Original scientific paper

UDK 616.98-039.71:578.8HPV
DOI 10.7251/COMEN2202178J

DEEP LEARNING ALGORITHM FOR CERVICAL CANCER DETECTION BASED ON IMAGES OF OPTOMAGNETIC SPECTRA

Branislava Jeftić[*1], Igor Hut[1], Ivana Stanković[1], Jovana Šakota Rosić[1], Lidija Matija[1], Đuro Koruga[1,2]

[1] University of Belgrade, Faculty of Mechanical Engineering, Belgrade, Serbia
[2] TFT Nano Center, Belgrade, Serbia

[*] Corresponding author: bjeftic@mas.bg.ac.rs

Abstract: In order to further investigate performance of Optomagnetic Imaging Spectroscopy in cervical cancer detection, deep learning algorithm has been used for classification of optomagnetic spectra of the samples. Optomagnetic spectra reflect cell properties and based on those properties it is possible to differentiate normal cells from cells showing different levels of dysplasia and cancer cells. In one of the previous research, Optomagnetic imaging spectroscopy has demonstrated high percentages of accuracy, sensitivity and specificity in cervical cancer detection, particularly in the case of binary classification. Somewhat lower accuracy percentages were obtained in the case of four class classification. Compared to the results obtained by conventional machine learning classification algorithms, proposed deep learning algorithm achieves similar accuracy results (80%), greater sensitivity (83.3%), and comparable specificity percentages (78%).

Keywords: Optomagnetic Imaging Spectroscopy, cervical cancer, deep learning, convolutional neural network.

1. INTRODUCTION

Optomagnetic imaging spectroscopy (OMIS) is a nano-physical method for material characterization. This relatively new approach involves exploration of sample magnetic properties based on the light-matter interaction [1]. It uses visible light for the diagnostics of sample state, ensuring non-invasiveness, i.e. that the considered sample won't be damaged during inspection, which is highly important in the case of biological sample diagnostics. Optomagnetic imaging spectroscopy has been introduced in previous researches as a tool for detection of different types of cancer such as cervical cancer, colorectal cancer, skin cancer and oral cancer [2-6]. Normal and cancerous cells and tissues were subjected to OMIS and based on their optomagnetic spectra classification into healthy/cancer group was made with various machine learning algorithms [2-6]. Performance evaluation of the classification algorithms was made based on sensitivity, specificity and accuracy measures. Matija et al proposed algorithm using Naïve Bayes classifier for cervical cancer detection which achieved sensitivity of 78.16%, specificity of 97.92%, and accuracy of 85.18% [2]. They have used stained Papanicolaou smears and characteristic OMIS spectra intensities and peak positions for classification of samples into two classes: healthy and cancerous, i.e. II Papanicolaou group and V Papanicolaou group. Jeftic et al investigated the application of Naïve Bayes classifier for abnormal cervical cell detection using unstained cervical smears and reported that the machine learning model gave sensitivity of 73%, specificity of 82% and accuracy of 81% for binary classification problem (II Papanicolaou group – III, IV and V Papanicolaou group classes) [3].

While classical machine learning models involve image segmentation step in order to provide characteristic features for classification problem solving, novel deep learning approach automatizes all steps, from segmentation to classification by learning complex features from the images of samples. Deep learning framework is currently being tested in numerous studies on cancer detection. Most of the proposed computer aided diagnostic systems are concerning medical images as a source of information relevant for the detection of abnormal cells and tissues which can enable early cancer diagnosis.

The challenge that deep learning systems for medical image classification poses is the demand of high amount of data and even distribution of data in multiclass classification problems. While the data deficiency may be compensated by pre-training the deep network on large available datasets through transfer learning, the problem of imbalanced datasets is usually addressed with methods based on under sampling, data augmentation (by rotation, translation, flipping, distortion, introducing noise, or scaling the images), as well as cost sensitive learning and model calibration [7-9].

Amongst vast amount of medical digital data available, scientific studies on deep learning implementation in diagnostics are mainly concentrated on the types of cancer such as breast cancer, melanoma, cervical cancer, brain tumor, colorectal cancer and lung cancer [10-15]. Regarding cervical cancer, there is a significant number of papers dealing with testing and discovering the best deep learning network architecture in terms of the highest accuracy measures, some of which will be presented in this paper. The aim is to improve conventional diagnostic tests, such as Papanicolaou test and liquid based cytology, which still have relatively high false negative rate, due to subjectivity of the test, poorly prepared samples, pathologist's workload, etc. [16].

Ghoneim et al proposed cervical cancer classification system based on convolution neural network (CNN) and extreme learning machines. They used pre-trained deep CNN models, namely VGG-16 Net and the CaffeNet, fine-tuned by dataset from Herlev database, as well as shallow CNN model. VGG-16 and CaffeNet were superior to shallow model to some extent, and in combination with extreme learning machine classifier VGG-16 and CaffeNet models achieved accuracy of 99.7% for cervical cell classification into two classes (normal/abnormal) and accuracy of 91.2% for classification into seven classes (different types of normal and abnormal cells) [17]. Rahaman et al obtained high percentages of accuracy for 2, 3 and 5 class classification using a hybrid deep feature fusion technique based on deep learning. Proposed DeepCervix framework tested on single cell cervical cytopathology images, achieved accuracy higher than 99% in all three classification problems: 99.85%, 99.38% and 99.14% for 2, 3 and 5 class classification problems, respectively. Among several deep learning models tested, VGG16 model proved to be superior in terms of performance compared to VGG19, ResNet50 and XceptionNet [18]. Hussain et al tested six deep convolutiona neural networks: AlexNet, VGG19, VGG19, ResNet50, ResNet101 and GoogleNet and ensemble classifier for multiclass cervical cancer diagnosis prediction. Ensemble classifier combined the three models out of six that showed best performance, namely ResNet50, ResNet101 and GoogleNet. The ensemble model outperformed other deep learning models with area under the curve (AUC) of 0.97 [19]. Alyafeai and Ghouti have developed a fully automated deep learning pipeline for cervical cancer classification using images of cervical region, cervigrams. For cervix detection, they used modified GoogLeNet classifier model and achieved 0.68 detection accuracy, while the convolutional neural network used for cancer classification gave AUC of 0.82 [20].

In most of the studies regarding cervical smear classification by deep learning models, small image datasets were used and the images were containing separate cervical cells. This does not reflect real clinical scenario, where cervical smears contain overlapped cells and some background staining interferences. By eliminating interfering factors, the application of such a model is narrowed in clinical practice. In order to overcome this problem, Martinez-Mas et al proposed Cell Merger Approach combined with convolutional neural network model and showed that by creating a dataset with overlapped and folded cells and applying deep model for classification, accuracy of 88.8%, sensitivity of 0.92 and specificity of 0.83 can be achieved [21].

In this paper, system for cervical sample classification based on Optomagnetic spectra of the sample and deep learning network is proposed. Based on the previous results obtained by Optomagnetic imaging spectroscopy and classical machine learning algorithms in the field of cancer detection and its

Table 1. *Papanicolaou test result findings and number of collected smears in each Papanicolaou group*

Papanicolaou groups	Description of the Papanicolaou test findings	Number of samples used in the study
Papanicolaou group II	Atypical cells with no evidence of malignancy	141
Papanicolaou group III	Abnormal, suspected cells	112
Papanicolaou group IV	Small number of malignant cells strongly indicating malignancy	42
Papanicolaou group V	Numerous malignant cells	146
Total number		441

potential as an accurate, fast and easy to use method, the aim of the study was to implement deep learning model in order to find the most optimal machine learning algorithm that would give the highest accuracy results in classifying cervical samples based on their optomagnetic properties. The images of optomagnetic sample spectra were used as an information carrier and fed into the deep neural network, contrary to standard cervical cell images used in other scientific studies. Since the optomagnetic spectra of normal, healthy cervical cells significantly differs from the optomagnetic spectra of cancer cells, the hypothesis of successful classification by deep learning network was made.

2. MATERIALS AND METHODS

Papanioclaou smears were collected from women routinely visiting two obstetric clinics in Serbia, Clinic of Gynaecology and Obstetrics – Narodni Front, Belgrade and clinic in Zajecar. In total, 441 cervical smears were used in the study. Papanicolaou test results were reported in one of the following groups: II, III, IV and V Papanicolau group. Number of each Papanicolaou group smears and descriptions of the findings are given in the Table 1.

Each Papanicolaou smear was prepared according to the standard procedure including fixation and staining of the cervical cells, prior to microscopic evaluation. Cancer cases were verified by histopathology. Papanicolaou smears were subjected to screening with Optomagnetic imaging spectroscopy and for each microscopic slide with cervical cells optomagnetic spectrum was obtained. The screening process followed the procedure that included collection of sample digital images in two modes (using diffuse white light and polarized white light), processing the images with convolution algorithm and production of optomagnetic spectra (Figure 1) [1].

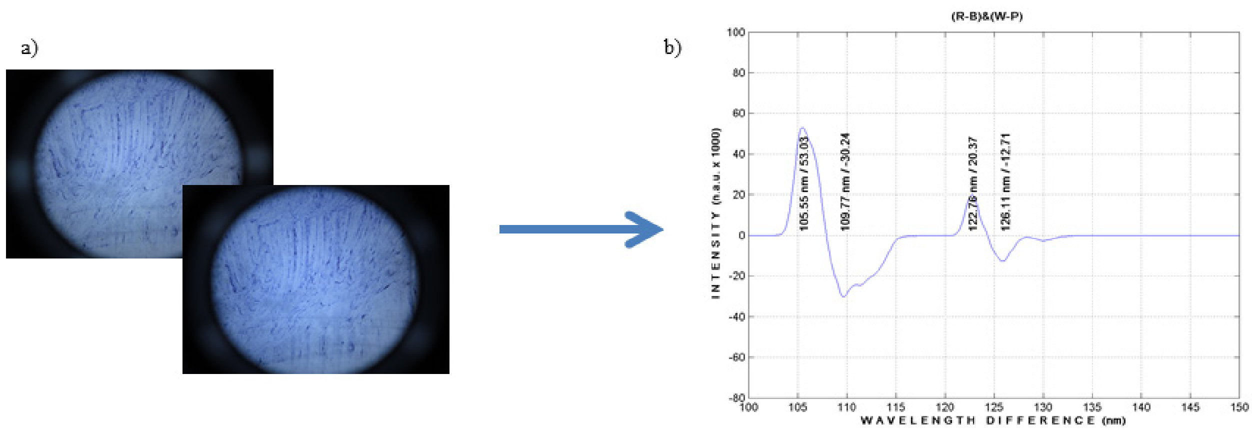

Figure 1. *a) A pair of digital images of the sample made with OMIS, using white diffuse and polarized white light, b) Optomagnetic spectrum obratined for a cervical smear sample (spectrum example for the II Papanicolaou group sample)*

Images of the sample optomagnetic spectra were used as an input in deep neural network. Samples were divided into two classes: class pap23 was made of sample spectra images of the Papanicolaou II and Papanicolaou III group smears, and class pap45 was made of sample spectra images of the Papanicolaou IV and Papanicolaou V group smears. The classes were labeled as “low-risk” and “high-risk” groups, respectively. The binary classification problem was chosen over four class problem mainly because of the uneven distribution of the samples across the Papanicolaou groups II-V. The deep neural network used in this study is VGG16 convolutional neural network, configuration D [22]. On the top of the VGG16 convolutional base, following layers of fully connected architecture were added: 256 neurons with ReLu activation function, dropout layer, batch normalization, layer of 4096 neurons with ReLu activation, batch norm layer, with last layer ending with softmax function. Since the dataset was relatively small, convolutional neural network with the Image-Net pretrained weights was used.

To assess performance of the proposed model for cervical cancer detection, we have used sensitivity, specificity and accuracy. Evaluation measures were calculated according to the following formulas:

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

where true positive (TP) is the number of correctly labeled positive samples, true negative (TN) is the number of correctly labeled negative samples, false negative (FN) is the number of true positive samples incorrectly labeled by classifier as negative and false positive (FP) is the number of true negative cases incorrectly labeled as positive cases.

3. RESULTS AND DISCUSSION

Binary classification of the cervical samples was done by obtaining Optomagnetic spectra of the cervical smear samples and subjecting optomagnetic spectra images to deep neural network VGG16. Results of the developed convolutional network are visualized using confusion matrix, where columns represent labels predicted by the network and rows represent true labels (Figure 2). Correctly classified samples are shown on the diagonal.

Based on the results shown in confusion matrix, sensitivity, specificity and accuracy are calculated. Proposed deep learning model for binary classification problem achieved sensitivity of 83.8%, specificity of 78% and accuracy of 80.4%. Compared to the results obtained by classical machine learning models tested on OMIS spectra in previous studies, results reported in this study are similar, with the improvement of the sensitivity percentages. Performance results of studies using Optomagnetic fingerprint of Papanicolaou smear samples for cervical cancer detection compared to the results achieved in this study are given in Table 2.

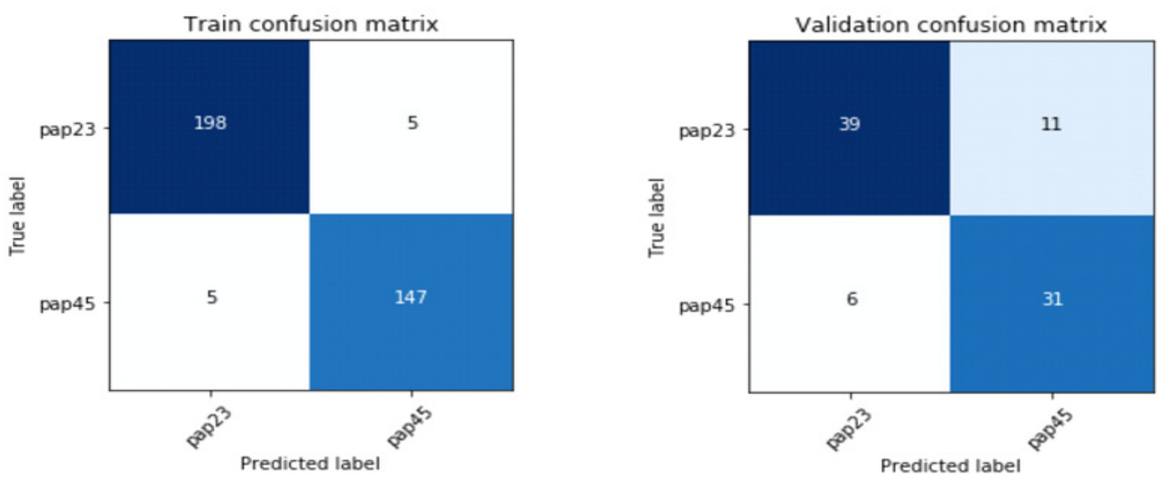

***Figure 2**. Confusion matrices for binary classification problem: Papanicolaou II and Papanicolaou III group cases labeled as “low-risk” vs. Papanicolaou IV and Papanicolaou V group cases labeled as “high risk” class*

Table 2. *Classification accuracy comparison of different machine learning models using Optomagnetic spectra parameters of stained samples and optomagnetic spectra images*

Ref.	Method	Classification problem	Sensitivity [%]	Specificity [%]
[2]	Naïve Bayes	2-class (II Pap – V Pap)	78.16	97.92
[23]	Random Forest	2-class (II, III Pap – IV, V Pap)	37.14	98.97
Proposed	CNN	2-class (II, III Pap – IV, V Pap)	83.3	78

Naïve Bayes classifier performed well in binary classification problem, with the sensitivity of 78.16% and specificity 97.92%. Still, classes used in the binary problem were made in the way to separate II Papanicolaou group samples (normal cells) from V Papanicolaou group samples (cancer cells) [2]. Since normal cells differ the most from the cancer cells, compared to low and high squamous intraepithelial lesions, it is expected that classification algorithm achieves highest scores in the normal/cancer classification problem (Table 2, Ref.2). However, if groups are formed around normal cells and low squamous intraepithelial lesion cells in one and carcinoma in situ and invasive cancer in another class, classification model will have more complex task, i.e. to find subtle differences between the degrees of cell abnormality leading to low risk case or more severe outcome. In this classification problem, Random Forest Model performed worse than proposed deep learning model (Table 2, Ref.23 and Proposed), having a higher number of false negative results leading to low sensitivity percentage of only 37.14. Deep neural network managed to overcome this problem, and achieved sensitivity of 83.3%. Specificity of proposed deep learning model is lower than specificity of other machine learning models used for cervical cancer detection based on Optomagnetic spectra, still in the range of conventional screening tests in detecting high grade squamous intraepithelial lesions and carcinoma (such as Papanicolaou test specificity of 85.58%, and HPV test specificity 54.92%) [24]. Giving that the dataset in this study was small in terms of deep learning principles, we could not ensure fine tuning of the model on more samples. Moreover, model could not be trained on more possible variations of the images themselves since the augmentation of the dataset by rotating, flipping or any other type of transformation would disrupt the position of characteristic peak position and intensity in Optomagnetic spectra.

4. CONCLUSION

This study introduces novel approach to cervical cancer detection based on Optomagnetic imaging spectroscopy and deep learning classification model. For the first time images of the Optomagnetic spectra of the Papanicolaou smears were used for binary classification problem into “low risk” and “high risk” classes. High risk class labels carcinoma in situ and invasive cervical cancer cases, while low risk class comprises of normal and low to mild grade squamous intraepithelial lesion cells. Compared to previous classification algorithms used for binary classification problems and OMIS, deep learning model gave better results in terms of sensitivity (83.3%), while the specificity was lower (79%). All performance measures are in the range of the conventional screening test measures with the expectancy that the sensitivity and specificity will improve with the increase of the data amount. Downside of using spectra images is that classical augmentation transformations cannot be done since they would interfere with the information captured inside of spectra peak position and intensity. To overcome this problem, more samples should be collected in following studies. Moreover, if we ensure that the classes are well balanced, classification into more than two classes will be possible.

5. ACKNOWLEDGEMENT

This research has been funded by the Ministry of Education, Science and Technological Development of the Republic of Serbia, through Project No. 451-03-68/2022-14/200105 from 04-Feb-2022.

Authors would also like to acknowledge Dr Milena Papic Obradovic, Clinic of Gynaecology and Obstetrics – Narodni Front, Kraljice Natalije 62, 11000 Belgrade, Serbia, and Dr Zoran Roško, Health

Clinic Zaječar, for providing us with the conventional Papanicolaou smears, without which this work would not have been successful.

6. REFERENCES

[1] Koruga, D. and Tomic, A., MYSKIN Inc, 2009. System and method for analysis of light-matter interaction based on spectral convolution. U.S. Patent Application 12/350,164.

[2] L. Matija, B. Jeftić, G. Nikolić, A. Dragičević, I. Mileusnić, J. Munćan, Dj. Koruga, *Nanophysical approach to diagnosis of epithelial tissues using Opto-magnetic imaging spectroscopy*, Nanomedicine, One Central Press (2014) 156-186.

[3] B. Jeftic, M. Papic-Obradovic, J. Muncan, L. Matija, Dj. Koruga, *Optomagnetic imaging spectroscopy application in cervical dysplasia and cancer detection: comparation of stained and unstained papanicolaou smears*, Journal of Medical and Biological Engineering, 37-6 (2017), 936-943.

[4] A. Dragicevic, L. Matija, Z. Krivokapic, I. Dimitrijevic, M. Baros, Dj. Koruga, *Classification of healthy and cancer states of colon epithelial tissues using opto-magnetic imaging spectroscopy*, Journal of Medical and Biological Engineering, 39-3 (2019), 367-380.

[5] I. Hut, B. Jeftic, L. Matija, Z. Cojbasic, Dj. Koruga, *Machine learning classification of cervical tissue liquid based cytology smear images by opto-magnetic imaging spectroscopy*, Tehnički vjesnik, 26-6 (2019) 1694-1699.

[6] B. Lisul, D. Jelovac, M. Petrovic, Z. Tepavcevic, Dj. Koruga, D. Grga, *Predictive Value of Opto-magnetic Imaging Spectroscopy in Discriminating Oral Squamous Cell Carcinoma from Non-tumor Tissue in Surgical Margins,* Journal of Medical and Biological Engineering, 39-6 (2019) 874-884.

[7] A. Bria, C. Marrocco, F. Tortorella, *Addressing class imbalance in deep learning for small lesion detection on medical images*, Computers in biology and medicine, 120 (2020) 103735.

[8] W. Qu, I. Balki, M. Mendez, J. Valen, J. Levman, P.N. *Tyrrell, Assessing and mitigating the effects of class imbalance in machine learning with application to X-ray imaging*, International journal of computer assisted radiology and surgery, 15-12, (2020) 2041-2048.

[9] S. Rajaraman, P. Ganesan, S. Antani, *Deep learning model calibration for improving performance in class-imbalanced medical image classification tasks*, PloS one, 17(1), (2022) p.e0262838.

[10] T.G. Debelee, S.R. Kebede, F. Schwenker, Z.M. Shewarega, *Deep learning in selected cancers' image analysis—a survey*, Journal of Imaging, 6-11, (2020) 121.

[11] Y. Li, L. Shen, *Skin lesion analysis towards melanoma detection using deep learning network*, Sensors, 18-2, (2018) 556.

[12] H.A. Haenssle, C. Fink, R. Schneiderbauer, F. Toberer, T. Buhl, A. Blum, A. Kalloo, A.B.H. Hassen, L. Thomas, A. Enk, L. Uhlmann, *Man against machine: diagnostic performance of a deep learning convolutional neural network for dermoscopic melanoma recognition in comparison to 58 dermatologists*, Annals of oncology, 29-8, (2018) 1836-1842.

[13] D. Bychkov, N. Linder, R. Turkki, S. Nordling, P.E. Kovanen, C. Verrill, M. Walliander, M. Lundin, C: Haglund, J. Lundin, *Deep learning based tissue analysis predicts outcome in colorectal cancer*, Scientific reports, 8-1, (2018) 1-11.

[14] L. Shen, L.R. Margolies, J.H. Rothstein, E. Fluder, R. McBride, W. Sieh, *Deep learning to improve breast cancer detection on screening mammography*, Scientific reports, 9-1, (2019) 1-12.

[15] W. Sun, B. Zheng, W. Qian, March. *Computer aided lung cancer diagnosis with deep learning algorithms,* Medical imaging, 9785 (2016) 241-248.

[16] X. Tan, K. Li, J. Zhang, W. Wang, B. Wu, J. Wu, X. Li, X. Huang, *Automatic model for cervical cancer screening based on convolutional neural network: a retrospective, multicohort, multicenter study.* Cancer Cell International, 21-1, (2021) 1-10.

[17] A. Ghoneim, G. Muhammad, M.S. Hossain, *Cervical cancer classification using convolutional neural networks and extreme learning machines*, Future Generation Computer Systems, 102 (2020) 643-649.

[18] M.M. Rahaman, C. Li, Y. Yao, F. Kulwa, X. Wu, X. Li, Q. Wang, *DeepCervix: a deep learning-based framework for the classification of cervical cells using hybrid deep feature fusion techniques*, Computers in Biology and Medicine, 136, (2021) 104649.

[19] E. Hussain, L.B. Mahanta, C.R. Das, R.K. Talukdar, *A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble*

deep convolutional neural network, Tissue and Cell, 65 (2020) 101347.

[20] Z. Alyafeai, L. Ghouti, *A fully-automated deep learning pipeline for cervical cancer classification,* Expert Systems with Applications, 141 (2020) 112951.

[21] J. Martinez-Mas, A. Bueno-Crespo, R. Martinez-Espana, M. Remezal-Solano, A. Ortiz-Gonzalez, S. Ortiz-Reina, J.P. Martinez-Cendan, *Classifying Papanicolaou cervical smears through a cell merger approach by deep learning technique,* Expert Systems with Applications, 160 (2020) 113707.

[22] K. Simonyan, A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556 (2014).

[23] B. Jeftić, L. Matija, Dj. Koruga, *New approach to detection of abnormal cervical cells*, Contemporary Materials, 10-2 (2019), 132-136.

[24] M. Kang, S.Y. Ha, H.Y. Cho, D.H. Chung, N.R. Kim, J. An, S. Lee, J.Y. Seok, J. Jeong, *Comparison of papanicolaou smear and human papillomavirus (HPV) test as cervical screening tools: can we rely on HPV test alone as a screening method? An 11-year retrospective experience at a single institution*, Journal of Pathology and Translational Medicine, 54-1, (2020) 112-118.

ПРИМЕНА АЛГОРИТМА ДУБОКОГ УЧЕЊА ЗА ДЕТЕКЦИЈУ КАРЦИНОМА ГРЛИЋА МАТЕРИЦЕ НА БАЗИ СЛИКА ОПТОМАГНЕТНИХ СПЕКТАРА

Сажетак: У циљу додатног испитивања потенцијала оптомагнетне имиџинг спектроскопије у детекцији канцера грлића материце, примењен је алгоритам дубоког учења за класификацију оптомагнетних спектара узорака. Оптомагнетни спектри садрже информацију о својствима ћелија и на основу тих својстава омогућено је разликовање нормалних ћелија, ћелија са различитим степеном дисплазије и ћелија канцера. У претходним истраживањима, показани су високи проценти тачности, сензитивности и специфичности са којима оптомагнетна имиџинг спектроскопија детектује канцер грлића материце у случају бинарне класификације, док су нешто нижи проценти постигнути у случају класификације у четири класе. У поређењу са резултатима добијеним конвенционалним алгоритмима машинског учења за класификацију, предложени алгоритам дубоког учења је дао сличне резултате по питању тачности (80%), већу сензитивност (83,3%) и компарабилну специфичност (78%).

Кључне речи: оптомагнетна имиџинг спектроскопија, канцер грлића материце, дубоко учење, конволуционе неуронске мреже.

Paper received: 24 August 2022 **Paper accepted: 12 October 2022**