Nuclear Engineering and Technology 54 (2022) 1439–1448

Contents lists available at ScienceDirect

Nuclear Engineering and Technology

journal homepage: www.elsevier.com/locate/net

Original Article

Feasibility study of deep learning based radiosensitivity prediction model of National Cancer Institute-60 cell lines using gene expression

Euidam Kim, Yoonsun Chung*

Department of Nuclear Engineering, Hanyang University, Seoul, Republic of Korea

ARTICLE INFO

Article history:
Received 3 December 2020
Received in revised form
24 September 2021
Accepted 14 October 2021
Available online 16 October 2021

Keywords:
Radiosensitivity
Prediction
Deep learning
Gene expression
Survival fraction at 2Gy
Convolutional neural network

ABSTRACT

Background: We investigated the feasibility of *in vitro* radiosensitivity prediction with gene expression using deep learning.
Methods: A microarray gene expression of the National Cancer Institute-60 (NCI-60) panel was acquired from the Gene Expression Omnibus. The clonogenic surviving fractions at an absorbed dose of 2 Gy (SF2) from previous publications were used to measure *in vitro* radiosensitivity. The radiosensitivity prediction model was based on the convolutional neural network. The 6-fold cross-validation (CV) was applied to train and validate the model. Then, the leave-one-out cross-validation (LOOCV) was applied by using the large-errored samples as a validation set, to determine whether the error was from the high bias of the folded CV. The criteria for correct prediction were defined as an absolute error<0.01 or a relative error<10%.
Results: Of the 174 triplicated samples of NCI-60, 171 samples were correctly predicted with the folded CV. Through an additional LOOCV, one more sample was correctly predicted, representing a prediction accuracy of 98.85% (172 out of 174 samples). The average relative error and absolute errors of 172 correctly predicted samples were 1.351±1.875% and 0.00596±0.00638, respectively.
Conclusion: We demonstrated the feasibility of a deep learning-based *in vitro* radiosensitivity prediction using gene expression.
© 2021 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

1. Introduction

Prediction and quantification of radiation response of normal tissue and tumor have been considered to be necessary for radiation risk assessment, radiological protection and radiotherapy. In the field of radiological protection, it is assumed that members of the population subjected to protection are equally sensitive to adverse health effects related to radiation exposure, which is the limitation of existing radiological protection practices [1]. An accurate and robust method to evaluate the radiosensitivity of individuals or subgroups is needed to improve the radiological protection under consideration of the various radiosensitivity among members in a protection group [1]. Likewise, in radiotherapy, patients would not indicate identical responses under the same physical dose due to the interpatient heterogenity of radiosensitivity [2]. Therefore, prediction of radiosensitivity would be beneficial for determining patient-specific treatment methods, doses, fractionation schedules, corresponding clinical outcomes, and reducing possible side effects of radiotherapy [2–4].

Researchers have revealed that the sensitivity to radiation damage of tumor cells depends on the type, characteristics, and genetic level of the tumor cells [2,5,6]. Moreover, advances in gene expression profiling technology have allowed the analysis of the expression level of numerous kinds of genes or proteins in not only normal tissues but also in tumor cells [7]. Several recent studies have shown that quantitative analysis of *in vitro* radiosensitivity based on gene expression profiling can be carried out, and they also have suggested models that can predict intrinsic radiosensitivity from gene expression data [4,6,8–11]. These studies improved our understanding of the relationship between gene expression and radiosensitivity. However, further discussion and research are still needed to establish a robust paradigm for predicting radiosensitivity [2,6,12].

Meanwhile, as a novel decision-making methodology, deep learning has recently emerged as a major tool for classification and prediction. The deep learning model updates itself using the hidden

* Corresponding author. Department of Nuclear Engineering, Hanyang University, 222 Wangsimni-ro, Seongdong-gu, Seoul, 04763, Republic of Korea
E-mail address: ychung@hanyang.ac.kr (Y. Chung).

https://doi.org/10.1016/j.net.2021.10.020
1738-5733/© 2021 Korean Nuclear Society, Published by Elsevier Korea LLC. This is an open access article under the CC BY-NC-ND license (http://creativecommons.org/licenses/by-nc-nd/4.0/).

relationships between the given data, which clearly exists but is hard to represent numerically. The gene expression data for radiosensitivity prediction is comprised of the expression value of numerous genes and proteins, which is very hard to measure the relationship between themselves or to the radiosensitivity using the conventional statistical methods. Even in previous studies dealing with gene expression, they extracted features by themselves with statistical methods such as significance analysis of microarray (SAM), unable to analyze the entire data according to these characteristics. However, with the high-level nonlinear feature learning and data processing of deep learning which overwhelms the conventional methods, we expected that the *in vitro* radiosensitivity prediction would be way more effective.

Therefore, in this study, we aimed to investigate the feasibility of *in vitro* radiosensitivity prediction using gene expression profiling data based on previously established deep learning modalities. Moreover, by comparing the performance of the prediction with the results of previous studies, we demonstrated the applicability and potential power of using deep learning technology to predict *in vitro* radiosensitivity from gene expression.

2. Materials and methods

2.1. Radiation response

Since the clonogenic cell surviving fraction of cells at an absorbed dose of 2 Gy (SF2) is widely used as a measurement of *in vitro* radiosensitivity, we also selected SF2 as an indicator of radiosensitivity in this study. The measured (true) SF2 values used in this study were obtained from previous publications [11,13].

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

The NCI-60 panel contains 60 cancer cell lines representing nine types of tumors. It was established by the US National Cancer Institute in the 1980s for *in vitro* drug screening [8]. The NCI-60 panel is now a valuable research resource, considering the continued use of this panel for investigations of radiation response analysis [11,14–16]. In this study, this panel was used as a platform representing multiple cancer cell lines to evaluate the performance of the radiosensitivity prediction model.

2.3. 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 [17]) 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 radiosensitivity. Excluding the melanoma cell line MDA-N, which was shown to be “not available” from the NCI-60, duplicated or triplicated 174 samples of the remaining 59 cell lines were used as inputs in the radiosensitivity prediction model.

2.4. Radiosensitivity prediction modeling

The deep learning-based radiosensitivity prediction model is based on the architecture of a convolutional neural network (CNN), which comprises two distinct components: a convolutional layer and a fully connected (FC) layer. A convolutional layer is a type of neural network that only connects nodes within a certain range, which leads to three distinct advantages: inherently prevent overfitting, effective use of calculation resources, and training with a relatively small amount of data [18]. In our study, we are in a situation where we are prone to have the curse of dimensionality, since the gene expression we are trying to deal with has 54,675 dimensions, compared to the only 174 data we have. Therefore, the CNN was selected to reduce the curse of dimensionality as much as possible and to provide sufficient learning with a relatively small amount of data. Since the gene expression is a one-dimensional vector, high-level feature vectors were extracted from gene expression using a one-dimensional convolutional layer with average pooling and no padding.

After convolutional layers, radiosensitivity is predicted via the FC layers with residual skip-connection [19]. This FC layer utilizes a skip connection designed to make calculated gradients propagate over several hidden layers along with the gradient descent algorithm, allowing the deep learning model to be constructed more deeply [19,20]. The residual block is applied by skipping each layer one by one. The overall structure of the prediction model is presented in Table 1.

For both components of the CNN model, a leaky rectified linear unit (leaky-ReLU) activation function was applied [21]. L2 regularization and dropout were also used at rates of 0.001 and 0.4 at the end of every hidden 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 [18,22–24].

2.5. Training and validation of the prediction model

The k-fold cross-validation method divides the entire dataset into k sub-datasets and uses each dataset in turn as a test set with the remaining k-1 datasets used as a training set to test the model. With the k-fold cross-validation, validation of the entire dataset is possible, maintaining the model bias and variance with the appropriate level. To prevent overfitting to data of a certain cancer cell line and to ensure that the data are correctly stratified, the folds of cross-validation were constructed such that if the first sample of a specific cell line was included in the K-th fold as a validation set, the rest of the samples of that cell line were excluded on that fold, and used as a training set for the fold, as shown in Table 2. We used the k as 6 since the entire number of the data samples was $174 = 6 \times 29$. 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. The hyperparameters of the model, such as learning rate, batch sizes, or

Table 1
Summary of the radiosensitivity prediction model structure.

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

Abbreviation: ReLU, Rectified Linear Unit; FC, Fully Connected.

Table 2
Data stratification of the 6-fold cross-validation.

Tissue of Origin	The number of samples for each fold						
	1st fold	2nd fold	3rd fold	4th fold	5th fold	6th fold	Total
Leukemia	3	3	3	3	3	3	18
NSCLC	5	5	4	4	4	4	26
Colon	3	3	4	4	4	3	21
CNS	3	3	3	3	3	3	18
Melanoma	5	4	4	4	4	5	26
Ovarian	3	4	4	4	3	3	21
Renal	4	4	4	3	4	4	23
Prostate	1	1	1	1	1	1	6
Breast	2	2	2	3	3	3	15
Total	29	29	29	29	29	29	174

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

kernel size were optimized through random searching using 1st fold as a validation set.

2.6. Measurement of model performance

The performance of the radiosensitivity prediction model was evaluated based on the calculation of the root mean squared deviation (RMSD). RMSD is defined as

$$\mathrm{RMSD} = \sqrt{\frac{\sum_{i=1}^{T}(\widehat{Y}_t - Y_t)^2}{T}}$$

where $\widehat{Y}_t$ represents the 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 was defined as the deviation between the measured and predicted SF2. The relative error was calculated as the absolute error divided by the measured SF2, as shown below.

$$Absolute\ error_{sample} = |predicted\ SF2_{sample} - \ measured\ SF2_{sample}|$$

$$Relative\ error_{sample} = \frac{absolute\ error_{sample}}{measured\ SF2_{sample}}$$

To evaluate the performance of the 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 [6]. However, it tended to be overly strict in cases with a low value of measured SF2. Therefore, we considered as a 'correct prediction' if either 1) the absolute error of the survival fraction is less than 1% or 2) the relative error between the measured and predicted SF2 is less than 10%, and an 'incorrect prediction' if the prediction cannot meet the criteria of the 'correct prediction'.

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.

2.7. Validation of the prediction

Additional validation through leave-one-out cross-validation (LOOCV) was conducted by using the incorrectly predicted sample as an independent test set and all the other samples as a broader training set. Through this additional LOOCV, we could determine whether the error was from the bias of the folded cross-validation, or the entire dataset was insufficient to predict the sample correctly. If the prediction was successful in the LOOCV, 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, which indicates that the entire dataset we have could not provide sufficient information to predict the data correctly.

2.8. 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 measured SF2 and predicted SF2. Statistical analysis was performed using GraphPad Prism version 7.03 for Windows, GraphPad Software, San Diego, California USA, www.graphpad.com.

3. Results

3.1. Development of a deep learning-based radiosensitivity prediction model

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

3.2. Performance of model and validation of prediction

Table 3 shows the average and standard deviation of the predicted radiosensitivity in five rounds of the 6-fold cross-validation. The training time per epoch was about 1.1 s/epoch with our system, and we stopped the training in 100,000 epochs with a mini-batch size of 29. 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%, 2–10%, and 10% or more, respectively. The model correctly predicted 171 samples out of the 174 samples, marking the initial prediction accuracy of the model as 98.28%. Three samples (red points in Fig. 3; one each from the cell lines MOLT-4, MDA-MB-435, and HL-60) with large error (527.59%,

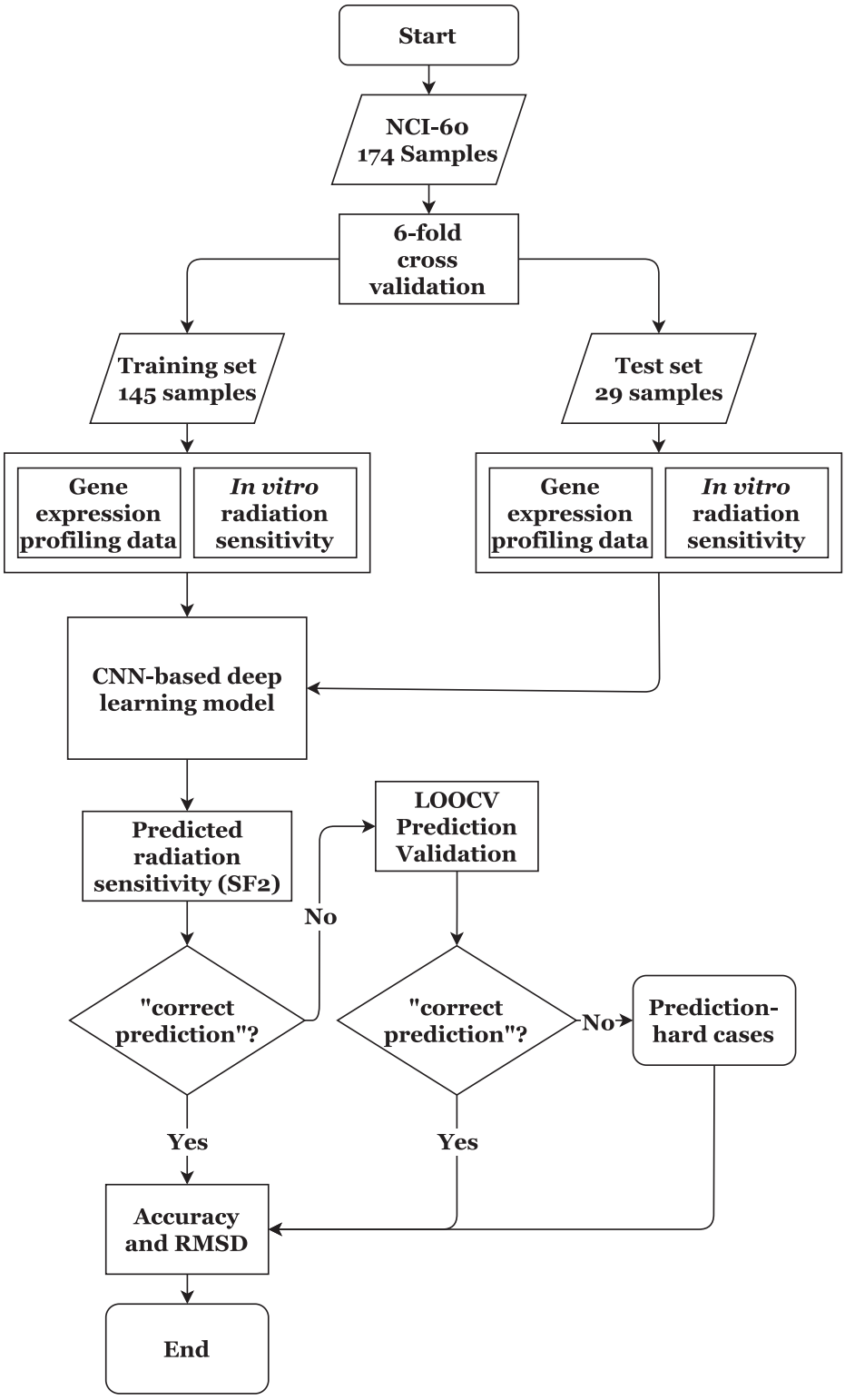

Fig. 1. Overall flowchart of the radiosensitivity prediction model.

129.11%, and 72.88% of relative error, respectively) were subjected to LOOCV of the 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.

As shown in Fig. 3, the predicted SF2 and the measured SF2 represented a distinct linear correlation, indicating that the model successfully predicted the radiosensitivity 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,

Table 3
Average, SD, relative error, absolute error, and prediction results of predicted SF2.

Cell lines	Tissue of Origin	Measured SF2	Predicted SF2 (average ± SD)	Relative Error (%)	Absolute Error	Prediction
CCRF-CEM	Leukemia	0.185	0.183 ± 0.012	0.842	0.001558	correct
			0.182 ± 0.013	1.715	0.003172	correct
			0.181 ± 0.007	2.013	0.003724	correct
HL-60	Leukemia	0.315	0.312 ± 0.013	1.011	0.003184	correct
			0.313 ± 0.010	0.512	0.001612	correct
			0.085 ± 0.013	72.879	0.229570	incorrect[a]
			0.301 ± 0.002	*4.542*	*0.014306*	*correct*[b,c]
K-562	Leukemia	0.050	0.049 ± 0.002	1.456	0.000728	correct
			0.060 ± 0.017	19.388	0.009694	correct
			0.052 ± 0.003	4.132	0.002066	correct
MOLT-4	Leukemia	0.050	0.051 ± 0.005	1.328	0.000664	correct
			0.050 ± 0.004	0.544	0.000272	correct
			0.314 ± 0.006	527.59	0.263794	incorrect[a]
			0.302 ± 0.001	*504.27*	*0.252136*	*incorrect*[b,d]
RPMI-8266	Leukemia	0.100	0.097 ± 0.005	3.258	0.003258	correct
			0.099 ± 0.007	1.166	0.001166	correct
			0.101 ± 0.004	1.246	0.001246	correct
SR	Leukemia	0.070	0.069 ± 0.008	0.923	0.000646	correct
			0.070 ± 0.003	0.149	0.000104	correct
			0.071 ± 0.005	1.103	0.000772	correct
A549	NSCLC	0.610	0.617 ± 0.009	1.130	0.006896	correct
			0.606 ± 0.006	0.611	0.003726	correct
			0.619 ± 0.021	1.446	0.008818	correct
EKVX	NSCLC	0.700	0.695 ± 0.002	0.677	0.004738	correct
			0.697 ± 0.004	0.487	0.003412	correct
			0.692 ± 0.004	1.201	0.008410	correct
HOP-62	NSCLC	0.164	0.170 ± 0.008	3.688	0.006048	correct
			0.163 ± 0.009	0.762	0.001250	correct
			0.178 ± 0.002	8.287	0.013590	correct
HOP-92	NSCLC	0.430	0.435 ± 0.005	1.054	0.004534	correct
			0.421 ± 0.005	2.099	0.009024	correct
			0.427 ± 0.007	0.650	0.002794	correct
NCI−H226	NSCLC	0.630	0.627 ± 0.013	0.428	0.002694	correct
			0.644 ± 0.018	2.147	0.013524	correct
NCI−H23	NSCLC	0.086	0.085 ± 0.004	1.172	0.001008	correct
			0.086 ± 0.006	0.144	0.000124	correct
			0.085 ± 0.005	1.651	0.001420	correct
NCI−H322 M	NSCLC	0.650	0.650 ± 0.010	0.044	0.000288	correct
			0.641 ± 0.012	1.401	0.009106	correct
			0.624 ± 0.012	3.936	0.025586	correct
NCI−H460	NSCLC	0.840	0.838 ± 0.010	0.186	0.001560	correct
			0.821 ± 0.022	2.233	0.018754	correct
			0.847 ± 0.024	0.816	0.006856	correct
NCI−H522	NSCLC	0.430	0.428 ± 0.006	0.350	0.001504	correct
			0.436 ± 0.012	1.420	0.006106	correct
			0.436 ± 0.017	1.369	0.005888	correct
COLO 205	Colon	0.690	0.686 ± 0.009	0.536	0.003696	correct
			0.693 ± 0.015	0.438	0.003024	correct
			0.683 ± 0.009	1.009	0.006964	correct
HCC-2998	Colon	0.440	0.442 ± 0.022	0.429	0.001886	correct
			0.439 ± 0.009	0.118	0.000518	correct
			0.433 ± 0.009	1.519	0.006682	correct
HCT-116	Colon	0.380	0.380 ± 0.012	0.109	0.000414	correct
			0.385 ± 0.011	1.425	0.005416	correct
			0.383 ± 0.008	0.891	0.003386	correct
HCT-15	Colon	0.400	0.405 ± 0.010	1.182	0.004726	correct
			0.398 ± 0.012	0.545	0.002178	correct
			0.410 ± 0.013	2.449	0.009794	correct
HT29	Colon	0.790	0.790 ± 0.010	0.025	0.000167	correct
			0.784 ± 0.010	0.743	0.005866	correct
			0.797 ± 0.011	0.948	0.007486	correct
KM12	Colon	0.420	0.428 ± 0.008	1.837	0.007714	correct
			0.421 ± 0.006	0.287	0.001204	correct
			0.425 ± 0.009	1.178	0.004948	correct
SW-620	Colon	0.620	0.616 ± 0.003	0.575	0.003568	correct
			0.611 ± 0.008	1.434	0.008892	correct
			0.607 ± 0.023	2.173	0.013472	correct
SF-268	CNS	0.450	0.451 ± 0.008	0.273	0.001230	correct
			0.446 ± 0.011	0.963	0.004334	correct
			0.447 ± 0.004	0.762	0.003428	correct
SF-295	CNS	0.730	0.717 ± 0.013	1.803	0.013164	correct
			0.728 ± 0.008	0.311	0.002270	correct
			0.734 ± 0.019	0.578	0.004216	correct

(*continued on next page*)

Table 3 (*continued*)

Cell lines	Tissue of Origin	Measured SF2	Predicted SF2 (average ± SD)	Relative Error (%)	Absolute Error	Prediction
SF-539	CNS	0.820	0.811 ± 0.017	1.102	0.009036	correct
			0.812 ± 0.014	0.940	0.007706	correct
			0.823 ± 0.009	0.332	0.002726	correct
SNB-19	CNS	0.430	0.439 ± 0.008	2.013	0.008656	correct
			0.427 ± 0.004	0.719	0.003092	correct
			0.437 ± 0.015	1.711	0.007358	correct
SNB-75	CNS	0.550	0.553 ± 0.009	0.515	0.002834	correct
			0.548 ± 0.016	0.324	0.001784	correct
			0.554 ± 0.012	0.644	0.003542	correct
U251	CNS	0.570	0.568 ± 0.017	0.379	0.002162	correct
			0.571 ± 0.011	0.229	0.001304	correct
			0.573 ± 0.008	0.601	0.003424	correct
LOX IMVI	Melanoma	0.680	0.687 ± 0.006	1.014	0.006894	correct
			0.680 ± 0.005	0.071	0.000480	correct
			0.681 ± 0.007	0.082	0.000556	correct
MALME-3M	Melanoma	0.800	0.800 ± 0.007	0.038	0.000306	correct
			0.789 ± 0.020	1.346	0.010764	correct
			0.792 ± 0.006	0.996	0.007970	correct
M14	Melanoma	0.420	0.440 ± 0.007	4.665	0.019594	correct
			0.427 ± 0.007	1.780	0.007476	correct
			0.432 ± 0.010	2.794	0.011736	correct
MDA-MB-435	Melanoma	0.179	0.186 ± 0.012	4.120	0.007374	correct
			0.173 ± 0.004	3.477	0.006224	correct
			0.410 ± 0.014	129.11	0.231102	incorrect[a]
			0.362 ± 0.005	*102.34*	*0.183192*	*incorrect*[b,d]
SK-MEL-2	Melanoma	0.660	0.663 ± 0.010	0.478	0.003156	correct
			0.667 ± 0.006	0.993	0.006552	correct
			0.647 ± 0.013	1.999	0.013196	correct
SK-MEL-28	Melanoma	0.740	0.736 ± 0.009	0.522	0.003862	correct
			0.723 ± 0.013	2.296	0.016990	correct
SK-MEL-5	Melanoma	0.720	0.729 ± 0.014	1.272	0.009156	correct
			0.711 ± 0.015	1.219	0.008774	correct
			0.723 ± 0.009	0.447	0.003216	correct
UACC-257	Melanoma	0.480	0.487 ± 0.006	1.372	0.006586	correct
			0.478 ± 0.006	0.407	0.001952	correct
			0.476 ± 0.012	0.820	0.003938	correct
UACC-62	Melanoma	0.520	0.515 ± 0.005	1.053	0.005474	correct
			0.521 ± 0.015	0.178	0.000928	correct
			0.528 ± 0.015	1.523	0.007920	correct
IGR-OV1	Ovarian	0.390	0.391 ± 0.010	0.233	0.000908	correct
			0.385 ± 0.004	1.193	0.004654	correct
			0.405 ± 0.009	3.836	0.014960	correct
OVCAR-3	Ovarian	0.550	0.549 ± 0.005	0.243	0.001338	correct
			0.546 ± 0.010	0.790	0.004346	correct
			0.542 ± 0.013	1.420	0.007812	correct
OVCAR-4	Ovarian	0.290	0.296 ± 0.006	2.234	0.006478	correct
			0.305 ± 0.011	5.329	0.015454	correct
			0.290 ± 0.016	0.121	0.000350	correct
OVCAR-5	Ovarian	0.408	0.407 ± 0.011	0.286	0.001168	correct
			0.406 ± 0.008	0.577	0.002356	correct
			0.405 ± 0.015	0.719	0.002932	correct
OVCAR-8	Ovarian	0.600	0.599 ± 0.014	0.137	0.000820	correct
			0.597 ± 0.010	0.490	0.002938	correct
			0.599 ± 0.018	0.118	0.000706	correct
NCI/ADR-RES	Ovarian	0.560	0.588 ± 0.011	4.993	0.027958	correct
			0.571 ± 0.011	1.956	0.010956	correct
			0.580 ± 0.008	3.650	0.020438	correct
SK-OV-3	Ovarian	0.900	0.881 ± 0.017	2.132	0.019188	correct
			0.887 ± 0.020	1.404	0.012640	correct
			0.877 ± 0.014	2.566	0.023092	correct
786-O	Renal	0.660	0.647 ± 0.015	2.018	0.013318	correct
			0.661 ± 0.005	0.163	0.001076	correct
			0.655 ± 0.005	0.760	0.005016	correct
A498	Renal	0.610	0.605 ± 0.012	0.805	0.004912	correct
			0.620 ± 0.011	1.708	0.010416	correct
			0.607 ± 0.019	0.457	0.002788	correct
ACHN	Renal	0.720	0.696 ± 0.007	3.266	0.023512	correct
			0.715 ± 0.008	0.633	0.004554	correct
			0.667 ± 0.008	7.302	0.052572	correct
CAKI-1	Renal	0.370	0.365 ± 0.015	1.280	0.004736	correct
			0.371 ± 0.028	0.315	0.001164	correct
RXF 393	Renal	0.670	0.674 ± 0.006	0.606	0.004058	correct
			0.670 ± 0.009	0.012	0.000080	correct
			0.667 ± 0.005	0.448	0.003002	correct
SN12C	Renal	0.620	0.625 ± 0.010	0.778	0.004826	correct

Table 3 (*continued*)

Cell lines	Tissue of Origin	Measured SF2	Predicted SF2 (average ± SD)	Relative Error (%)	Absolute Error	Prediction
			0.624 ± 0.002	0.640	0.003970	correct
			0.616 ± 0.011	0.630	0.003906	correct
TK-10	Renal	0.520	0.520 ± 0.007	0.078	0.000408	correct
			0.525 ± 0.009	0.904	0.004700	correct
			0.522 ± 0.009	0.333	0.001732	correct
UO-31	Renal	0.620	0.624 ± 0.004	0.602	0.003732	correct
			0.621 ± 0.004	0.082	0.000506	correct
			0.625 ± 0.004	0.799	0.004954	correct
PC-3	Prostate	0.484	0.474 ± 0.013	2.036	0.009854	correct
			0.490 ± 0.009	1.175	0.005688	correct
			0.487 ± 0.007	0.679	0.003286	correct
DU-145	Prostate	0.520	0.520 ± 0.008	0.088	0.000460	correct
			0.517 ± 0.011	0.644	0.003350	correct
			0.517 ± 0.011	0.560	0.002910	correct
MCF7	Breast	0.576	0.566 ± 0.010	1.769	0.010190	correct
			0.565 ± 0.007	1.880	0.010830	correct
			0.574 ± 0.010	0.406	0.002336	correct
MDA-MB-231	Breast	0.630	0.635 ± 0.012	0.839	0.005286	correct
			0.637 ± 0.009	1.141	0.007186	correct
			0.626 ± 0.013	0.642	0.004042	correct
HS 578T	Breast	0.790	0.791 ± 0.017	0.142	0.001122	correct
			0.787 ± 0.003	0.399	0.003156	correct
			0.800 ± 0.005	1.208	0.009544	correct
BT-549	Breast	0.630	0.627 ± 0.008	0.551	0.003474	correct
			0.635 ± 0.015	0.731	0.004606	correct
			0.619 ± 0.002	1.742	0.010974	correct
T-47D	Breast	0.520	0.524 ± 0.009	0.743	0.003866	correct
			0.523 ± 0.007	0.645	0.003354	correct
			0.528 ± 0.010	1.522	0.007912	correct

Abbreviation: NSCLC, Non-Small Cell Lung Cancer; CNS, Central Nervous System; LOOCV, Leave-One-Out Cross-Validation; SD, Standard Deviation; SF2, Survival Fraction at 2 Gy.

[a] Subjected to prediction validation experiment due to initial classification of incorrect prediction.

[b] Result of the LOOCV of prediction validation, presented in the 4th row of the HL-60, MOLT-4, and MDAMB-435 in the italic text.

[c] Changed into 'correct prediction' in the LOOCV of prediction validation.

[d] Classified as 'prediction-hard' cases due to the result of LOOCV of prediction validation.

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 LOOCV of prediction validation was 98.85% (172 out of 174 were correct), and the RMSD was 0.0252 with prediction-hard cases and 0.00867 without the prediction-hard cases.

4. Discussion

Deep learning is an emerging research field that gained prominence as hardware advanced. It is widely used for decision-making, prediction, and classification. In this study, we proposed the feasibility of deep learning as a novel methodology of *in vitro* radiosensitivity prediction by developing a deep learning-based *in vitro* radiosensitivity 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* radiosensitivity prediction.

In their analyses of model accuracy, Torres-Roca et al. and Zhang et al. who similarly tried to predict the radiosensitivity of the NCI-60 cancer cell lines both set the criteria of the 'correct prediction' when the predicted SF2s were within 10% of the measured values [6,11]. 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 similar but more reasonable criteria, 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. [6,11]. These results indicate that complex biological nonlinear genetic interactions influencing the radiosensitivity of a cancer cell lines are likely to be well represented by deep learning.

Three samples with large errors, one each from the cell lines MOLT-4, MDA-MB-435, and HL-60, were subjected to LOOCV of prediction validation because it might not be due to merely deviation 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 LOOCV of prediction validation, the sample from the HL-60 cell line represented a significantly improved prediction result, which indicates that the large error of the HL-60 sample from the initial prediction appears to be due to the high bias of the fold (the fold cannot represent the whole dataset). For the remaining 'prediction-hard' cases, MOLT-4 and MDA-MB-435, we were unable to determine whether it was due to the high bias of the fold, or if there were other possible problems that could not be investigated in this study, such as mislabeling issues. Therefore, further research would be needed to address it. However, what is noteworthy, even if these 'prediction-hard' cases are due to the high bias of the fold, 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.

Since this study is focused on how the deep learning based *in vitro* radiosensitivity prediction methodology is feasible and applicable compared to the other previously published results, we needed whole prediction results of all of the data to compare the result with the other studies, to confirm whether this methodology

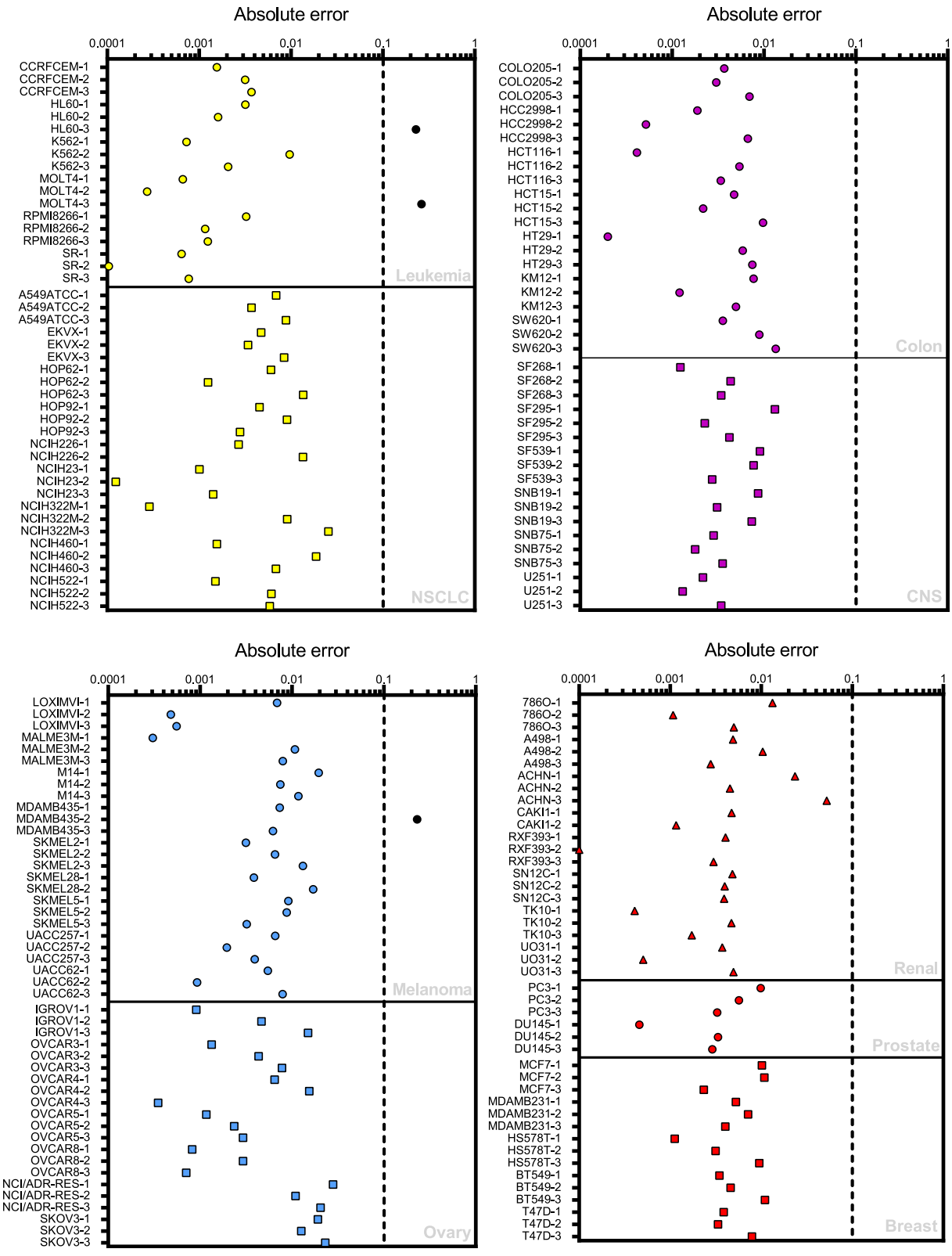

Fig. 2. Plotted absolute error of all predicted triplicated samples of NCI-60 cell lines (log-scale). Dark circles are representing three samples (HL-60, MOLT-4, and MDAMB-435) that were subjected to prediction validation. The vertical dotted line is a threshold for the correct prediction of absolute error. The order and classification of the samples were based on the NCI-60 panel.

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

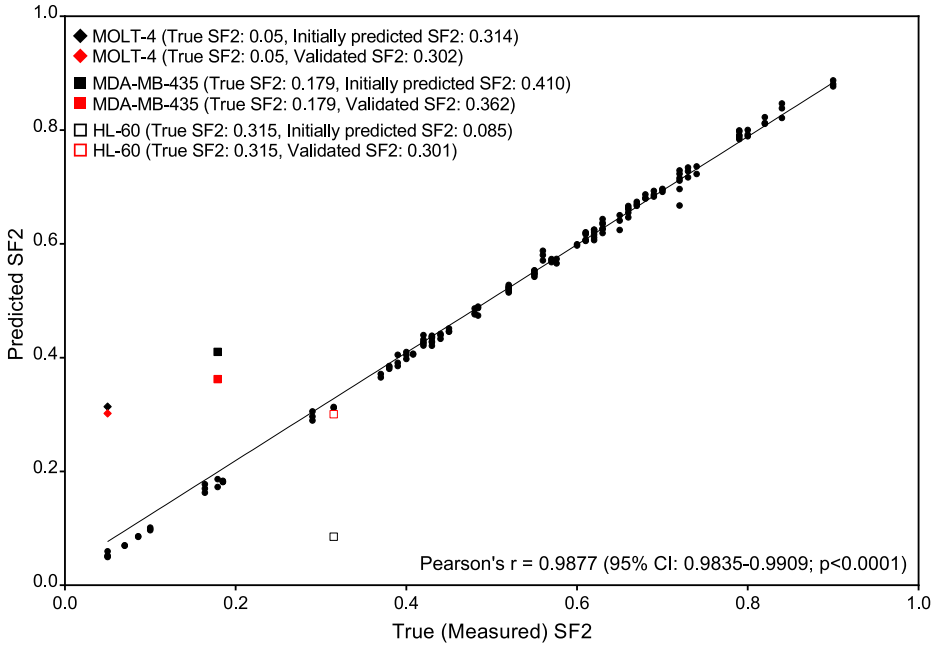

Fig. 3. Linear correlation between true (measured) and predicted survival fraction at 2 Gy. Dark circles are representing the predicted SF2s of every triplicated sample.

is scientifically worth to go further. However, with the original cross-validation which completely separates the train, validation set with the test set, we could not get the whole prediction results of all of the data we have. Under these circumstances, as we confirmed that the authors obtained prediction results of all of the data through cross-validation in previous publications [6,11,28], our prediction results of all data were obtained through the 6-fold cross-validation.

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 [25]. However, in this study, the limited number of cell lines samples with survival data available for training may not have fully demonstrated the overall characteristics of gene expression and therefore may not have fully derived the whole potential of deep learning. Thus, it seems necessary to further boost the performance of this methodology by additional training using a large number of radiosensitivity datasets from not only the NCI-60 cell lines but also the other types of cancer cell lines. Second, the use of classical microarray analysis rather than ribonucleic acid (RNA) sequencing, which is the latest gene expression profiling method, can be considered as one of the limitations of this study. Although the microarray is a little outdated method and is constantly being replaced with RNA sequencing, we used microarray data to demonstrate the feasibility of deep learning aided radiosensitivity prediction through comparison with previous studies. In this perspective, further research is needed regarding the prediction model using RNA sequencing data, rather than the microarray.

Despite these limitations, several improvements in radiosensitivity prediction analyses are expected from this 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 and represent a direct and transparent relationship between the input genes and radiosensitivity, since the input data has undergone the only minimal process [26]. Second, the deep learning model can further learn (trained) from additional data presented after training, which enables deep learning to self-correct and make itself more robust [27]. 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 or machine learning-based model, which leads to more valid and accurate prediction results.

Regarding the scientific validation of the prediction by this method, in deep learning, causes and results are the only information provided. One of their characteristics is that they maintain "black boxes" concerning their internal processes even though they provide good results. Deep learning used in this study is also very useful for its ability to predict radiosensitivity with high accuracy, but it is not providing any scientific explanation for how such predictions are made by far. Therefore, further research regarding the interpretability of the radiosensitivity prediction using deep learning would be needed as many recent studies do, to identify the genetic mechanisms of how organisms react to radiation exposure [28].

5. Conclusions

In summary, this study successfully demonstrated the feasibility of a deep learning-based *in vitro* radiosensitivity prediction using gene expression profiling data. We established a CNN-based deep learning model and compared the prediction performance of our method with other previously published methods with the same data. With additional research and external validation, this method could be expanded ultimately to the *in vivo* radiosensitivity prediction.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

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).

References

[1] S.D. Bouffler, Evidence for variation in human radiosensitivity and its potential impact on radiological protection, Ann. ICRP 45 (2016) 280–289.
[2] J.G. Scott, G. Sedor, P. Ellsworth, J.A. Scarborough, K.A. Ahmed, D.E. Oliver, et al., Pan-cancer prediction of radiotherapy benefit using genomic-adjusted radiation dose (GARD): a cohort-based pooled analysis, Lancet Oncol. 22 (2021) 1221–1229.
[3] D.G. Hirst, T. Robson, Molecular biology: the key to personalised treatment in radiation oncology? Br. J. Radiol. 83 (2010) 723–728.
[4] H.S. Kim, S.C. Kim, S.J. Kim, C.H. Park, H.C. Jeung, Y.B. Kim, et al., Identification of a radiosensitivity signature using integrative metaanalysis of published microarray data for NCI-60 cancer cells, BMC Genom. 13 (2012) 348.
[5] A.C. Begg, F.A. Stewart, C. Vens, Strategies to improve radiotherapy with targeted drugs, Nat. Rev. Cancer 11 (2011) 239–253.
[6] J.F. Torres-Roca, S. Eschrich, H. Zhao, G. Bloom, J. Sung, S. McCarthy, et al., Prediction of radiosensitivity using a gene expression classifier, Cancer Res. 65 (2005) 7169–7176.
[7] S. Ramaswamy, T.R. Golub, DNA microarrays in clinical oncology, J. Clin. Oncol. 20 (2002) 1932–1941.
[8] S.A. Amundson, K.T. Do, L.C. Vinikoor, R.A. Lee, C.A. Koch-Paiz, J. Ahn, 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 (2008) 415–424.
[9] J. Khan, J.S. Wei, M. Ringner, L.H. Saal, M. Ladanyi, F. Westermann, et al., Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks, Nat. Med. 7 (2001) 673–679.
[10] K. Ogawa, S. Murayama, M. Mori, Predicting the tumor response to radiotherapy using microarray analysis (Review), Oncol. Rep. 18 (2007) 1243–1248.
[11] C. Zhang, L. Girard, A. Das, S. Chen, G. Zheng, K. Song, Nonlinear quantitative radiation sensitivity prediction model based on NCI-60 cancer cell lines, Sci. World J. 2014 (2014) 903602.
[12] L.J. Peters, W.A. Brock, J.D. Chapman, G. Wilson, Predictive assays of tumor radiocurability, Am. J. Clin. Oncol. 11 (1988) 275–287.
[13] S. Eschrich, H. Zhang, H. Zhao, D. Boulware, J.H. Lee, G. Bloom, et al., Systems biology modeling of the radiosensitivity network: a biomarker discovery platform, Int. J. Radiat. Oncol. Biol. Phys. 75 (2009) 497–505.
[14] M.R. Boyd, K.D. Paull, Some practical considerations and applications of the national cancer institute *in vitro* anticancer drug discovery screen, Drug Dev. Res. 34 (1995) 91–109.
[15] M. Burkard, Integrating the NCI-60 data with “omics” for drug discovery, Drug Dev. Res. 73 (2012).
[16] R.H. Shoemaker, The NCI60 human tumour cell line anticancer drug screen, Nat. Rev. Cancer 6 (2006) 813–823.
[17] T.D. Pfister, W.C. Reinhold, K. Agama, S. Gupta, S.A. Khin, R.J. Kinders, 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 Therapeut. 8 (2009) 1878–1884.
[18] A. Krizhevsky, I. Sutskever, G.E. Hinton, Imagenet classification with deep convolutional neural networks, Adv. Neural Inf. Process. Syst. (2012) 1097–1105.
[19] K. He, X. Zhang, S. Ren, J. Sun, Deep residual learning for image recognition, Proceedings of the IEEE conference on computer vision and pattern recognition (2016) 770–778.
[20] X. Glorot, Y. Bengio, Understanding the difficulty of training deep feedforward neural networks, in: T. Yee Whye, T. Mike (Eds.), Proceedings of the Thirteenth International Conference on Artificial Intelligence and Statistics. Proceedings of Machine Learning Research, PMLR, 2010, pp. 249–256.
[21] A.L. Maas, A.Y. Hannun, A.Y. Ng, Rectifier nonlinearities improve neural network acoustic models, Proc icml: Cités (2013) 3.
[22] G.E. Dahl, T.N. Sainath, G.E. Hinton, Improving deep neural networks for LVCSR using rectified linear units and dropout, in: IEEE International Conference on Acoustics, Speech and Signal Processing 2013, 2013, pp. 8609–8613.
[23] G.E. Hinton, N. Srivastava, A. Krizhevsky, I. Sutskever, R.R. Salakhutdinov, Improving Neural Networks by Preventing Co-adaptation of Feature Detectors, 2012 arXiv preprint arXiv:12070580.
[24] V. Nair, G.E. Hinton, Rectified linear units improve restricted Boltzmann machines, in: Proceedings of the 27th International Conference on Machine Learning (ICML-10), 2010, pp. 807–814.
[25] D. Silver, A. Huang, C.J. Maddison, A. Guez, L. Sifre, G. van den Driessche, et al., Mastering the game of Go with deep neural networks and tree search, Nature 529 (2016) 484–489.
[26] P. Meyer, V. Noblet, C. Mazzara, A. Lallement, Survey on deep learning for radiotherapy, Comput. Biol. Med. 98 (2018) 126–146.
[27] S.J. Pan, Q. Yang, A survey on transfer learning, IEEE Trans. Knowl. Data Eng. 22 (2010) 1345–1359.
[28] J.H. Oh, W. Choi, E. Ko, M. Kang, A. Tannenbaum, J.O. Deasy, PathCNN: interpretable convolutional neural networks for survival prediction and pathway analysis applied to glioblastoma, Bioinformatics 37 (2021) i443–i450.