cross-validation method divides the entire dataset into six sub-datasets and uses each dataset in turn as a test set with the remaining five datasets used as a training set to test the model. To prevent overfitting to data of a certain tumor cell line, a 6-fold cross-validation was designed such that the data of a particular cell line were not included in the same fold. The final predicted SF2 was determined as the average of five rounds of independent cross-validations to increase the stability and reduce the deviation of the predicted value.

Measurement of model performance

The performance of the radiation sensitivity prediction model was evaluated based on calculation of

the root mean squared deviation (RMSD). RMSD is defined as

$$\text{RMSD} = \sqrt{\frac{\sum_{i=1}^T (\hat{Y}_t - Y_t)^2}{T}}, \quad (1)$$

where \hat{Y}_t represents the true (measured) SF2 of sample t , Y_t represents the SF2 value predicted by the model, and T represents the number of samples.

The absolute error and the relative error of the SF2 prediction were defined as the absolute deviation between the true and predicted SF2 and the absolute error divided by the true SF2, respectively.

In order to evaluate the performance of this prediction model, the “correct prediction” criteria were defined. In previous studies, the correct prediction was defined using only relative error, following the known variability of the clonogenic cell survival assay ⁴. However, it tended to be overly strict in cases with a low value of true SF2. Therefore, we classified a “correct prediction” if either the absolute error of the sample is less than 0.01 (1% in terms of survival fraction, not considered to be clinically significant) or the relative error between the measured and predicted SF2 is less than 10%.

The model was evaluated and trained with the NVIDIA TITAN RTX and the TensorFlow 1.14.0 framework based on Python version 3.6.8.

Validation of the prediction

If a particular sample cannot be classified as correctly predicted, it is necessary to determine whether the error was caused by insufficient training set data used for sample prediction due to the folded cross-validation, or whether the total data used in the study could not provide sufficient explanation to predict that sample correctly. Therefore, additional experiments were conducted to identify such “prediction-failed” samples, using them as an independent test set and all the other samples as a broader training set.

If the prediction was successful in this additional experiment, it could then be determined that the corresponding fold was not able to provide sufficient evidence to predict the data correctly.

Conversely, if the prediction failed again, the data could be classified as “prediction-hard” cases.

Statistical analysis

Statistical analysis was utilized to evaluate the predictive performance of the model. Two-tailed Pearson correlation analysis with a 95% confidence interval was used to investigate the correlation between true (measured) SF2 and predicted SF2. Statistical analysis was performed using GraphPad Prism version 7.03 (GraphPad Software, San Diego, CA, USA).

Data availability

The datasets analysed during the current study are available in the gene expression omnibus (GEO; series accession number GSE 32474) repository, <https://www.ncbi.nlm.nih.gov/sites/GDSbrowser>. All data generated during this study are included in this published article (and its Supplementary Information files).

References

1. Bouffler SD. Evidence for variation in human radiosensitivity and its potential impact on radiological protection. *Annals of the ICRP* **45**, 280-289 (2016).
2. Hirst DG, Robson T. Molecular biology: the key to personalised treatment in radiation oncology? *Br J Radiol* **83**, 723-728 (2010).
3. Kim HS, *et al.* Identification of a radiosensitivity signature using integrative metaanalysis of published microarray data for NCI-60 cancer cells. *BMC Genomics* **13**, 348 (2012).
4. Torres-Roca JF, *et al.* Prediction of radiation sensitivity using a gene expression classifier. *Cancer Res* **65**, 7169-7176 (2005).
5. Ramaswamy S, Golub TR. DNA microarrays in clinical oncology. *J Clin Oncol* **20**, 1932-1941 (2002).
6. Amundson SA, *et al.* Integrating global gene expression and radiation survival parameters across the 60 cell lines of the National Cancer Institute Anticancer Drug Screen. *Cancer Res* **68**, 415-424 (2008).
7. Khan J, *et al.* Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks. *Nat Med* **7**, 673-679 (2001).
8. Ogawa K, Murayama S, Mori M. Predicting the tumor response to radiotherapy using microarray analysis (Review). *Oncol Rep* **18**, 1243-1248 (2007).

9. Zhang C, Girard L, Das A, Chen S, Zheng G, Song K. Nonlinear quantitative radiation sensitivity prediction model based on NCI-60 cancer cell lines. *ScientificWorldJournal* **2014**, 903602 (2014).
10. Peters LJ, Brock WA, Chapman JD, Wilson G. Predictive assays of tumor radiocurability. *Am J Clin Oncol* **11**, 275-287 (1988).
11. Silver D, *et al.* Mastering the game of Go with deep neural networks and tree search. *Nature* **529**, 484-489 (2016).
12. Meyer P, Noblet V, Mazzara C, Lallement A. Survey on deep learning for radiotherapy. *Comput Biol Med* **98**, 126-146 (2018).
13. Pan SJ, Yang Q. A Survey on Transfer Learning. *IEEE Transactions on Knowledge and Data Engineering* **22**, 1345-1359 (2010).
14. Eschrich S, *et al.* Systems biology modeling of the radiation sensitivity network: a biomarker discovery platform. *Int J Radiat Oncol Biol Phys* **75**, 497-505 (2009).
15. Boyd MR, Paull KD. Some practical considerations and applications of the national cancer institute in vitro anticancer drug discovery screen. *Drug Development Research* **34**, 91-109 (1995).
16. Burkard M. Integrating the NCI-60 Data with “Omics” for Drug Discovery. *Drug*

Development Research **73**, (2012).

17. Shoemaker RH. The NCI60 human tumour cell line anticancer drug screen. *Nat Rev Cancer* **6**, 813-823 (2006).
18. Pfister TD, *et al.* Topoisomerase I levels in the NCI-60 cancer cell line panel determined by validated ELISA and microarray analysis and correlation with indenoisoquinoline sensitivity. *Mol Cancer Ther* **8**, 1878-1884 (2009).
19. Krizhevsky A, Sutskever I, Hinton GE. Imagenet classification with deep convolutional neural networks. In: *Advances in neural information processing systems*) (2012).
20. He K, Zhang X, Ren S, Sun J. Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*) (2016).
21. Glorot X, Bengio Y. Understanding the difficulty of training deep feedforward neural networks. In: *Proceedings of the thirteenth international conference on artificial intelligence and statistics*) (2010).
22. Maas AL, Hannun AY, Ng AY. Rectifier nonlinearities improve neural network acoustic models. In: *Proc. icml*) (2013).
23. Dahl GE, Sainath TN, Hinton GE. Improving deep neural networks for LVCSR using rectified linear units and dropout. In: *2013 IEEE International Conference on Acoustics, Speech and Signal Processing*) (2013).

24. Hinton GE, Srivastava N, Krizhevsky A, Sutskever I, Salakhutdinov RR. Improving neural networks by preventing co-adaptation of feature detectors. *arXiv preprint arXiv:12070580*, (2012).

25. Nair V, Hinton GE. Rectified linear units improve restricted boltzmann machines. In: *Proceedings of the 27th international conference on machine learning (ICML-10)* (2010).

Acknowledgments

This research was supported by Basic Science Research Program through the National Research Foundation of Korea (KRF) funded by the Ministry of Education (NRF-2018R1D1A1B07049228) and by the research fund of Hanyang University (HY-2018).

Author contributions

YC conceived and designed the study, and. EK performed the study, analyzed the data. YC and EK prepared all figures and tables, and composed the main manuscript. All authors reviewed the manuscript.

Conflict of interest statement

The authors declare no conflict of interests.

Materials & Correspondence

Yoonsun Chung, Ph.D.

Department of Nuclear Engineering, Hanyang University,

222 Wangsimni-ro, Seongdong-gu, Seoul 04763, Republic of Korea

Tel: 82-2-2220-0463; Fax: 82-2-2220-0463, E-mail: ychung@hanyang.ac.kr

Figures

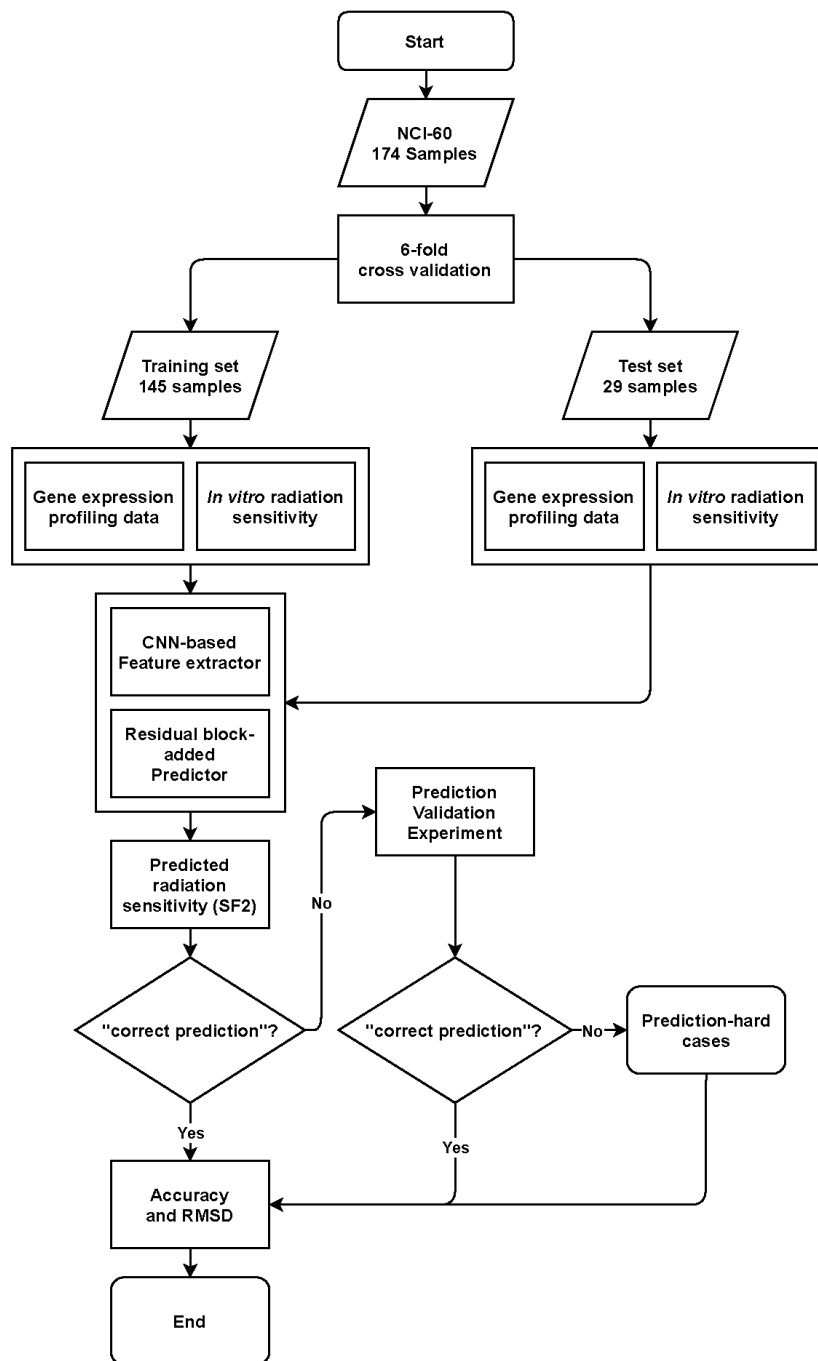


Figure 1. Overall flowchart of the radiation sensitivity prediction model.

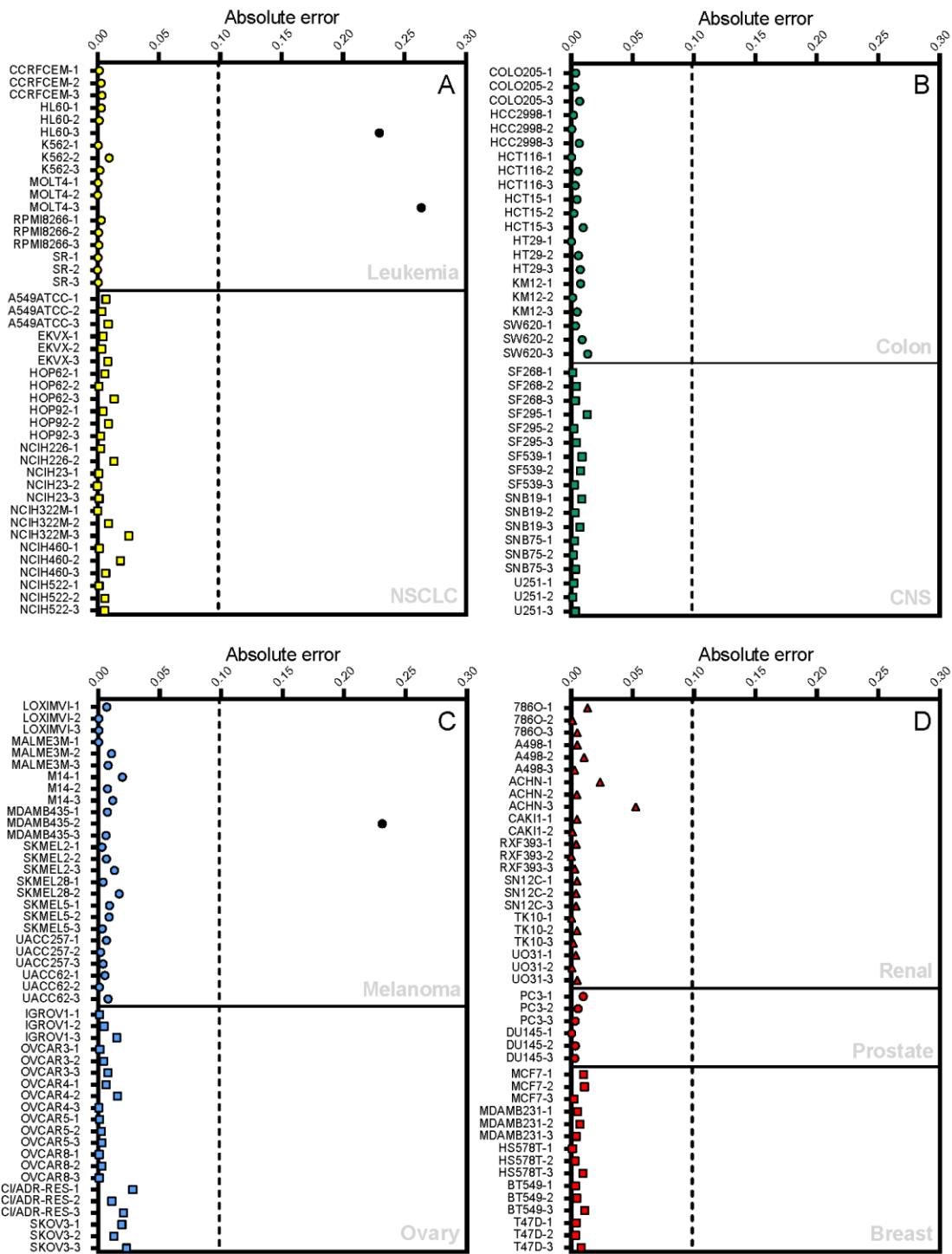


Figure 2. Plotted absolute error of all predicted triplicated samples of NCI-60 cell lines. Dark circles are representing three samples (HL-60, MOLT-4, and MDAMB-435) who were subjected to prediction validation. Vertical dotted line is a threshold for correct prediction of absolute error. The order and classification of the samples were based on NCI-60 panel. Abbreviation: NSCLC = Non-

Small Cell Lung Cancer; CNS = Central Nervous System

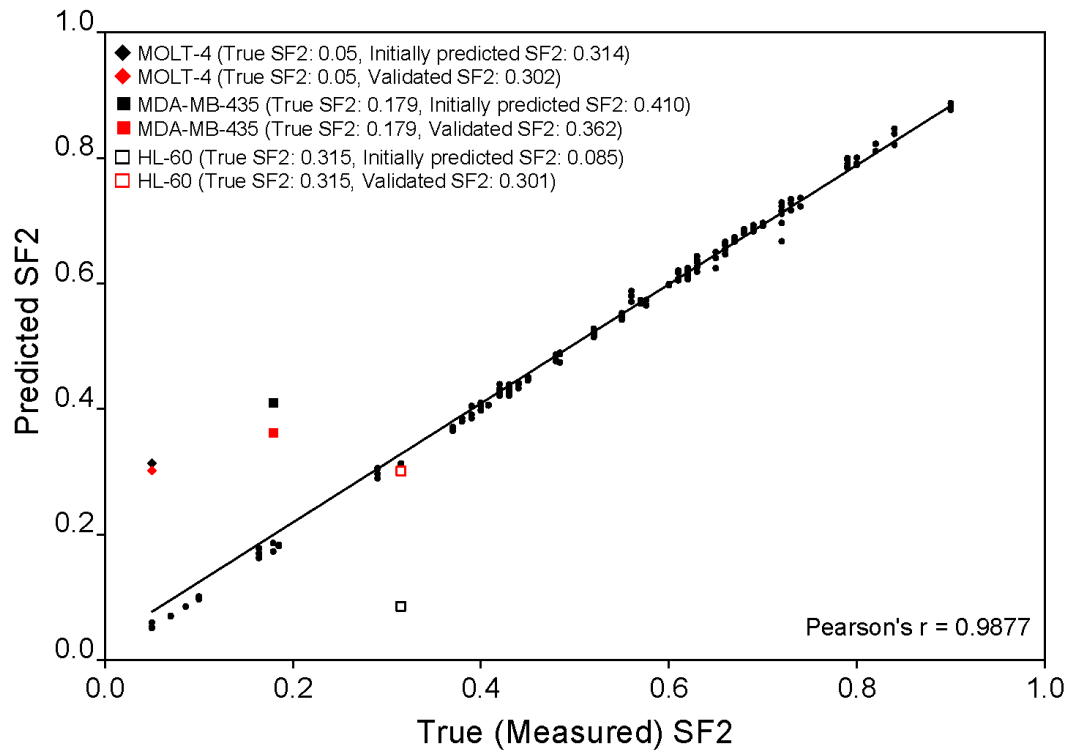


Figure 3. Linear correlation between true (measured) and predicted survival fraction at 2 Gy. Dark circles are representing the predicted SF2s of every triplicated samples.

Main Tables

Table 1. Predicted survival fraction at 2 Gy (SF2) value by the model

Cell lines	True SF2	Predicted SF2	Cell lines	True SF2	Predicted SF2	Cell lines	True SF2	Predicted SF2	Cell lines	True SF2	Predicted SF2
CCRF-CEM	0.185	0.183	COLO205	0.69	0.686	MALME-3M	0.8	0.8	786-O	0.66	0.647
		0.182			0.693			0.789			0.661
		0.181			0.683			0.792			0.655
HL-60	0.315	0.312	HCC-2998	0.44	0.442	M14	0.42	0.44	A498	0.61	0.605
		0.313			0.439			0.427			0.62
		0.085			0.433			0.432			0.607
K-562	0.05	0.049	HCT-116	0.38	0.38	MDA-MB-435	0.179	0.186	ACHN	0.72	0.696
		0.06			0.385			0.173			0.715
		0.052			0.383			0.41			0.667
MOLT-4	0.05	0.051	HCT-15	0.4	0.405	SK-MEL-2	0.66	0.663	CAKI-1	0.37	0.365
		0.05			0.398			0.667			0.371
		0.314			0.41			0.647			0.674
RPMI-8266	0.1	0.097	HT29	0.79	0.79	SK-MEL-28	0.74	0.736	RXF393	0.67	0.67
		0.099			0.784			0.723			0.667
		0.101			0.797			0.729			0.625
SR	0.07	0.069	KM12	0.42	0.428	SK-MEL-5	0.72	0.711	SN12C	0.62	0.624
		0.07			0.421			0.723			0.616
		0.071			0.425			0.487			0.52
A549	0.61	0.617	SW-620	0.62	0.616	UACC-257	0.48	0.478	TK-10	0.52	0.525
		0.606			0.611			0.476			0.522
		0.619			0.607			0.515			0.624
EKVX	0.7	0.695	SF-268	0.45	0.451	UACC-62	0.52	0.521	UO-31	0.62	0.621
		0.697			0.446			0.528			0.625
		0.692			0.447			0.391			0.474
HOP-62	0.164	0.17	SF-295	0.73	0.717	IGR-OV1	0.39	0.385	PC-3	0.484	0.49
		0.163			0.728			0.405			0.487
		0.178			0.734			0.549			0.52
HOP-92	0.43	0.435	SF-539	0.82	0.811	OVCAR-3	0.55	0.546	DU-145	0.52	0.517
		0.421			0.812			0.542			0.517
		0.427			0.823			0.296			0.566
NCI-H226	0.63	0.627	SNB-19	0.43	0.439	OVCAR-4	0.29	0.305	MCF7	0.576	0.565
		0.644			0.427			0.29			0.574
		0.085			0.437			0.407			0.635
NCI-H23	0.086	0.086	SNB-75	0.55	0.553	OVCAR-5	0.408	0.406	MDA-MB-231	0.63	0.637
		0.085			0.548			0.405			0.626
		0.65			0.554			0.599			0.791
NCI-H322M	0.65	0.641	U251	0.57	0.568	OVCAR-8	0.6	0.597	HS578T	0.79	0.787
		0.624			0.571			0.599			0.8
		0.838			0.573			0.588			0.627
NCI-H460	0.84	0.821	LOXIMVI	0.68	0.687	NCI/ADR-RES	0.56	0.571	BT-549	0.63	0.635
		0.847			0.68			0.58			0.619
		0.428			0.681			0.881			0.524
NCI-H522	0.43	0.436	-	-	-	SK-OV-3	0.9	0.887	T-47D	0.52	0.523
		0.436						0.877			0.528

Abbreviations: NSCLC, Non-Small Cell Lung Cancer; CNS, Central Nervous System

Table 2. Summary of the radiation sensitivity prediction model structure

Category	Layers	Output Size	Activation
Convolutional Feature Extractor	Input	$1 \times 54765 \times 1$	None
	Convolution Layer 1	$1 \times 23242 \times 10$	Leaky ReLU
	Pooling Layer 1	$1 \times 11621 \times 10$	None
	Convolution Layer 2	$1 \times 4788 \times 20$	Leaky ReLU
	Pooling Layer 2	$1 \times 2394 \times 20$	None
	Convolution Layer 3	$1 \times 942 \times 40$	Leaky ReLU
	Pooling Layer 3	$1 \times 471 \times 40$	None
	Convolution Layer 4	$1 \times 172 \times 80$	Leaky ReLU
	Pooling Layer 4	$1 \times 86 \times 80$	None
	Convolution Layer 5	$1 \times 28 \times 160$	Leaky ReLU
Pooling Layer 5	$1 \times 14 \times 160$	None	
Fully Connected Predictor	Flattening Layer	1×2240	None
	Fully-connected Layer 1	1×800	Leaky ReLU
	Fully-connected Layer 2	1×256	Leaky ReLU
	Fully-connected Layer 3	1×100	Leaky ReLU
	Fully-connected Layer 4	1×32	Leaky ReLU
	Output	1×1	Absolute value

Abbreviation: ReLU, Rectified Linear Unit

Supplementary Files

This is a list of supplementary files associated with this preprint. Click to download.

- [SupplementaryInformation.docx](#)