Medicinska istraživanja 2025; 58(1):41-47 | DOI: 10.5937/medi58-55403

OPEN ACCESS

Medical Research | Published by Faculty of Medicine University of Belgrade

ORIGINAL ARTICLE

A novel support vector machine learning approach using fractal and run-length matrix indicators for identifying nuclear changes in laryngeal cancer

Svetlana Valjarevic[1], Milan B. Jovanovic[1], Jovana Paunovic Pantic[2], Igor Pantic[3, 4]

[1]University of Belgrade, Faculty of Medicine, Clinical Hospital Center "Zemun", Belgrade, Serbia
[2] University of Belgrade, Faculty of Medicine, Department of Pathophysiology, Belgrade, Serbia
[3] University of Belgrade, Faculty of Medicine, Department of Medical Physiology, Belgrade, Serbia
[4]University of Haifa, Haifa, Israel

Received: 13 December 2024
Revised: 26 December 2024
Accepted: 30 January 2025

Check for updates

Funding information:
The authors acknowledge support from the Science Fund of the Republic of Serbia, grant No. 7739645 "Automated sensing system based on fractal, textural and wavelet computational methods for detection of low-level cellular damage", SensoFracTW. We also acknowledge the support of the Ministry of Science, Technological Development and Innovation of the Republic of Serbia grant 451-03-66/2024-03/200110 (subgrant entitled "Development of artificial intelligence models based on the random forest algorithm for the detection of discrete structural changes in the cell nucleus").

Competing interests:
The authors have declared that no competing interests exist

Correspondence to:
Igor V. Pantic
University of Belgrade, Faculty of Medicine, Institute of Medical Physiology,
26/II Visegradska Street, 11000 Belgrade, Serbia
University of Haifa, 199 Abba Hushi Blvd., Mount Carmel, Haifa, IL-3498838, Israel;
Email: igor.pantic@med.bg.ac.rs; igorpantic@gmail.com

Summary

Introduction/Aim: We aimed to propose a novel and innovative concept of a support vector machine learning algorithm that employs fractal and run-length matrix indicators of nuclear structure to identify malignant squamous epithelial cells in laryngeal cancer.

Material and Methods: Regions of interest in micrographs of laryngeal cancer and chronic laryngitis were analyzed using the box-counting fractal and run-length matrix textural techniques. For each nucleus, we quantified fractal dimension values, lacunarity, long-run emphasis, and short-run emphasis. These features were used as input data for training and testing the support vector machine model in the "Scikit-learn" library for Python.

Results: The support vector machine model produced relatively good performance indicators. The classification accuracy of the model was 0.83, indicating its adequate ability to distinguish cancer cells from non-cancer cells in our sample. The F1 score (the harmonic mean of precision and recall) was 0.83, suggesting a relatively good balance between these two metrics. The value of the Matthews Correlation Coefficient for this model was 0.65, which indicated moderate agreement between the predicted and actual labels and balanced performance across the two classes.

Conclusion: The proposed model provides a solid foundation for further developing artificial intelligence systems for signal analysis in cancer research. If the limitations of this concept are addressed, future research can focus on developing a more comprehensive machine-learning model for identifying laryngeal epithelial cancer cells.

Keywords: artificial intelligence, machine learning, nucleus, chromatin, fractal

Cite this article as: Valjarevic Svetlana, B. Jovanovic M, Paunovic Pantic J, Pantic I. A novel support vector machine learning approach using fractal and run-length matrix indicators for identifying nuclear changes in laryngeal cancer; Medicinska istraživanja 2025; 58(1):41-47 DOI: 10.5937/medi58-55403

INTRODUCTION

Novel Information and Communication Technologies (ICTs) have enabled a significant increase in the level of automation when analyzing physiological and pathological signals. One such technology is support vector machine (SVM) learning for identification of patterns present in two-dimensional signal data (1, 2). SVM is a supervised learning approach where a computer model during development is presented with numerous correctly associated inputs and outputs. In time, the model learns of new associations and patterns and can predict the correct outcome based on the new inputs. In the case of SVM, the model uses a decision boundary, a line, a plane, or a hyperplane to divide data points, which is particularly useful for classification tasks. In biomedical sciences, the input data usually belong to multidimensional spaces so hyperplane is commonly used for the separation of data points. The closest data points with the greatest impact on the shape and location of the hyperplane are often referred to as support vectors (3, 4).

The SVM approach is commonly used in clinical medicine to discriminate between physiological and pathological states and conditions. Here, various kernel functions can be used to separate data, including polynomial, sigmoid, and radial basis kernel functions. Inputs can include data on two-dimensional patterns such as cell size and shape, fractal parameters, textural indicators and wavelet transform quantifiers. The outputs can include a class of the cell (i.e., damaged or intact), patient status, diagnostic category, prognostic indicators, and other data that can represent a descriptor of a data class. Less frequently, SVMs can be built for regression purposes, for prediction of a continuous physiological or pathological variable (5, 6).

When using SVM and machine learning in general for identification of data patterns associated with cancer, probably the best approach would be to apply a mathematical analysis of textural or related changes in cell structure. Previously, such concepts have been introduced in machine learning, particularly in applying co-occurrence matrix analysis for the training of supervised machine learning models and the models based on decision trees (7). Apart from the co-occurrence matrix, run-length matrix (RLM) and fractal approaches also have certain potential in training both SVM and alternative models for cell classification in other pathologies (8, 9). The potential rationale for using SVM would be its greater ability to handle high-dimensional spaces and better utilization of computational and processing power.

In our previous work, we have demonstrated that, based on wavelet and gray-level co-occurrence matrix (GLCM) textural data related to nuclear organization, it is possible to train SVM and random forest models that can be useful for the identification of squamous epithelial cancer cells in laryngeal cancer (LC) tissue (7). This research further raised questions if such a model can be developed to use other types of inputs. Hereby, we present a concept of an SVM model that utilizes a combination of nuclear fractal and run-length matrix parameters in order to differentiate between intact and malignant squamous epithelial cells (SECs) in laryngeal tissue. The proposed model uses 4 mathematical parameters: fractal dimension, lacunarity, RLM short-run emphasis, and RLM long-run emphasis. Our initial results indicate that when applied to nuclear regions of interest, the developed SVM model may have satisfactory performance.

MATERIALS AND METHODS

The research is a continuation of our previous work (7), where digital micrographs were obtained from biopsy samples of 50 patients diagnosed with laryngeal squamous cell carcinoma and the control group consisting of 50 patients established to have only chronic laryngitis. The patients were previously diagnosed and treated at the University Clinical Hospital Center Zemun, Belgrade, Serbia, and the researchers obtained approval from the Ethical Commission of the University of Belgrade, Faculty of Medicine Serbia (Approval No. 17/I-17, 12-Jan-2023.). Inclusion and exclusion criteria for this retrospective study as well as micrograph creation and characteristics were explained earlier (7).

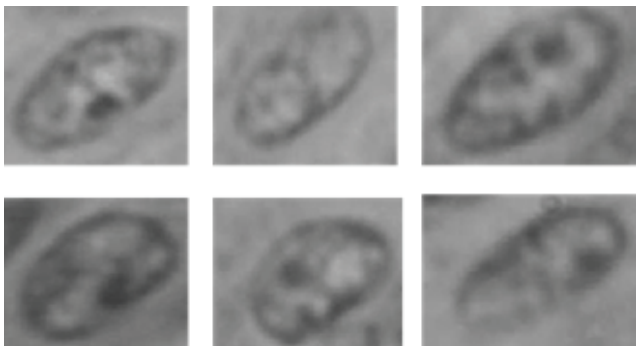

Figure 1. Example of squamous epithelial cell (SEC) nuclei in a format suitable for binarization and fractal analysis using the box-counting method.

Fractal and run-length matrix analysis was performed on a total of 2000 SEC nuclear regions of interest (ROIs), of which 1000 belonged to the experimental and 1000 to the control group. The contrast, brightness and other image indicators were previously adjusted to achieve the optimal binarization thresholds and performance of fractal analysis (Figure 1). For the fractal analysis of cell nuclei, we used FracLac software, previously developed by Audrey Karperien and Charles Sturt University (10, 11). Fractal dimension (D) was calculated using the traditional box-counting technique (10, 12-14) where the object is covered by a multitude of boxes of different sizes (ε) after which the calculations are done from a slop of

logarithmic regression line where box numbers (N) and the sizes are taken into account:

$$D=\lim\varepsilon \rightarrow 0[\log N_{\varepsilon}/\log \varepsilon$$

While fractal dimension is an indicator of complexity, fractal lacunarity (λ) is an indicator of heterogeneity and largely depends on the number, size and other characteristics of architectural gaps. It is determined from the variation coefficient (CV) for resolution unit mass considering the grid position (g) and scale:

$$\lambda = CV_{\varepsilon,g}{}^{2} = (\sigma_{\varepsilon,g}/_{\mu\varepsilon,g})^{2}$$

Run-length matrix analysis was done in MaZda software (version 4.6) created by the authors at the Institute of Electronics, Technical University of Lodz (TUL), Poland (15-18). Briefly, this method analyzes sequences of consecutive resolution units (also called "runs") in order to quantify elements of textural anisotropy, as well as other aspects of spatial relationships within the two-dimensional signals. In this work, we focused on two major quantifiers of RLM, Long Run Emphasis (LngREmph) and Short Run Emphasis (ShrtREmph). Both indicators are dependent of respective frequencies of runs with length j that have a gray value of i, or p(i,j):

$$\text{ShrtREmph} = \left(\sum_{i=1}^{N_g}\sum_{j=1}^{N_r}\frac{p(i,j)}{j^2}\right)/C$$

$$\text{LngREmph} = \left(\sum_{i=1}^{N_g}\sum_{j=}^{N_r} j^2 p(i,j)\right)/C$$

In these formulas, coefficient C is determined as:

$$C = \sum_{i=1}^{N_g}\sum_{j=1}^{N_r} p(i,j)$$

Support vector machine algorithm was proposed within the Google Colab and Jupyter Notebook service, using the scikit-learn machine learning library (version 1.3) for Python programming language (19-22). The obtained values of nuclear ROI fractal dimension, lacunarity, LngREmph and ShrtREmph were used as input parameters, while the output consisted of the cell class related to its affiliation to the experimental or the control group. P-dimensional mathematical vectors were the representations of specific data points. Consequently, separation of the data was achieved using a (p-1)-dimensional hyperplane(s). For the model, in Scikit-learn library, we calculated the values of the F1 score (harmonic mean of precision and recall), Matthews Correlation Coefficient (MCC, measure of the quality of binary classifications), and the area under the Receiver Operating Characteristic (ROC) curve, a performance measurement for classification problems at various thresholds settings. Matplotlib, a plotting library for Python was used for visualizing model performance. NumPy (Numerical Python), a core scientific computing library, was used for numerical computation and other data operations.

RESULTS

All four parameters that we quantified in this study were suitable for supervised machine learning purposes and for the development of support vector machine models. The average nuclear ROI long run emphasis value was 5.417 ± 1.997 in the LC group and 16.134 ± 10.479 in the controls. Conversely, values of short-run emphasis were lower in controls compared to the LC group and equaled 0.461 ± 0.085 and 0.505 ± 0.093, respectively. A statistically highly significant difference ($p<0.01$) was observed between the groups for both run-length matrix indicators. We also observed significant differences between the groups in values of fractal parameters of cell nuclei ($p<0.01$). The average nuclear ROI fractal dimension value was 1.522 ± 0.130 in the LC group and 1.585 ± 0.119 in the controls. The average value of fractal lacunarity was 0.505 ± 0.093 in the LC group and 0.461 ± in the controls. These results indicated that the nuclear patterns in squamous epithelial cells in laryngeal cancer tissue are characterized by reduced complexity and possibly increased levels of fractal heterogeneity.

After 5-fold cross-validation, the support vector machine model produced relatively good performance indicators. The classification accuracy of the model equaled 0.83, suggesting the adequate ability of the model to separate cancer from non-cancer cells in our sample. The precision and recall of the model were 0.79 and 0.87, respectively, highlighting the model's strong performance in the identification of true positives but with room for improvement. The value of the F1 score (harmonic mean of precision and recall) was 0.83, indicating that there is a relatively good balance between the two metrics. The value of the Matthews Correlation Coefficient for this model was 0.65, which suggested a moderate agreement between the predicted and actual labels and balanced performance across the two classes. The area under the Receiver Operating Characteristics curve was 0.89 which suggested a good discriminatory power of the SVM classifier. The ROC curve is presented in **Figure 2**.

Feature importance analysis was performed taking into account the best hyperparameters during hyperparameter optimization. For linear kernel, the most influential feature for model development and performance was the fractal dimension, with a score of 1.78. The second most important feature was nuclear ROI lacunarity, with a score of 1.50. Short-run emphasis had a fracture importance of 1.44, while long-run emphasis had a feature importance of 0.31. During the creation of the support vector machine learning model, we also performed the stratified k-fold cross-validation on five splits with the values of the

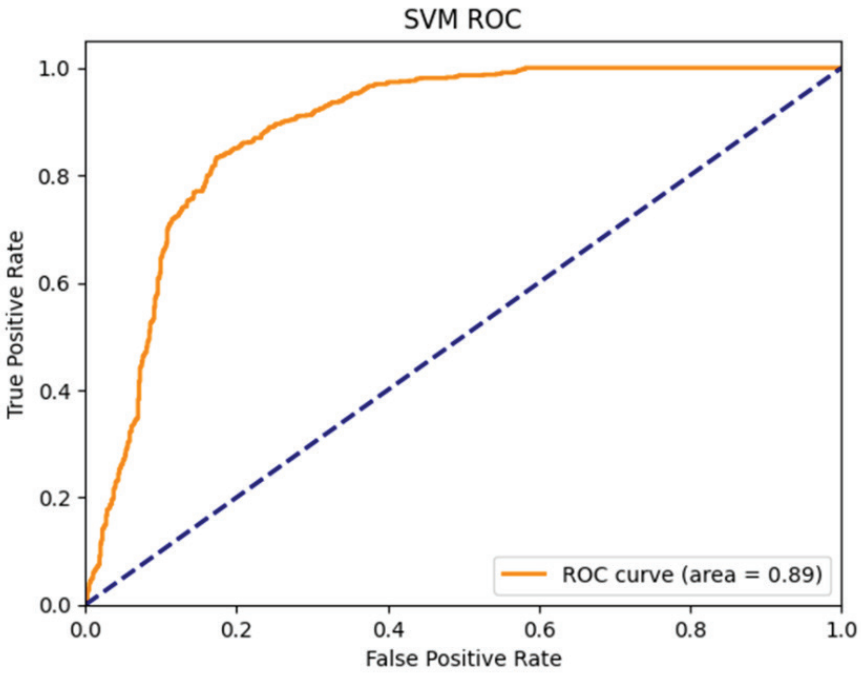

Figure 2. Receiver Operating Characteristics curve of the support vector machine model.

"shuffle" parameter set to "True" in order to increase the degree of randomness. Various hyperparameters for this model have been considered, including the Regularization parameter, shrinking heuristic, Tolerance for stopping criteria, Maximum number of iterations, and Kernel types. The proposed segment of the code is shown in **figure 3**.

DISCUSSION

In this work, we explain the concept of a support vector machine algorithm intended for differentiation between malignant and non-malignant laryngeal squamous epithelial cells. We present a model that functions based on fractal and run-length matrix inputs obtained through mathematical analysis of cell nuclear architecture. The quantified indicators of the model performance suggest that the SVM concept, as a whole, in these clinical and experimental settings, holds some potential, especially when considering its future integration with similar artificial intelligence-based computational systems.

This research is a continuation of our recently published work on the application of nuclear GLCM and wavelet mathematical inputs for supervised machine learning (7). There, we created an SVM model using nuclear inverse difference moment, angular second moment, contrast, correlation, sum variance, as well as coefficient energies of the discrete wavelet transform. Similarly, in our present study, the SVM approach was used to separate malignant from non-malignant laryngeal squamous epithelial cells. The developed SVM model demonstrated very good performance since its classification accuracy was 83% and the area under the receiver operating characteristic curve was 0.89. This performance was similar to the alternative decision tree–based, random forest model described in the previous research (7).

```
from sklearn.svm import SVC
from sklearn.model_selection import cross_val_predict, cross_val_score, StratifiedKFold, GridSearchCV
from sklearn.metrics import roc_curve, auc, f1_score, precision_score, recall_score, matthews_corrcoef, classification_report, accuracy_score
import matplotlib.pyplot as plt
import numpy as np

X = df[['Horzl_LngREmph', 'Horzl_ShrtREmp', 'FD', 'Lac']]
y = df['target']


param_grid = {
    'C': [0.01, 0.1, 1, 10, 100, 1000],
    'kernel': ['linear', 'rbf', 'poly', 'sigmoid'],
    'gamma': ['scale', 'auto', 0.001, 0.01, 0.1, 1],
    'degree': [2, 3, 4, 5],
    'coef0': [0, 0.1, 0.5, 1, 2],
    'shrinking': [True, False],
    'class_weight': [None, 'balanced'],
    'tol': [1e-4, 1e-3, 1e-2],
    'max_iter': [-1, 1000, 5000]
}
```

Figure 3. The segment of the Python Scikit-learn code with optional hyperparameter tuning.

This is not the first research where the run-length matrix and fractal parameters of cell nuclei have been used for machine learning classification tasks. In 2023, the values of fractal dimension were considered as inputs for the creation of an SVM classifier for the detection of structural changes of cell nuclei following exposure to a hyperosmotic environment (23). When combining this indicator with GLCM and DWT quantifiers, the resulting model was shown to have a relatively good classification accuracy of 71.7%. Although this performance is not considered excellent, it may still hold scientific potential since human-based identification of structural changes is much less effective. In recent work, nuclear textural features based on run length matrix and wavelet analyses were used for the detection of discrete changes in chromatin distribution associated with iron nanoparticle exposure (8). Therein, it was shown that for hepatocytes, it is possible to create a supervised machine learning model, specifically those based on random forest and gradient boosting architecture, to distinguish between intact and possibly damaged cells.

The rationale for using run-length matrix and fractal parameters for the detection of change in nuclear structure lies in the fact that these techniques can objectively quantify structural alterations in digital micrographs and their regions of interest. This is particularly the case with ROIs of nuclei and nuclear chromatin, where the changes are almost invisible to the human eye during a conventional microscopy assessment. Nuclear chromatin distribution is governed by numerous intracellular and intranuclear mechanisms (24-27), and cells exposed to toxic environments or malignant transformation may exhibit changes in euchromatin and heterochromatin patterns detectable using contemporary computational methods. In the case of cancer cells, one can expect significant alterations in nuclear morphology, which are often noticeable by the human eye and may include extensive redistribution of heterochromatin, chromosomal rearrangements, and translocations, as well as chromatin marginalization. Also, in malignant cells, nuclei may have irregular contours or be distorted, and overall DNA content and density may be increased, leading to a phenomenon called hyperchromasia. Finally, nucleoli may be prominent, enlarged and more conspicuous. These alterations may be detectable using fractal and textural methods, particularly bearing in mind that in the past, the techniques have been successfully used for the detection of nuclear changes in both physiological and pathological conditions (28-30).

Limitations of our approach include the relatively small ROI sample size used for training the model, as well as the usual drawbacks of the development of machine learning models in biomedical research. Support vector machines, although relatively useful when handling multidimensional data in medicine, may still suffer from generalizability to other cell populations and clinical conditions. In other words, the model may adequately identify and classify cancer cells in this sample of malignant squamous epithelial cells, but this effectiveness may not be manifested in other circumstances. Also, like in many other approaches in supervised machine learning, the interpretability of SVMs remains relatively low since the inner workings of the model, due to its multidimensionality and overall complexity, remain elusive. Finally, the fact that the model is focused on cell nuclear ROIs rather than on the patient and technical issues related to the reproducibility of fractal and RLM analysis also pose a significant limitation. Therefore, our results and the SVM approach should be considered more as a preliminary concept than a fully applicable and developed machine learning model.

CONCLUSION

In conclusion, we present a support vector machine learning concept intended to differentiate between nuclear regions of interest of malignant and non-malignant laryngeal squamous epithelial cells. The model is based on nuclear values of fractal dimension, fractal lacunarity, run-length matrix short-run emphasis, and long-run emphasis. Preliminary results show that the model could reach acceptable discriminatory power and other measures of performance, giving it the potential to be integrated with other machine-learning approaches for the identification of cancer cells. If the limitations of this concept are overcome, future research can be focused on the development of a more comprehensive and effective artificial intelligence system based on run-length matrix and fractal indicators with potential applications in clinical medicine.

Conflicts of interest

None to declare.

Author contributions

SV and IP contributed to the conception and design of the work, the acquisition, analysis, and interpretation of data, and the preparation of the manuscript draft. MBJ and JPP contributed to the conception and design of the work and the preparation of the manuscript draft.

Ethical approval

The researchers obtained approval from the Ethical Commission of the University of Belgrade, Faculty of Medicine Serbia (Approval No. 17/I-17, 12-Jan-2023.). All procedures performed in studies involving human participants were in accordance with the ethical standards of the institutional and/or national research committee and with the 1964 Helsinki Declaration and its later amendments or comparable ethical standards.

REFERENCES

1. Xu L, Liang G, Shi S, Liao C. SeqSVM: A Sequence-Based Support Vector Machine Method for Identifying Antioxidant Proteins. Int J Mol Sci. 2018;19(6).
2. Cai CZ, Han LY, Ji ZL, Chen X, Chen YZ. SVM-Prot: Web-based support vector machine software for functional classification of a protein from its primary sequence. Nucleic Acids Res. 2003;31(13):3692-7.
3. Lee HK, Kim CH, Bhattacharjee S, Park HG, Prakash D, Choi HK. A Paradigm Shift in Nuclear Chromatin Interpretation: From Qualitative Intuitive Recognition to Quantitative Texture Analysis of Breast Cancer Cell Nuclei. Cytometry A. 2020.
4. AlKubeyyer A, Ben Ismail MM, Bchir O, Alkubeyyer M. Automatic detection of the meningioma tumor firmness in MRI images. J Xray Sci Technol. 2020;28(4):659-82.
5. Jiang H, Ching WK. Classifying DNA repair genes by kernel-based support vector machines. Bioinformation. 2011;7(5):257-63.
6. Vidya KS, Ng EY, Acharya UR, Chou SM, Tan RS, Ghista DN. Computer-aided diagnosis of Myocardial Infarction using ultrasound images with DWT, GLCM and HOS methods: A comparative study. Comput Biol Med. 2015;62:86-93.
7. Valjarevic S, Jovanovic MB, Miladinovic N, Cumic J, Dugalic S, Corridon PR, et al. Gray-Level Co-occurrence Matrix Analysis of Nuclear Textural Patterns in Laryngeal Squamous Cell Carcinoma: Focus on Artificial Intelligence Methods. Microscopy and Microanalysis. 2023;29(3):1220-7.
8. Paunovic Pantic J, Vucevic D, Radosavljevic T, Corridon PR, Valjarevic S, Cumic J, et al. Machine learning approaches to detect hepatocyte chromatin alterations from iron oxide nanoparticle exposure. Scientific reports. 2024;14(1):19595.
9. Pantic I, Paunovic Pantic J. Artificial Intelligence in Chromatin Analysis: A Random Forest Model Enhanced by Fractal and Wavelet Features. Fractal and Fractional. 2024;8(8):490.
10. Karperien AL, Jelinek HF. Box-Counting Fractal Analysis: A Primer for the Clinician. Adv Neurobiol. 2024;36:15-55.
11. Karperien A. FracLac for ImageJ. 1999-2023 [Available from: http://rsb.info.nih.gov/ij/plugins/fraclac/FLHelp/Introduction.htm.
12. Revittser A, Selin I, Negulyaev Y, Chubinskiy-Nadezhdin V. The analysis of F-actin structure of mesenchymal stem cells by quantification of fractal dimension. PloS one. 2021;16(11):e0260727.
13. Eser S, Saribas E. Anatomical assessment of the trabecular structure of the alveolar bone in periodontal disease by fractal analysis method. Folia Morphol (Warsz). 2024;83(1):157-67.
14. Yavuz E, Yardimci S. Comparison of periapical radiography, panoramic radiography, and CBCT in the evaluation of trabecular bone structure using fractal analysis. Oral Radiol. 2024.
15. Strzelecki M, Szczypinski P, Materka A, Klepaczko A. A software tool for automatic classification and segmentation of 2D/3D medical images. Nuclear Instruments & Methods In Physics Research A. 2013;702:137-40.
16. Szczypinski P, Strzelecki M, Materka A, Klepaczko A. MaZda-A software package for image texture analysis. Computer methods and programs in biomedicine. 2009;94(1):66-76.
17. Szczypinski P, Strzelecki M, Materka A. MaZda - a Software for Texture Analysis. Proc of ISITC 2007, November 23-23, 2007; Republic of Korea2007. p. 245-9.
18. Kociołek M, Materka A, Strzelecki M, Szczypinski P, editors. Discrete wavelet transform –derived features for digital image texture analysis. Proc of Interational Conference onSignals and Electronic Systems, 18-21 September 2001; 2001; Lodz, Poland.
19. Pedregosa F, Varoquaux G, Gramfort A, Michel V, Thirion B, Grisel O, et al. Scikit-learn: Machine Learning in Python. J Mach Learn Res. 2011;12:2825–30.
20. Ahn S. Building and analyzing machine learning-based warfarin dose prediction models using scikit-learn. Transl Clin Pharmacol. 2022;30(4):172-81.
21. Kondo Y. [[Fundamentals] 6. Calculation of Machine Learning-related Statistics and Evaluation Values Using Scikit-learn]. Nihon Hoshasen Gijutsu Gakkai Zasshi. 2023;79(12):1387-94.
22. Yang Y, Hu P, Chen SR, Wu WW, Chen P, Wang SW, et al. Predicting the Activity of Oral Lichen Planus with Glycolysis-related Molecules: A Scikit-learn-based Function. Curr Med Sci. 2023;43(3):602-8.
23. Pantic I, Valjarevic S, Cumic J, Paunkovic I, Terzic T, Corridon PR. Gray Level Co-Occurrence Matrix, Fractal and Wavelet Analyses of Discrete Changes in Cell Nuclear Structure following Osmotic Stress: Focus on Machine Learning Methods. Fractal and Fractional. 2023;7(3):272.
24. Pantic I, Paunovic Pantic J. Artificial Intelligence in Chromatin Analysis: A Random Forest Model Enhanced by Fractal and Wavelet Features. Fractal and Fractional [Internet]. 2024; 8(8).
25. Espinoza JA, Kanellis DC, Saproo S, Leal K, Martinez JF, Bartek J, et al. Chromatin damage generated by DNA intercalators leads to degradation of RNA Polymerase II. Nucleic Acids Res. 2024.
26. Minami K, Iida S, Maeshima K. Chromatin organization and DNA damage. Enzymes. 2022;51:29-51.
27. Wang D, Lu X, Jiang Y, Pan L, Zhu F, Yu A, et al. The chromatin remodeling protein BPTF mediates cell cycle, proliferation and apoptosis in porcine ovarian granulosa cells. Theriogenology. 2023;211:172-81.
28. Mattos AC, Florindo JB, Adam RL, Lorand-Metze I, Metze K. The Fractal Dimension Suggests Two Chromatin Configurations in Small Cell Neuroendocrine Lung Cancer and Is an Independent Unfavorable Prognostic Factor for Overall Survival. Microsc Microanal. 2022:1-5.
29. Dincic M, Todorovic J, Nesovic Ostojic J, Kovacevic S, Dunderovic D, Lopicic S, et al. The Fractal and GLCM Textural Parameters of Chromatin May Be Potential Biomarkers of Papillary Thyroid Carcinoma in Hashimoto's Thyroiditis Specimens. Microsc Microanal. 2020;26(4):717-30.
30. Gupta S, Savala R, Gupta N, Dey P. Fractal dimension and chromatin textural analysis to differentiate follicular carcinoma and adenoma on fine needle aspiration cytology. Cytopathology. 2020;31(5):491-3.

NOV PRISTUP MAŠINSKOM UČENJU POMOĆU POTPORNIH VEKTORA KOJI KORISTI FRAKTALNE I MATRIČNE POKAZATELJE DUŽINE NIZA ZA IDENTIFIKACIJU JEDARNIH PROMENA KOD KARCINOMA LARINKSA

Svetlana Valjarević[1], Milan B. Jovanović[1], Jovana Paunović Pantić[2], Igor Pantić[3, 4]

Sažetak

Uvod/Cilj: Cilj istraživanja je bio da predložimo novi i inovativni koncept algoritma mašinskog učenja sa potpornim vektorima koji koristi fraktalne i matrične pokazatelje dužine niza strukture jedra za identifikaciju malignih skvamoznih epitelnih ćelija kod laringealnog karcinoma.

Materijal i metode: Analizirane su regije od interesa na mikrografima laringealnog karcinoma i hroničnog laringitisa korišćenjem fraktalne tehnike brojanja kvadrata i teksturalne tehnike matrice dužine niza. Za svako jedro kvantifikovane su vrednosti fraktalne dimenzije, lakunarnosti, naglašenosti dugih nizova i naglašenosti kratkih nizova. Ove karakteristike korišćene su kao ulazni podaci za treniranje i testiranje modela mašinskog učenja sa potpornim vektorima u biblioteci *Scikit-learn* za Python.

Rezultati: Model mašinskog učenja sa potpornim vektorima pokazao je relativno dobre pokazatelje performansi. Tačnost klasifikacije modela iznosila je 0,83 što ukazuje na adekvatnu sposobnost modela da razlikuje kancerogene od nekancerogenih ćelija u našem uzorku. Vrednost F1 ocene (harmonijska sredina preciznosti i osetljivosti) iznosila je 0,83, što sugeriše relativno dobar balans između ova dva pokazatelja. Vrednost Metjuzovog koeficijenta korelacije za ovaj model iznosila je 0,65 što ukazuje na umerenu saglasnost između predviđenih i stvarnih oznaka i uravnotežene performanse modela unutar dve klase.

Zaključak: Predloženi model pruža solidnu osnovu za dalji razvoj sistema veštačke inteligencije za analizu signala u istraživanjima raka. Ako se ograničenja ovog koncepta prevaziđu, buduća istraživanja mogu biti usmerena na razvoj sveobuhvatnijeg modela mašinskog učenja za identifikaciju epitelnih ćelija laringealnog karcinoma.

Ključne reči: veštačka inteligencija, mašinsko učenje, jedro, hromatin, fraktal

Primljen: 13.12.2024. I **Revizija:** 26.12.2024. I **Prihvaćen:** 30.01.2025.

Medicinska istraživanja 2025; 58(1):41-47