DISTANCE BETWEEN SPECIES BY CONFUSION OPERATORS OF MULTI-CLASS CLASSIFIERS

Nicolae TARBĂ[1], Ionela N. IRIMESCU[2], Ana M. PLEAVĂ[3], Eugen N. SCARLAT[4], Mona MIHĂILESCU[5]

We introduce a method to evaluate the similarities between classes of objects based on the confusion matrices coming from the multi-class machine learning (ML) predictors that operate in the vector space generated by the classes. Operation in the confusion domain allows us to use the distance as evaluation metric of the closeness of the relevant properties carried by the classes. We applied the method to semi-quantitatively evaluate the effects of X-ray and carbon ion radiation on cultured SW1353 chondrosarcoma cells with respect to the non-irradiated reference. We prove that it is possible to draw useful information on the strength of cells nuclei changes.

Keywords: machine learning classifiers, features, confusion matrix, confusion domain, confusion operator, cross-confusion coefficient, distance.

1. Introduction

Supervised machine learning (ML) classifiers are largely used to sort cells species or tissues using features drawn from optical microscopy images [1-4]. Since hyperspectral microscopy images span across a wide spectrum ranging between 400 and 1000nm in each pixel, the properties extracted from the hyperspectral images (HSI) were used in association with ML algorithms to obtain effective and robust classifiers for accurate cell sorting [5-7]. For example, in the case of irradiated/non-irradiated cells, the classification was done based on spatial features taken on spectral sub-intervals [8]. In multiple-category scenarios, the classifiers compute the probabilities of belonging to each category, such that the greatest likelihood gives the category the cell is predicted to belong to.

[1] Phd stud. Computer Sciences Doctoral School, Computer Science and Engineering Department, Faculty of Automatic Control and Computers, National University of Science and Technology POLITEHNICA Bucharest, Romania, nicolae.tarba@upb.ro
[2] Phd stud., Applied Sciences Doctoral School, National University of Science and Technology POLITEHNICA Bucharest, Romania, ionela.irimescu@stud.fsa.upb.ro
[3] Res., CAMPUS Research Center, National University of Science and Technology POLITEHNICA Bucharest, Romania.
[4] Assoc. Prof., Physics Dept., National University of Science and Technology POLITEHNICA Bucharest, Romania.
[5] Prof., Physics Dept, Research Center for Applied Sciences in Engineering, National University of Science and Technology POLITEHNICA Bucharest, Romania.

Besides accurate sorting, we prove that classifiers carry information which reveal the similarity among the cells species. The extent to which the species are similar is given by the distance metric in a vector space mapped onto itself by the confusion operators generated by the confusion matrices of the ML classifiers. Therefore, the coordinates in vector space are defined not by the raw probabilities calculated by the classifiers in the sorting process, but by the weights of the classification errors in the confusion matrices.

In the application presented here, the features the ML classifiers work with are extracted from the HSIs of the investigated samples. General features used so far in the analysis of various types of images (microscopic, macroscopic, satellite) are based on shape, texture, edge and corner detection, brightness, or spectral response. In most cases the initial number of features is large enough to perform the classification with no error, i.e., with accuracy Acc=1. When reducing the number of features, the accuracy falls below unity Acc<1 and the process enters into the so-called "confusion domain" where the distance could be considered an indicator of the differences/similarities between the relevant features that characterize one species.

After the presentation of the method, we describe its application in the case of three cells species: irradiated with carbon ions (CI), with X-rays (XR), and non-irradiated (REF). Since there is a large debate on the effectiveness of XR therapy versus CI therapy [9, 10], every piece of information related to their comparative properties becomes important. Particularly the optical methods for cell characterization gain more and more ground due to high automatization, high processing speeds, and workforce reduction [11].

2. Method

2.1 Multi-class confusion matrices

Table 1 shows the general form of a $N\times N$ confusion matrix corresponding to a ML classifier that sorts N species of objects S_k k=1,…N having $\#S_k$ elements each. $E_{Sk/Sj}$ means the number of elements of the species S_k predicted to belong to the species S_j.

The accuracy metric maintains unity value Acc=1 down to a minimal combination of f_{min} features that still ensure the classification without errors. An additional removal of any of the remaining f_{min} features, $f<f_{min}$, leads to classification errors and Acc<1. Therefore, the last f_{min} features are considered the relevant features, and the domain $f<f_{min}$ is considered the confusion domain. The closer the values of the features of the species, the greater the confusion.

Table 1

Multi-class confusion matrix

		PREDICTED		
		$\mathbf{S_1}$	**…** $\mathbf{S_j}$ **…**	$\mathbf{S_N}$
ACTUAL	$\mathbf{S_1}$ (#S_1)	$E_{S1/S1}$	… $E_{S1/Sj}$ …	$E_{S1/SN}$
	….	…	…	…
	$\mathbf{S_k}$ (#S_k)	$E_{Sk/S1}$	… $E_{Sk/Sj}$ …	$E_{Sk/SN}$
	….	….	…	…
	$\mathbf{S_N}$ (#S_N)	$E_{SN/S1}$	… $E_{SN/Sj}$ …	$E_{SN/SN}$

Inside the confusion domain $f<f_{min}$, the number of possible combinations to select f features from the available ones f_{min} is $C_{f_{min}}^{f}$; hence the number of confusion matrices is the same $C_{f_{min}}^{f}$. For every $f<f_{min}$, an averaged confusion matrix is computed whose elements are the mean values over the number of confusion matrices. The standard errors of the mean values are computed accordingly. Next, the confusion matrix of normalized values is obtained by dividing each entry of a specific row by the sum of the values within it. Consequently, the sums per row become one, and the normalized values denoted with "e" can be interpreted as estimators of the multi-class classification probabilities. The confusion matrix of normalized values is then the same as the confusion matrix of probabilities (see Table 2). The corresponding standard errors of the probabilities result from the standard errors of the mean values by using the error propagation relationships. For $f.\geq f_{min}$ the classification errors are zero, and the matrices have unit values along the main diagonal and zero elsewhere. The notation $e_{Sk/Sj}$ means the probability an element belonging to species S_k to be predicted and classified in the category S_j.

Table 2

Multi-class confusion matrix of probabilities

		PREDICTED		
		$\mathbf{S_1}$	**…** $\mathbf{S_j}$**…**	$\mathbf{S_N}$
ACTUAL	$\mathbf{S_1}$ (#S_1)	$e_{S1/S1}$	…$e_{S1/Sj}$…	$e_{S1/SN}$
	….	…	…	…
	$\mathbf{S_k}$ (#S_k)	$e_{Sk/S1}$	…$e_{Sk/Sj}$…	$e_{Sk/SN}$
	….	….	…	…
	$\mathbf{S_N}$ (#S_N)	$e_{SN/S1}$	…$e_{SN/Sj}$…	$e_{SN/SN}$

Since the sums per row are one and assuming Euclidean metric, the square roots of the elements of the confusion matrix of the probabilities can be considered the coordinates of an unit vector used in classification process.

2.2 The confusion operator and the associated matrix

The confusion operator formally expresses the classification action in the *N*-dimensional space of unit vectors where the species are represented by *N*-dimensional vectors. The operator maps the actual elements of the species into predicted vectors of unit size located in the first hyper-octant. In the supervised classification process all elements that certainly belong to the species S_k are inputs for ML classifiers and are written as line vectors of coordinates 1 on the *k* position and zero elsewhere. The vectors are of the line type such that the representation be consistent with the meanings ACTUAL (on lines)/PREDICTED (on columns) in the confusion matrix. The set $\{S_1, \ldots, S_N\}$ constitutes an orthonormal basis.

For $f<f_{min}$, the elements of the matrix **C** associated with the confusion operator $\mathcal{C}$ are the square roots of the corresponding elements of the confusion matrix of probabilities; for the sake of simplifying the notations, the dependence on f is omitted:

$$\mathbf{C}=\begin{pmatrix} \sqrt{e_{S_1/S_1}} & \cdots & \sqrt{e_{S_1/S_N}} \\ \vdots & \ddots & \vdots \\ \sqrt{e_{S_N/S_1}} & \cdots & \sqrt{e_{S_N/S_N}} \end{pmatrix}, \quad f<f_{min}. \tag{1}$$

The process of classifying an element from S_k is written in the form $\mathcal{C}(S_k)=S_k\cdot\mathbf{C}$, the result of the sorting action being a predicted species represented by a vector of unit size:

$$\mathcal{C}(S_k) = S_k\cdot\mathbf{C} = \sum_{j=1}^{N} \sqrt{e_{S_k/S_j}} \cdot S_j, \qquad \sum_{j=1}^{N} e_{S_k/S_j} = 1. \tag{2}$$

Therefore

$$S_k \xrightarrow{\text{classification}} \mathcal{C}(S_k) = (\sqrt{e_{S_k/S_1}}, \ \ldots, \sqrt{e_{S_k/S_j}}, \ldots, \sqrt{e_{S_k/S_N}}), \tag{3}$$

i.e., the predicted species $\mathcal{C}(S_k)$ is confounded with the actual species S_j with the probability e_{S_k/S_j}.

For two vectors $U=(u_1, \ldots, u_N)$ and $V=(v_1, \ldots, v_N)$, the dot product is:

$$\langle U, V\rangle = U\cdot V^T = u_1 v_1 + \ldots + u_N v_N. \tag{4}$$

The predicted state $\mathcal{C}(S_k)$ expands on the basis $\{S_1, \ldots, S_N\}$, the expansion coefficients being computed by the help of the dot product $\langle \mathcal{C}(S_k), S_j\rangle = \sqrt{e_{S_k/S_j}}$.

2.3 Cross-confusion coefficient and distance

For $f<f_{min}$ features and $S_k \neq S_j$, the cross-confusion coefficient denoted $K(S_k,S_j)$ (f omitted) indicates the extent to which the predicted species $\mathcal{C}_f(S_k)$ is confounded with the species S_j, and is computed as the dot product

$$K(S_k,S_j) = \langle \mathcal{C}_f(S_k), S_j \rangle, \quad k \neq j. \tag{5}$$

The predicted species $\mathcal{C}_f(S_k)$ has properties similar to those of the species S_j to the extent $\langle \mathcal{C}_f(S_k), S_j \rangle$. Similarly, the confusion operator predicts the species S_j as $\mathcal{C}_f(S_j)$ which has properties similar to those of the species S_k to the extent $\langle \mathcal{C}_f(S_j), S_k \rangle$:

$$K(S_j,S_k) = \langle \mathcal{C}_f(S_j), S_k \rangle, \quad k \neq j. \tag{5'}$$

Since $\langle \mathcal{C}_f(S_k), S_j \rangle \neq \langle \mathcal{C}_f(S_j), S_k \rangle$, the two coefficients are not necessarily equal to each other $K(S_k, S_j) \neq K(S_j, S_k)$.

The distance between two states U and V is:

$$d(U,V) = |\, U-V \,| = (\langle U-V, U-V \rangle)^{1/2} = (1- \langle U, V \rangle)^{1/2}. \tag{6}$$

The maximum distance is $\sqrt{2}$. The normalized distance is $D=d/\sqrt{2}$. The normalized distance between a predicted species $\mathcal{C}_f(S_j)$ and an actual species S_k can be expressed with the help of $K(S_k, S_j)$ from Eq.(5):

$$D(\mathcal{C}_f(S_j), S_k) = (1- \langle \mathcal{C}_f(S_j), S_k \rangle)^{1/2} = (1- K(S_j, S_k))^{1/2}. \tag{6'}$$

The bigger the confusion coefficients, the smaller the distances.

3. Application

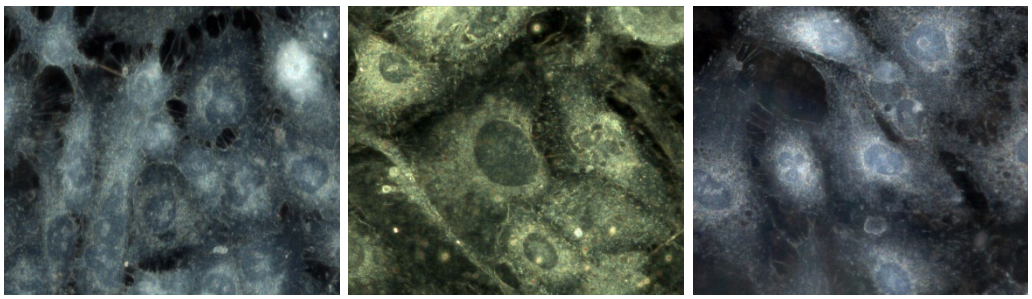
We used Microsoft's ML.NET AutoML [12] for multi-class classification with LogLoss as the optimization metric to explore various models with various hyper-parameters [13-15]. After the training and testing stages, the best models were chosen based on their LogLoss values. Here the ML classifiers use an initial set of 49 optical features drawn from the 1D hyperspectral curve of SW1353 chondrosarcoma cells nuclei. This is different from [8] where the features were extracted from 2D images; besides, here the classifications are multi-class not only binary. The features number is progressively reduced through permutation feature importance (PFI) [16], removing the features that increase LogLoss the least when their values are randomly permuted. We found a minimal combination of 6 features ($f_{min}=6$) for which there is no classification error. Additional elimination of any of the remaining features leads to classification errors and subunit accuracy. The interval $f<6$ is the confusion domain. The number of confusion matrices used to generate the matrices of the confusion operators is shown in Table 3.

Table 3

The number of confusion matrices in the confusion domain, $f_{min}=6$

f	5	4	3	2	1
# confusion matrices	6	15	20	15	6

The samples consist of nuclei from 38 cells of CI species, 52 cells of XR species, and 78 cells of REF species. For irradiated cells the dose was 4Gy. In Fig. 1 are hyperspectral microscopy images of cells from each of the three species. We focus on the confusion operators that confuse the species CI with REF, respectively XR with REF in the confusion domain.



a) b) c)

Fig. 1 Experimental hyperspectral images for a) REF, b) IC, c) XR species.

For three species, the confusion matrices from Eq.(1) reduce to 3×3 matrices (again, for the sake of simplifying the notations, the dependence on f is omitted):

$$\mathbf{C}=\begin{pmatrix} \sqrt{e_{CI/CI}} & \sqrt{e_{CI/XR}} & \sqrt{e_{CI/REF}} \\ \sqrt{e_{XR/CI}} & \sqrt{e_{XR/XR}} & \sqrt{e_{XR/REF}} \\ \sqrt{e_{REF/CI}} & \sqrt{e_{REF/XR}} & \sqrt{e_{REF/REF}} \end{pmatrix}. \tag{7}$$

4. Results

Table 4a shows the matrices of the confusion operators built as explained in Sec.2.2 according to the meaningfulness stated by Eq.(7). The corresponding standard errors for $f\leq 6$ are given in Table 4b. The case $f_{min}=6$ is only to mark the threshold of the confusion domain. For $f_{min}=6$ the accuracy is still unity with no contribution to confusion matrices nor to confusion operators.

Table 4a

The matrices of the confusion operators

$f_{min}=6$	f=5	f=4
$\begin{pmatrix}1.000 & 0.000 & 0.000\\ 0.000 & 1.000 & 0.000\\ 0.000 & 0.000 & 1.000\end{pmatrix}$	$\begin{pmatrix}0.992 & 0.073 & 0.103\\ 0.000 & 0.994 & 0.113\\ 0.000 & 0.065 & 0.998\end{pmatrix}$	$\begin{pmatrix}0.989 & 0.111 & 0.103\\ 0.062 & 0.988 & 0.139\\ 0.058 & 0.109 & 0.992\end{pmatrix}$
f=3	f=2	f=1
$\begin{pmatrix}0.967 & 0.178 & 0.185\\ 0.088 & 0.959 & 0.270\\ 0.119 & 0.188 & 0.975\end{pmatrix}$	$\begin{pmatrix}0.914 & 0.258 & 0.313\\ 0.179 & 0.879 & 0.441\\ 0.163 & 0.241 & 0.957\end{pmatrix}$	$\begin{pmatrix}0.747 & 0.320 & 0.580\\ 0.233 & 0.770 & 0.594\\ 0.191 & 0.300 & 0.936\end{pmatrix}$

Table 4b

The standard errors of the matrices of the confusion operators

f=6	f=5	f=4
$\begin{pmatrix}0.000 & 0.000 & 0.000\\ 0.000 & 0.000 & 0.000\\ 0.000 & 0.000 & 0.000\end{pmatrix}$	$\begin{pmatrix}0.007 & 0.001 & 0.001\\ 0.000 & 0.006 & 0.057\\ 0.000 & 0.021 & 0.091\end{pmatrix}$	$\begin{pmatrix}0.005 & 0.020 & 0.035\\ 0.017 & 0.007 & 0.044\\ 0.029 & 0.029 & 0.005\end{pmatrix}$
f=3	f=2	f=1
$\begin{pmatrix}0.006 & 0.017 & 0.028\\ 0.011 & 0.008 & 0.026\\ 0.024 & 0.021 & 0.007\end{pmatrix}$	$\begin{pmatrix}0.019 & 0.022 & 0.050\\ 0.041 & 0.033 & 0.052\\ 0.029 & 0.022 & 0.009\end{pmatrix}$	$\begin{pmatrix}0.056 & 0.040 & 0.062\\ 0.086 & 0.080 & 0.074\\ 0.020 & 0.050 & 0.019\end{pmatrix}$

4.1 The distances between predicted XR/CI and actual REF

When the classifiers sort the elements of CI and XR species, the results are the predictions $\mathcal{C}_f$(CI) and $\mathcal{C}_f$(XR) computed according to Eq.(3). One can remark that whatever f=5, 4, 3, 2, 1, the following relationships stand (see Eq.(7) and Table 4a, the dependence on f is omitted as before)

$$\sqrt{e_{XR/REF}} > \sqrt{e_{CI/REF}} \quad \text{for all } f\leq 5, \tag{8}$$

meaning that the predictor $\mathcal{C}_f$(XR) confuses REF to a greater extent than $\mathcal{C}_f$(CI) confuses REF:

$$K(\text{XR, REF}) > K(\text{CI, REF}), \tag{8'}$$

equivalent to

$$\langle \mathcal{C}_f(\text{XR}), \text{REF}\rangle > \langle \mathcal{C}_f(\text{CI}), \text{REF}\rangle. \tag{8''}$$

In terms of normalized distances, by using Eq.(6'), one finds (see Table 5)

$$D(\mathcal{C}_f(\text{XR}), \text{REF}) < D(\mathcal{C}_f(\text{CI}), \text{REF}) \text{ for all } f\leq 5. \tag{8'''}$$

Table 5

Normalized distances between predicted XR/CI and actual REF

f	5	4	3	2	1
$D(\mathcal{C}_f(\text{CI}), \text{REF})$	0.994±0.001	0.947±0.018	0.903±0.016	0.829±0.030	0.648±0.048
$D(\mathcal{C}_f(\text{XR}), \text{REF})$	0.942±0.030	0.928±0.024	0.854±0.015	0.747±0.035	0.637±0.058

Consequently the predicted state $\mathcal{C}_f(\text{XR})$ is closer to REF than the predicted state $\mathcal{C}_f(\text{CI})$.

4.2 The distances between predicted REF and actual CI/XR

The classifier now sorts out the non-irradiated REF cells. Intuitively, if $\mathcal{C}_f(\text{XR})$ mimics REF better than $\mathcal{C}_f(\text{CI})$ does, then one might expect $\mathcal{C}_f(\text{REF})$ to mimic better the XR cells than the CIs. In compact form, this consistency relationship should be:

$$(\sqrt{e_{\text{REF/XR}}} - \sqrt{e_{\text{REF/CI}}})(\sqrt{e_{\text{XR/REF}}} - \sqrt{e_{\text{CI/REF}}}) > 0. \quad (9)$$

Indeed, whatever f = 5, 4, 3, 2, 1, the results are consistent with those from Eq.(8) (once again, see Eq.(7) and Table 4a):

$$\sqrt{e_{\text{REF/XR}}} > \sqrt{e_{\text{REF/CI}}}, \text{ consistent with } \sqrt{e_{\text{XR/REF}}} > \sqrt{e_{\text{CI/REF}}}. \quad (10)$$

The corresponding equivalent forms are related to the values of the *K*s coefficients

$$K(\text{REF, XR}) > K(\text{REF, CI}), \quad (10')$$

$$\langle \mathcal{C}_f(\text{REF}), \text{XR} \rangle > \langle \mathcal{C}_f(\text{REF}), \text{CI} \rangle \quad (10'')$$

which are consistent with Eq.(8') and Eq.(8'') respectively.

The normalized distances can be compared in Table 6:

$$D(\mathcal{C}_f(\text{REF}), \text{XR}) < D(\mathcal{C}_f(\text{REF}), \text{CI}) \quad (10''').$$

Eq.(10''') is also consistent with Eq.(8''').

Table 6

Normalized distances between predicted XR/CI and actual REF

f	5	4	3	2	1
$D(\mathcal{C}_f(\text{REF}), \text{CI})$	1.000±0.000	0.970±0.015	0.938±0.013	0.915±0.016	0.899±0.011
$D(\mathcal{C}_f(\text{REF}), \text{XR})$	0.967±0.011	0.944±0.015	0.901±0.011	0.871±0.013	0.839±0.030

The predicted species $\mathcal{C}_f(\text{REF})$ is closer to XR than to CI for all classifiers in the confusion domain.

4.3 The distances between predicted CI/XR and predicted REF

For all f≤5, the normalized distances between the predicted species are indicated in Table 7. Since

$$D(\mathcal{C}_f(\text{XR}), \mathcal{C}_f(\text{REF})) < D(\mathcal{C}_f(\text{CI}), \mathcal{C}_f(\text{REF})), \tag{11}$$

ML classifiers notice greater differences between $\mathcal{C}_f(\text{REF})$ and $\mathcal{C}_f(\text{CI})$ than between $\mathcal{C}_f(\text{REF})$ and $\mathcal{C}_f(\text{XR})$.

Table 7

Normalized distances between predicted CI/XR and predicted REF

f	5	4	3	2	1
$D(\mathcal{C}_f(\text{CI}), \mathcal{C}_f(\text{REF}))$	0.945±0.063	0.910±0.051	0.820±0.045	0.699±0.072	0.467±0.180
$D(\mathcal{C}_f(\text{XR}), \mathcal{C}_f(\text{REF}))$	0.907±0.062	0.867±0.059	0.739±0.048	0.580±0.109	0.414±0.290

To summarize, there is a general trend that emerges from Tables 8a and 8b. Except for the unitary normalized distances between the actual species, stated by the working hypotheses, all distances involving XR and REF species (actual or predicted) are smaller than those involving CI and REF species (actual or predicted). Since both CI and XR are obtained from the irradiation of REF cells, it can be said that X-rays produce less optical changes on the relevant features than carbon ion beams.

Table 8a

Normalized distances involving XR and REF species

XR	$D(\text{REF}, \text{XR})$	$D(\text{XR}, \mathcal{C}_f(\text{REF}))$	$D(\text{REF}, \mathcal{C}_f(\text{XR}))$	$D(\mathcal{C}_f(\text{XR}), \mathcal{C}_f(\text{REF}))$
f=5	1	0.967	0.942	0.907
f=4	1	0.944	0.928	0.867
f=3	1	0.901	0.854	0.739
f=2	1	0.871	0.747	0.580
f=1	1	0.839	0.637	0.414

Table 8b

Normalized distances involving CI and REF species

CI	D(REF, CI)	D(CI, $\mathcal{C}_f$(REF))	D(REF, $\mathcal{C}_f$(CI))	D($\mathcal{C}_f$(CI), $\mathcal{C}_f$(REF))
f=5	1	1.000	0.994	0.945
f=4	1	0.970	0.947	0.910
f=3	1	0.938	0.903	0.820
f=2	1	0.915	0.829	0.699
f=1	1	0.899	0.648	0.467

5. Conclusions

We propose a method to evaluate the similarities/differences among the features of N species with the help of multi-class ML classifiers working in ~~a~~ supervised classification mode. For a number of features above a certain threshold, the classification performs without errors with Acc=1. Any further reduction of the number of features leads to entering into the confusion domain defined by the occurrence of errors and the appearance of confusion matrices with Acc<1. By generating the confusion operators in a N-dimensional vector space, we proved it is possible to use the distance metric and the cross-confusion coefficients between the actual and predicted species as semi-quantitative indicators of the similarities/differences between the relevant properties of the species.

We applied the method in the case of three cell species, two irradiated with CI and XR, and one non-irradiated REF, and we computed the cross-confusion coefficients as well as the distances between predicted and actual species. All the findings endorse the general conclusion that REF species, actual or predicted, are closer to XR species, also actual or predicted, than to CI species. Since both CI and XR are obtained from the irradiation of REF cells, we conclude that X-ray radiation produces weaker changes on the relevant optical properties than CI radiation.

Acknowledgement

The authors acknowledge the groups who prepared the cultured cells IFIN HH Magurele Romania under the project 543PED/2019 and irradiated samples with carbon ions, and with X-rays, i.e., IRABAT Ganil Caen, France, and XSTRAHL X-Ray Generator Ratingen, Germany respectively. The hyperspectral images were recorded using CytoViva system acquired through INOVABIOMED Project No. P_36_611, MySMIS code 107066.

REFERENCES

[1]. *S.K. Melanthota, D. Gopal, S. Chakrabarti, A.A. Kashyap, R. Radhakrishnan, N. Mazumder,* "Deep learning-based image processing in optical microscopy", Biophys Rev 14, 2022, pp.463–481.

[2]. *V.K. Lam, T. Nguyen, T. Phan, B.M. Chung, G. Nehmetallah, C.B. Raub*, "Machine Learning with Optical Phase Signatures for Phenotypic Profiling of Cell Lines", Cytometry A. Epub 2019 Apr 22. PMID: 31008570; PMCID: PMC7068951, 2019 Jul;95(**7**), 2019, pp.757-768.

[3]. *V. Calin, M. Mihailescu, G. Petrescu, M.G. Lisievici, N. Tarba, D. Calin, V.G. Ungureanu, D. Pasov, F. Brehar, R. Gorgan, M. Moisescu, T. Savopol*, "Grading of glioma tumors using digital holographic microscopy", Heliyon, Volume **10**, Issue 9, 2024, pp.e29897.

[4]. *M. Boldeanu, C. Burileanu, H. Cucu, L Marmureanu*, "Pollen classification using classical ML algorithms on fluorescence and scattering data ", U.P.B. Sci. Bull., Series C, Vol. 84, Iss. 4, 2022, pp.123-136.

[5]. *K. Liu, Z. Ke, P. Chen, S. Zhu, H. Yin, Z. Li, Z. Chen*, "Classification of two species of Gram-positive bacteria through hyperspectral microscopy coupled with machine learning," Biomed. Opt. Express 12, 2021, pp.7906-7916.

[6]. *S-C. Yoon, N. Ekramirad,* "Machine learning–assisted multispectral and hyperspectral imaging, Machine Learning and Artificial Intelligence" in Chem.&Biol. Sensing", 2024, pp. 227-258.

[7]. *S.V. Parasca, M.A. Calin, D. Manea, R. Radvan*, "Hyperspectral imaging with machine learning for in vivo skin carcinoma margin assessment: a preliminary study", Phys Eng Sci Med., Sep; 47 (**3**), 2024, pp.1141-1152.

[8]. *R. Negoiță, M. Ilișanu, I. Irimescu, R. Popescu, M. Tudor, M. Mihailescu, E.N. Scarlat, A. M. Pleavă, A. Dinischiotu, D. Savu*, "Specific spectral sub-images for machine learning evaluation of optical differences between carbon ion and X ray radiation effects," Heliyon, Vol.**10**, Iss. 15, 2024, pp.e35249.

[9]. *K. Sekihara, H. Himuro, N. Saito, Y. Ota, T. Kouro, Y. Kusano, S. Minohara, R. Hirayama, H. Katoh, T. Sasada, D. Hoshino*, "Evaluation of X-ray and carbon-ion beam irradiation with chemotherapy for the treatment of cervical adenocarcinoma cells in 2D and 3D cultures", Cancer Cell Int. 2022 Dec. 9; 22 (**1**), 2022, pp. 391-406.

[10]. *S.Y. Chung, H. Takiyama, J.H. Kang, S. Chang, B.S.Min, H. Tsuji, S. Yamada, W. S. Koom*, "Comparison of clinical outcomes between carbon ion radiotherapy and X-ray radiotherapy for reirradiation in locoregional recurrence of rectal cancer", Sci Rep 12, 2022, pp.1845-1860.

[11]. *G. Garty, A.W Bigelow, M. Repin, H.C. Turner, D. Bian, A.S Balajee, O.V. Lyulko, M. Taveras, Y.L. Yao, D.J Brenner,* "An automated imaging system for radiation biodosimetry", Microsc Res. Tech. 2015 Jul; 78 (**7**), pp.587-98.

[12]. *Z. Ahmed, S. Amizadeh, M. Bilenko, R. Carr, W.S. Chin, Y. Dekel, X. Dupre, V. Eksarevskiy, S. Filipi, T. Finley, A. Goswami*, "Machine learning at Microsoft with ML. NET". In Proceedings of the 25th ACM SIGKDD international conference on knowledge discovery & data mining, July 2019, pp.2448-2458.
[13]. *G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, T.Y. Liu,* "LightGBM: A highly efficient gradient boosting decision tree", Advances in neural information processing systems, 2017, 30, pp.3149-3157.
[14]. *R.K. Vinayak, R. Gilad-Bachrach,* "Dart: Dropouts meet multiple additive regression trees", In Artificial Intelligence and Statistics, PMLR, February 2015, pp.489-497.
[15]. *S. Shalev-Shwartz, T. Zhang,* "Stochastic dual coordinate ascent methods for regularized loss minimization", Journal of Machine Learning Research, 14 (**1**), 2013, pp.567-599.
[16]. *L. Breiman,* "Random forests", Machine learning, 45, 2001, pp.5-32.