Computer Methods and Programs in Biomedicine 219 (2022) 106755

Contents lists available at ScienceDirect

Computer Methods and Programs in Biomedicine

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

Classification of skin cancer using convolutional neural networks analysis of Raman spectra

Ivan A. Bratchenko[a,†,*], Lyudmila A. Bratchenko[a,†], Yulia A. Khristoforova[a], Alexander A. Moryatov[b,c], Sergey V. Kozlov[b,c], Valery P. Zakharov[a]

[a] Department of Laser and Biotechnical Systems, Samara University, 34 Moskovskoe Shosse, Samara, 443086, Russian Federation
[b] Department of Oncology, Samara State Medical University, 159 Tashkentskaya Street, Samara, 443095, Russian Federation
[c] Department of Visual Localization Tumors, Samara Regional Clinical Oncology Dispensary, 50 Solnechnaya Street, Samara, 443095, Russian Federation

ARTICLE INFO

Article history:
Received 3 November 2021
Revised 21 January 2022
Accepted 11 March 2022

Keywords:
Skin cancer
Optical biopsy
Convolutional neural networks
Projection on latent structures with discriminant analysis
Raman spectroscopy
Malignant tumor

ABSTRACT

Background and objective: Skin cancer is the most common malignancy in whites accounting for about one third of all cancers diagnosed per year. Portable Raman spectroscopy setups for skin cancer "optical biopsy" are utilized to detect tumors based on their spectral features caused by the comparative presence of different chemical components. However, low signal-to-noise ratio in such systems may prevent accurate tumors classification. Thus, there is a challenge to develop methods for efficient skin tumors classification.

Methods: We compare the performance of convolutional neural networks and the projection on latent structures with discriminant analysis for discriminating skin cancer using the analysis of Raman spectra with a high autofluorescence background stimulated by a 785 nm laser. We have registered the spectra of 617 cases of skin neoplasms (615 patients, 70 melanomas, 122 basal cell carcinomas, 12 squamous cell carcinomas and 413 benign tumors) in vivo with a portable Raman setup and created classification models both for convolutional neural networks and projection on latent structures approaches. To check the classification models stability, a 10-fold cross-validation was performed for all created models. To avoid models overfitting, the data was divided into a training set (80% of spectral dataset) and a test set (20% of spectral dataset).

Results: The results for different classification tasks demonstrate that the convolutional neural networks significantly ($p<0.01$) outperforms the projection on latent structures. For the convolutional neural networks implementation we obtained ROC AUCs of 0.96 (0.94 – 0.97; 95% CI), 0.90 (0.85–0.94; 95% CI), and 0.92 (0.87 – 0.97; 95% CI) for classifying a) malignant vs benign tumors, b) melanomas vs pigmented tumors and c) melanomas vs seborrheic keratosis respectively.

Conclusions: The performance of the convolutional neural networks classification of skin tumors based on Raman spectra analysis is higher or comparable to the accuracy provided by trained dermatologists. The increased accuracy with the convolutional neural networks implementation is due to a more precise accounting of low intensity Raman bands in the intense autofluorescence background. The achieved high performance of skin tumors classifications with convolutional neural networks analysis opens a possibility for wide implementation of Raman setups in clinical setting.

© 2022 Elsevier B.V. All rights reserved.

1. Introduction

Skin tumors are one of the most common cancer type in whites and more than a third of all diagnosed cancers are caused by skin cancer [1]. While the number of new cases of malignant melanoma and other types of skin cancer are slowly growing, the mortality rates caused by different skin cancer types are decreasing with the application of new drugs and therapeutic approaches [2]. In general, early detection of any tumor and adequate therapy and/or surgery is the best strategy in successful patient treatment. For many years, classification of skin tumors is started from a naked-eye examination and may be complemented with dermoscopic investigation of suspicious tissues [3]. At the same time, the results of dermoscopic investigation strongly depend on the experience of the clinical staff and previous experience. Deep and precise analysis of tumor morphology and other important clinical features may

* Corresponding author.
E-mail address: iabratchenko@gmail.com (I.A. Bratchenko).
† Ivan A. Bratchenko and Lyudmila A. Bratchenko should be considered joined first author.

https://doi.org/10.1016/j.cmpb.2022.106755
0169-2607/© 2022 Elsevier B.V. All rights reserved.

be analyzed only by a qualified specialist with extensive experience in tumors classification [4].

Today, "optical biopsy" [5] is one of several promising methods being utilized to detect tumors based on their spectral features caused by the comparative presence of different chemical components. Various optical modalities including terahertz spectroscopy [6], ultraviolet and visible fluorescence spectroscopy [7], Raman spectroscopy (RS) [8,9], diffuse reflectance spectroscopy [10] demonstrated a possibility to diagnose cancer, including rare tumors and tumors with atypical clinical features [11,12]. In recent decades, progress in technical equipment for skin cancer detection and diagnosis proposed several sensitive approaches using RS combined with native fluorescence or autofluorescence (AF) [8–12]. RS is a well-known technique based on inelastic scattering of light when interacting with molecular bonds of tissues chemical elements. The detected Raman spectra comprises unique bands, that carries information of certain chemical bonds in the analyzed tissue, which makes RS an outstanding approach for precise detection of malignant and benign tissues [13]. It is important to note, that in addition to the Raman shift, the scattered light contains an AF signal received after interaction with the tissues. The spectral properties of the AF signal depend on the presence of endogenous or native fluorophores in the tissues [14]. However, the AF tumor studies demonstrate lower performance of tumors classification in comparison with RS-based approaches [15].

Previously [8] we developed a portable RS system equipped a commercially available Raman probe. The signal to noise ratio (SNR) for the extraction of Raman bands with the proposed portable system was about 3 and could be treated common in comparison to the systems with deeply cooled detectors providing a high SNR Raman signal [16]. We demonstrated in vivo applicability of the developed portable system in clinical studies on the cohort of 617 cases of skin neoplasms (70 melanomas, 122 basal cell carcinomas, 12 squamous cell carcinomas and 413 benign tumors). The performed study enrolled patients examined in local clinics by general practitioners (GPs) and directed to a specialized Oncology Dispensary with skin tumors that were suspicious as malignant and showed in some way signs of malignancy. The spectra were classified with a projection on latent structures and discriminant analysis (PLS-DA). We obtained ROC AUCs of 0.75 (0.71 – 0.79; 95% CI), 0.69 (0.63 – 0.76; 95% CI), and 0.81 (0.74 – 0.87; 95% CI) for classifying: a) malignant vs benign tumors, b) melanomas vs pigmented tumors and c) melanomas vs seborrheic keratosis respectively. The positive and negative predictive values ranged from 20% to 52% and from 73% to 99% respectively. The biopsy ratio varied from 0.92:1 to 4.08:1 for a fixed sensitivity levels at 90%, 95% and 99%. The achieved performance of automatic analysis with the proposed system was higher than the accuracy of GPs and medical trainees, and was comparable or lower compared to the accuracy of trained oncological specialists [8].

As the achieved accuracies of skin cancer differentiation were also lower than the accuracies of the accuracy of skin cancer analysis with deeply cooled RS setups [8,17]. In order to improve the quality of RS-based "optical biopsy", a classification based on artificial neural networks implementation can be utilized. Artificial neural networks demonstrated a great potential in spectral data analysis including Raman and AF spectra classification [18,19]. At the same time there is a need for further investigation of artificial neural networks that can be useful in the analysis of Raman spectral datasets collected with portable devices. Few studies have demonstrated the application of convolutional neural networks (CNN) for the analysis of such spectral data [20,21], whereas datasets collected with portable devices containing spectral information about malignant and benign tissues have never been analyzed. In this paper we demonstrate the results and performance of the CNN classifier for the analysis of "optical biopsy" data of melanoma and other skin neoplasms. The analysis of the obtained spectroscopic data was performed for joint RS and AF implementation and various diagnostic tasks. We compared the CNN-based classifier accuracy with the accuracy of standard PLS-DA analysis of "optical biopsy" data as well as with an accuracy of medical staff and other approaches utilizing neural networks in the detection and classification of skin cancer.

2. Material and methods

2.1. Patients

We collected spectral data from 617 samples of skin tumors. Each patient was examined by a specialized oncologist both in local clinics and in specialized Oncology Dispensary. The oncologist made a provisional diagnosis of the studied skin neoplasm on the basis of visual examination. After the examination of oncologist, we utilized the portable RS system for the registration of neoplasms spectra. The final (ground truth) diagnosis for each tested lesion was defined on the basis of histopathology. The final spectral dataset analyzed in this study consisted of 204 malignant neoplasms (70 malignant melanomas (MM), 122 basal cell carcinomas (BCC), 12 squamous cell carcinomas (SCC)) and 413 benign neoplasms (26 dermatofibromas (DF), 62 papillomas (PP), 40 hemangiomas (HE), 113 seborrheic keratosis (SK), 170 nevi (NE), 1 cutaneous horn and 1 benign tumor of epidermal appendage). The detailed distribution of human skin lesions, including information on tumor locations, patients, etc., can be found elsewhere [16].

2.2. Spectroscopic setup and spectra processing

Details of the experimental portable Raman system that allows for simultaneous Raman and AF signal registration in the near-infrared (NIR) region can be found elsewhere [16]. Briefly, we utilize 785 nm diode laser (LuxxMaster LML-785.0RB-04, PD-LD, New Jersey, USA) for the stimulation of skin tissues Raman and AF response and collected the scattered radiation with the handheld commercially available Raman probe (RPB785, InPhotonics, Massachusetts, USA). The collected radiation was filtered and delivered to the portable spectrometer (QE65Pro, Ocean optics, Florida, USA). The laser power density on the skin was about 0.3 W/cm^2 and did not caused any damage to the skin or discomfort in patients. After spectra registration in the SpectraSuite software (Ocean optics, Florida, USA) in the 780–1000 nm region within 60 s, the spectra were preprocessed before the statistical analysis to exclude background signal and other possible noises.

2.3. Statistical analysis

In this study we analyzed the spectral data containing RS and AF parts, as previously [8] we observed no statistical difference between analysis of the Raman part and the combined Raman and AF parts from the studied skin neoplasms. This study included three dichotomic diagnostic tasks, that allowed for estimation of CNN and PLS-DA performance:

1. Discrimination of malignant (MM, BCC, SCC) and benign (DF, PP, HE, SK, NE, cutaneous horn, benign tumor of epidermal appendage) tumors;
2. Discrimination of MM and benign pigmented skin tumors (NE+SK);
3. Discrimination of MM and SK.

For the described three diagnostic tasks, all collected in vivo spectroscopic data were utilized. The estimation of the positive predictive value (PPV – number of true-positives divided to the number of true-positives and number of false-positives results),

the negative predictive value (NPV – number of true-negative divided to the number of true-negatives and number of false-negative results) and the biopsy ratio (the number of negative biopsies that are conducted for each true-positive biopsy showing skin cancer – false positives number divided by true positives number) was performed at sensitivity levels of 90%, 95%, and 99%. The results were presented using the receiver operating characteristic (ROC) curves area under the curve (AUC) with the 95% confidence interval (CI). The ROC curves were plotted using the pROC package within the *R studio* software [22]. The significance of the AUCs and comparisons between different AUCs were tested in a standard manner [23].

2.4. PLS-DA

The collected in vivo spectral data was analyzed with the help of projection on latent structures with discriminant analysis (PLS-DA) [24]. It is possible to construct a classification model to identify the skin tissue types based on the PLS-DA utilizing the most significant differences observed in the registered spectra. Each dataset was divided into a training set (80% of set) and a test set (20%). The 10-fold cross-validation was used to check the stability of the obtained classification PLS-DA models. The number of principal components (PCs) for the constructed PLS-DA models was chosen by the criterion of the local minimum of the root mean square error (RMSE) of the performed 10-fold cross-validation. Despite the fact that in many cases the increase in the PCs number leads to the minimization of RMSE and the increase in accuracy of the classification model, RMSE in cross-validation starts varying in comparison to RMSE for the obtained model, which allows for using only a number of first PCs that do not cause RMSE variations during the cross-validation process. In addition, we calculated the variable importance in projection (VIP) scores [25]. To determine the differentiation accuracy of the analyzed neoplasm types, the PLS-DA predictors were calculated to present the numeric value of the neoplasm diagnosis in the model. To implement the SIMPLS algorithm for the PLS portion of the PLS-DA method [26], a multivariate analysis was carried out using the *MDAtools* package available within the *R studio* software.

2.5. CNN

Each dataset was divided into a training set (on which the constructed model is trained – 80% of spectral dataset) and a test set (for evaluating the model and its additional training – 20% of spectral dataset). To analyze each classification model, a k-fold cross-validation was performed in the training set. The k-fold cross-validation at $k = 10$ is performed according to the following algorithm: 10% of the samples are selected from the analyzed set as a test block, 90% as a training block. The model is constructed on the training block and predicts the test block. The test block is then returned to the analyzed set. From the analyzed set, the other 10% of the samples are randomly selected as a test block. Samples selected as a test block in one iteration was not re-selected into a test block in subsequent iterations. The cycle is thus repeated 10 times. The final characteristic of the model is calculated as an average of 10 blocks. Cross-validation allows for determining the optimal parameters of the model to avoid overfitting.

Solving the posed classification tasks by means of deep learning is implemented by separate CNN models, each of which corresponds to a separate problem. The choice of the CNN architecture for recognition of current Raman dataset of skin neoplasms consisted of consecutive stages. At the first stage, the verified CNN configurations and advanced deep learning practice based on CNN were analysed. Analyzing the published results has shown that the following CNN configurations are characterized by the prospects for recognition of Raman spectra: simple sequential CNNs [27], CNNs containing the Inception module [28], CNNs with residual connections [29], ensembled CNNs [30], CNNs based on a combination of convolutional layers with recurrent layers [31]. It should be noted that for Raman spectra recognition, recurrent neural networks can be used not only in combination with CNN, but also as an independent model. For instance, Yu et al. [32] demonstrated the application of the LSTM-based recurrent network to analyze a Raman dataset in the task of marine pathogens classification. Utilization of LSTM-based recurrent networks has also been demonstrated for recognition and classification of SERS spectra of nine clinically important *Staphylococcus* species [27] and for recognition of generating synthetic Raman data in the cancer tissue classification problem [33]. One of the features of utilizing the artificial neural networks is the need for an empirical choice of the network topology and optimization of hyperparameters for the solved task. Therefore, empirical optimization of the topology and hyperparameters was then performed. As a measure of success and criteria for evaluating the tested CNN configurations were: the stability of the model, the magnitude of the error, the accuracy of the classification, the visualization of variables importance distribution. The 10-fold cross validation was used as the evaluation protocol. Thus, in the current work, network topology is chosen as a combination of a residual one-dimensional convolutional base, a recurrent layer, and a fully connected classification level. The residual one-dimensional convolutional base is a directed acyclic graph with two branches, one of which implements the residual connection, the other is a sequential combination of the convolutional layer, the pooling layer, the convolutional layer, and the pooling layer. Convolutional layers learn filter maps for different types of patterns in the input. The pooling layers extract the most prominent structures. The residual connection [34] is implemented using the convolution of the input tensor with the convolution window equal to 1 for linear reduction of the dimension. The residual one-dimensional convolutional base is followed by a recurrent layer characterized by 64-dimensionality of the output space, the fraction of the units to drop for the linear transformation equals 0.1. The *binary_crossentropy* loss function is used as a feedback signal for training the weight tensors. The CNN is trained using the *adamax* algorithm. The number of epochs was determined by the local minimum of the loss function during cross-validation. We utilized 110 epochs in Malignant vs Benign model, 70 epochs in MM vs NE + SK model, and 75 epochs in MM vs SK model. CNN analysis was carried out using the *KERAS* package within the *R studio* software [35]. The informativeness of individual predictors is visualised by means of the "vip" package within the R studio software [36]. In order to avoid a "black box" classification, model explanation and evaluating the variables contribution to the constructed model are required. In the current work, an algorithm of calculating permutation-based variable importance was used. The permutation-based method for assessing variable importance in constructed model is used in traditional machine learning; the perspective of its application and its adaptation for analyzing the deep learning models are described in Refs [37–39]. Fig. 1 demonstrates a schematic diagram of the proposed approach for analysing the Raman spectral dataset and corresponding structural features using the proposed CNN.

3. Results

Table 1 demonstrates the achieved performance for the classification models based on the PLS-DA and CNN implementation. The application of CNN leads to the overall increase in skin malignant and benign tumors classifications. The achieved improvement in ROC AUC for CNN is significant ($p<0.01$) in comparison to the PLS-DA implementation for all the three discussed tasks. For discrim-

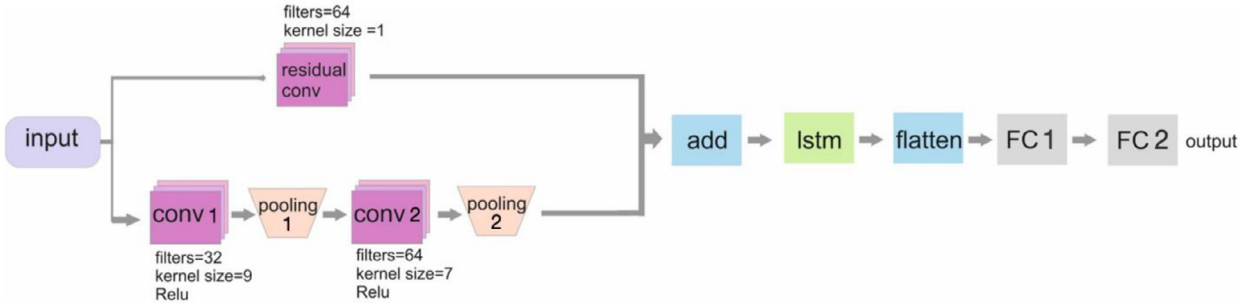

Fig. 1. Schematic diagram of the architecture of one-dimensional CNN for the recognition of Raman spectra and classification of skin tissues where: **input** is an input layer with output shape (912, 1); **conv1** is a one-dimensional convolutional layer with padding = "same" and output shape (912, 32); **pooling1** is an average pooling layer with pool_size = 4, strides = 4 and output shape (228, 32); **conv2** is a one-dimensional convolutional layer with padding = "same" and output shape (228, 64); **pooling2** is an average pooling layer with pool_size = 4, strides = 4 and output shape (57, 64); **residual conv** is a one-dimensional convolutional layer with strides = 16 and output shape (57, 64); **add** is a layer that adds residual tensor and convolutional base, output shape (57,64); **lstm** is a recurrent layer; **flatten** is a flatten layer; **FC1** is a fully connected layer with units = 128 and *Relu* activation function; **FC2** is a fully connected layer with units = 1 and *sigmoid* activation function.

Table 1
Skin tissue classification for three diagnostic tasks – 1) Malignant (MM, BCC, SCC) vs Benign (DF, PP, HE, SK, NE, cutaneous horn, benign tumor of epidermal appendage); 2) MM vs NE+SK; 3) and MM vs SK – with PLS-DA and CNN classifiers.

	ROC AUC (95% CI)	
Diagnosis classification task	PLS-DA	CNN
Malignant (n = 204) vs benign neoplasms (n = 413)	0.75 (0.71 – 0.79)	0.96 (0.94 – 0.97)
MM (n = 70) vs NE+SK (n = 283)	0.66 (0.59 –0.74)	0.90 (0.85 –0.94)
MM (n = 70) vs SK (n = 113)	0.81 (0.74 –0.87)	0.92 (0.87 – 0.97)

Table 2
Summary of diagnostic parameters derived from ROCs according to various levels of sensitivity in skin cancer discrimination (PPV, ratio of true-positives to the total of true-positives and false-positives; NPV, ratio of true-negative to the total of true-negatives and false-negative; and "biopsy ratio", the number of negative biopsies that are conducted for each true-positive biopsy showing skin cancer).

Diagnosis classification task	Sensitivity (95% CI)	Specificity (95% CI)	PPV	NPV	Biopsy ratio
PLS-DA					
Malignant (n = 204) vs	0.99 (0.97 – 1.00)	0.03 (0.01 – 0.08)	0.34	0.86	1.98:1
benign neoplasms	0.95 (0.91 – 0.98)	0.14 (0.06 – 0.30)	0.35	0.85	1.84:1
(n = 413)	0.90 (0.85 –	0.32 (0.15 – 0.45)	0.40	0.85	1.53:1
MM (n = 70) vs NE+SK	0.99 (0.91 – 1.00)	0.06 (0.04 – 0.11)	0.21	0.96	3.84:1
(n = 283)	0.95 (0.89 – 1.00)	0.10 (0.04 – 0.29)	0.21	0.89	3.82:1
	0.90 (0.79 – 0.96)	0.26 (0.10 – 0.41)	0.23	0.89	3.32:1
MM (n = 70) vs SK	0.99 (0.93 – 1.00)	0.13 (0.07 – 0.29)	0.41	0.96	1.41:1
(n = 113)	0.95 (0.87 – 1.00)	0.27 (0.210 – 0.45)	0.45	0.90	1.23:1
	0.90 (0.80 – 0.97)	0.40 (0.25 – 0.66)	0.48	0.87	1.08:1
CNN					
Malignant (n = 204) vs	0.99 (0.97 – 1.00)	0.64 (0.26 – 0.78)	0.58	0.99	0.74:1
benign neoplasms	0.95 (0.91 – 0.98)	0.80 (0.73 – 0.88)	0.71	0.97	0.42:1
(n = 413)	0.90 (0.85 – 0.94)	0.89 (0.82 – 0.94)	0.80	0.94	0.25:1
MM (n = 70) vs NE+SK	0.99 (0.91 – 1.00)	0.10 (0.06 – 0.50)	0.21	0.97	3.69:1
(n = 283)	0.95 (0.89 – 1.00)	0.50 (0.26 – 0.68)	0.32	0.97	2.12:1
	0.90 (0.79 – 0.96)	0.65 (0.48 – 0.87)	0.39	0.96	1.59:1
MM (n = 70) vs SK	0.99 (0.93 – 1.00)	0.11 (0.05 – 0.20)	0.41	0.94	1.46:1
(n = 113)	0.95 (0.87 – 1.00)	0.19 (0.07 – 0.89)	0.42	0.86	1.38:1
	0.90 (0.80 – 0.97)	0.83 (0.17 – 0.97)	0.77	0.93	0.30:1

inating Malignant and Benign neoplasms, the achieved ROC AUC with CNN implementation is as high as 0.96 (0.94 – 0.97; 95% CI). For the two other dichotomy classification tasks, the ROC AUCs are slightly lower. Table 2 demonstrates the exact values of specificity, PPV, NPV and Biopsy ratios for the fixed values of sensitivity in cases of PLS-DA and CNN spectral dataset analysis. Once again, the CNN shows the better values of classification performance in comparison to the PLS-DA. The values of PPV, NPV and Biopsy ratios achieved with CNN outperform the similar values obtained with PLS-DA almost in all cases (however, in some cases they demonstrate almost the same values).

Fig. 2a demonstrates the examples of raw registered spectra (containing both RS and AF signals), while Fig. 2b shows the extracted Raman signal. It is of interest that the AF of skin tissues stimulated by a 785 nm laser is a non-linear decreasing function whereas the low-intensity Raman signal is a set of separate tiny bands combined into wider peaks. Due to a complex biochemical composition of skin tissues, the precise determination of exact chromophores is quite complicated and, for skin tissue classification, a simple analysis of separate AF peaks or Raman bands is ineffective [16]. Therefore, the registered spectra were analyzed using the PLS-DA and the CNN-based approaches. Figs. 3 demonstrate VIPs for all constructed PLS-DA and CNN models. In the PLS-DA case, the VIP is a weighted sum of loadings and merely highlights all spectral features from all loadings obtained in the classification model. Fig. 3a shows that in the case of PLS-DA, the VIPs utilize the data from narrow bands in the 1125–1185, 1437–1470, and 1645–1681 cm^{-1} regions (related to the Raman spectra), while in the

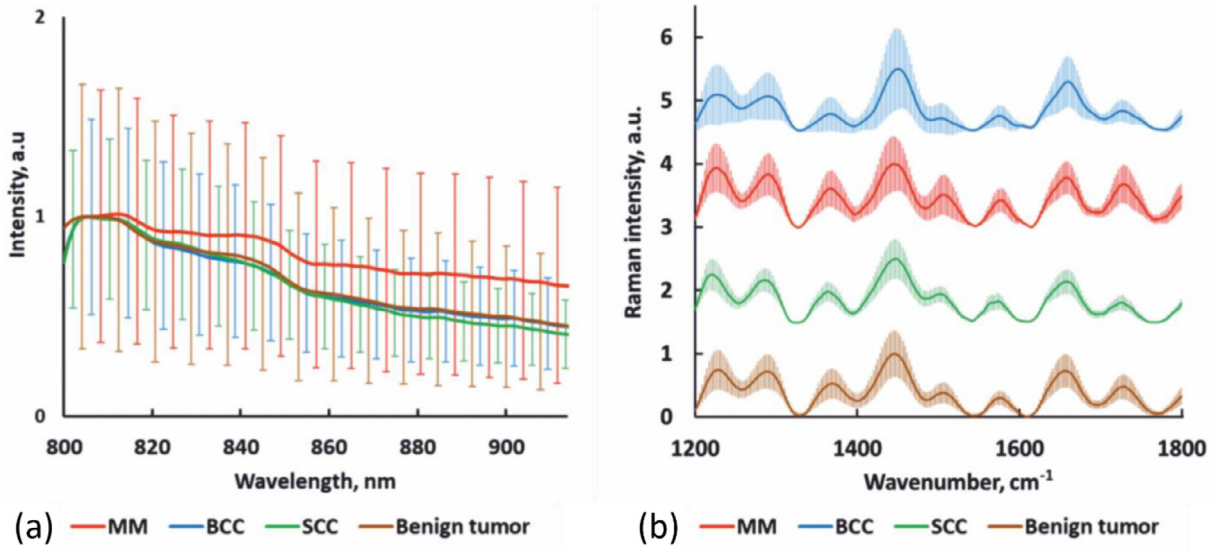

Fig. 2. Mean values and standard deviation for: (a) raw registered spectra (Raman signal and autofluorescence) and (b) extracted Raman signal of skin tissues stimulated by a 785 nm laser.

low-wavelength region (300–1000 cm^{-1}), the PLS-DA captures only AF features. The PLS-DA VIPs utilize the data based on the curvature of wide AF bands in the analyzed spectra. The VIPs of mixed Raman and AF spectra analyzed in this study are a combination of Raman and AF properties where the narrow low-intensity Raman bands are masked with wide high-intensity AF peaks. For the analysis of pigmented lesions (MM vs pigmented benign and MM vs SK), strong AF dominates weak Raman bands and the PLS-DA analysis utilizes mostly wide AF features. It is important to note, that for all classification tasks the PLS-DA utilized from 1 to 4 principal components (PCs). The loadings of the obtained PCs for each constructed model highlight the most informative spectral bands that could be useful in tissue classification. The CNN VIPs also highlight the most informative bands that provide the best performance for the obtained classifier. One can see that CNN utilizes more Raman bands in the analysis in comparison to PLS-DA. For Malignant vs Benign neoplasms, the differentiation CNN not only uses information from the bands in the 1000 – 1800 cm^{-1} region, but also information from the bands in the 380 – 1000 cm^{-1} region. To distinguish between MM and Pigmented lesions, the CNN utilizes the exact Raman bands instead of wide AF information. The PLS-DA VIPs highlight the most informative Raman bands near 1450 and 1650 cm^{-1} and primarily use information from the AF region near 470–1050 $cm^{-1}$for Malignant and Benign neoplasms discrimination. To distinguish between MM and Pigmented lesions (NE+SK), the PLS-DA utilizes the bands near 1300 and 1650 cm^{-1}, and near 570, 720 and 1000 cm^{-1} for MM and SK discrimination. However, in case of the MM and SK discrimination, the mentioned bands are strongly affected by a high AF signal.

4. Discussion

Applying CNN to analyze Raman spectra leads to increased performance of the proposed portable Raman setup. The ROC AUCs values have increased by 0.11 to 0.24 in MM vs SK and MM vs Pigmented neoplasm discriminations respectively; biopsy ratios decreased from 1.53:1 to 0.25:1 in case of Malignant vs Benign neoplasm classifications. Thus, CNN analysis demonstrates higher efficiency in comparison to PLS-DA.

Moreover, applying CNN to the analysis of Raman spectra leads to a more precise evaluation of experimental data in comparison to PLS-DA. Fig. 3 demonstrates that the obtained VIPs for CNN application utilizes information from the Raman bands both in the low-wavenumber region (up to 1000 cm^{-1}) and in the high-wavenumber region (over 1000 cm^{-1}). Figs. 3b and c demonstrate that PLS-DA analysis utilizes information from the wide AF signal to discriminate the pigmented neoplasms. Only in the case of Malignant and Benign neoplasm discrimination, PLS-DA is able to use Raman data. In contrast, CNN helps to track and analyze Raman bands across the whole spectra for all the three discussed classification tasks. Typically, Raman spectra are contaminated with the AF signal that masks low-intensity Raman bands, whereas PLS-DA analysis is able to utilize primarily just the AF features from the spectra. At the same time, CNN provides an ability to analyze both the AF and the Raman components for a more precise classification of skin neoplasms. It is important to note, that the Raman signal is basically distinguished from the AF signal using different techniques [16,40]. However, in some cases the AF signal carries important information for skin tissue classification and its elimination affects the classification accuracy [41]. The discussed algorithm of CNN analysis of skin spectral data does not require clearing the recorded spectra of the AF signal and automatically utilizes the information both from the Raman and the AF parts of the spectra. This feature makes CNN analysis of Raman spectral data an outstanding instrument that provides a higher accuracy of analysis.

If we compare the accuracy of CNN spectral analysis of Raman spectra with the accuracy of other optical approaches to skin neoplasms classification, we may see that the proposed approach appears favorable in many aspects. For example, application of deeply-cooled Raman systems provides a ROC AUC of 0.91 (0.89 – 0.93; 95% CI) [42] for distinguishing between Benign and Malignant tumors with PLS-DA, while the proposed approach helps to achieve ROC AUC of 0.96 (0.94 – 0.97; 95% CI). The obtained values of ROC AUCs demonstrate that even portable Raman systems can compete with high SNR Raman systems (with deeply-cooled detectors) when the spectral data is analyzed with a CNN. Note that Zhao et al. [42] demonstrated further increase in the ROC AUCs values by adding the patients' demographic data to the analysis. Such data also can be added to the CNN analysis to increase the performance of skin cancer detection.

The biopsy ratios for differentiating between Malignant vs Benign neoplasms and between MM vs SK demonstrated by the CNN are as low as 0.25:1 and 0.30:1 respectively. Our study has shown that the biopsy ratios are quite low in comparison to the

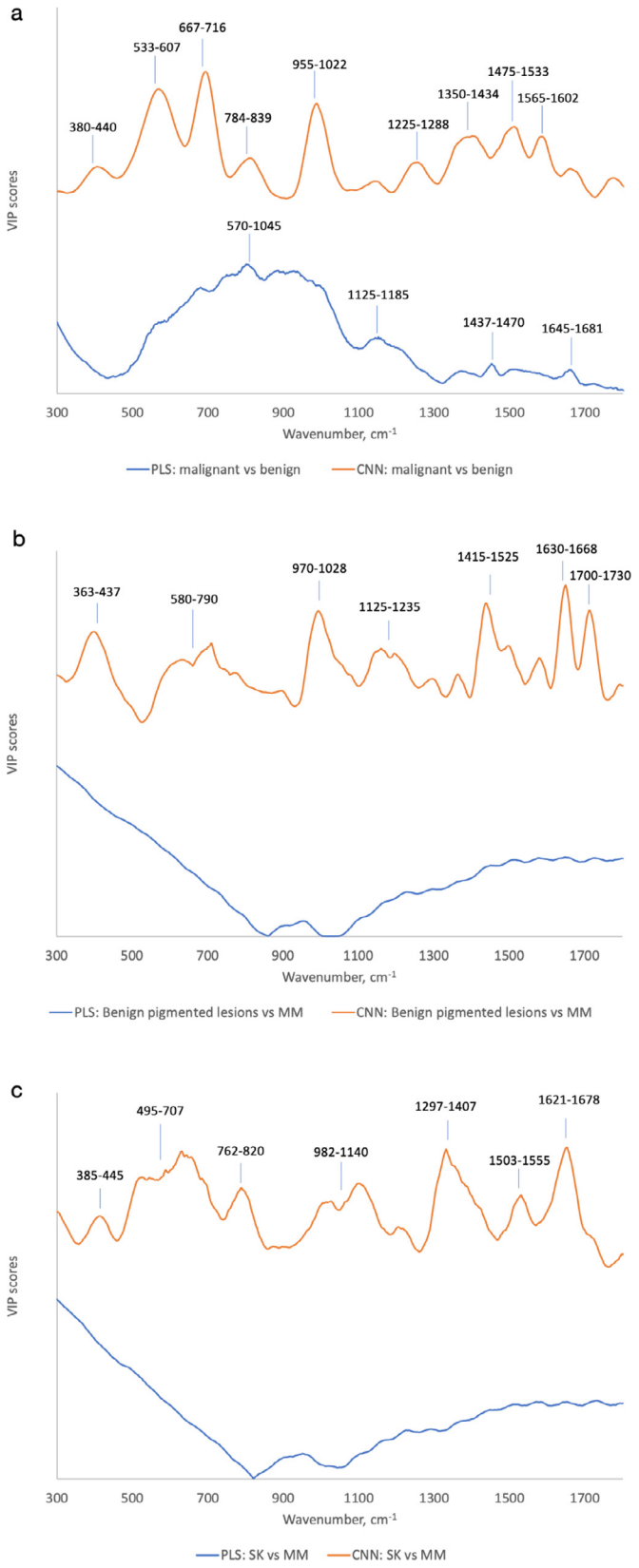

Fig. 3. VIPs for classification models: (a) Malignant vs Benign, (b) MM vs Pigmented (NE+SK), (c) MM vs SK; MM – malignant melanoma, NE – nevus, SK - seborrheic keratosis.

biopsy ratios when diagnosing melanomas by visual inspection by experienced dermoscopy users (biopsy ratio near 8:1) and other health care professionals (over 30:1) [43–46]. At the same time, we should note that further studies with appropriate design are required for the correct comparison of dermatologists biopsy ratios with biopsy ratios provided by Raman setups combined with CNN. Multispectral imaging of skin cancers with the *MelaFind* device demonstrated a 10.8:1 biopsy ratio in the evaluation of pigmented lesions in a large multicenter prospective study [47]. Our results were further compared by means of dermoscopy and spectroscopy-based computer-assisted diagnosis (CAD) techniques. The Cochrane review [48] of CAD systems applications for diagnosing skin cancer in adults demonstrated the mean sensitivity of 90% (84% – 94%; 95% CI) with the mean specificity of 74% (64%–83%; 95% CI) for the dermoscopy-based CAD systems where the biopsy ratio was 1.14:1. For the spectroscopy-based CAD systems, the mean sensitivity was 93% (84% – 97%; 95% CI) and the mean specificity was 44% (25% – 65%; 95% CI) with the biopsy ratio of 2.42:1. The values of sensitivity and specificity demonstrated by CAD systems are slightly lower than the corresponding values in our study, however, the biopsy ratios in case of CNN implementation are appreciably better than the biopsy ratios obtained by the existing CAD systems. Recent in vivo Raman studies of pigmented skin lesions provided 4.1:1 biopsy ratio with 100% sensitivity [49], however only 7 melanomas were studied in this survey.

Further studies may compare the obtained results with the analysis accuracy performed by medical specialists. The PLS-DA analysis for the fixed sensitivity of 90%, 85%, 81%, 62% provided a 32%, 51%, 58%, 77% specificity in malignant vs benign neoplasm discrimination and a 40%, 60%, 64%, 82% specificity in MM vs SK discrimination. Thus, the PLS-DA is able to classify skin neoplasms with a mean accuracy higher than the accuracy provided by GPs and trainees, and with a comparable or lower accuracy compared to trained dermatologists and experts (based on the data published in *Cochrane review* [50]) [8]. CNN provides significantly ($p<0.01$) higher values of the obtained specificities for the same values of sensitivity: 89%, 94%, 97%, 100% in case of Malignant vs Benign neoplasm discrimination and 83%, 89%, 95%, 100% in case of MM vs SK discrimination. Thus, utilizing CNN for Raman spectra analysis helps to increase the sensitivity and specificity of skin tumors detection.

Fig. 4 demonstrates the achieved ROC curves for CNN discrimination between Malignant and Benign tumors as well as the ROC curves obtained in different studies and the accuracies provided by medical specialists. The presented results show that CNN-based analysis of Raman spectra can produce ROC AUCs that outperform or at least equal to the results provided by medical specialists. Moreover, the proposed CNN analysis of the Raman spectral data even outperforms the approaches based on CNN analysis of skin tumor images. In the study by Haenssle et al. [51], the CNN algorithm classified 60 cases of benign lesions (30 NE, 10 SK, 10 solar lentigo, 5 angiomas, 5 dermatofibromas) vs 40 cases of malignant and premalignnat lesions (15 MM, 10 BCC, 5 SCC, 5 actinic keratosis, 5 Bowen's disease). All 100 cases included pigmented/non-pigmented and melanocytic/non-melanocytic skin lesions. The CNN approach for tumor image analysis revealed a sensitivity, specificity, and ROC AUC of 95% (84–99%; 95% CI), 77% (65–86%; 95% CI), and 0.92 (0.87–0.97; 95% CI). In addition, the study by Haenssle et al. provided the results of skin tumors classification provided by 96 trained dermatologists with different experience (beginner, skilled, expert). Dermatologists were asked to indicate their dichotomous diagnosis (malignant/premalignant, benign) on the basis of dermatoscopic image of the lesion. The dermatologists' mean sensitivity of 89% (87–91%; 95% CI) and specificity of 81% (79–83%; 95% CI) are presented in Fig. 3.

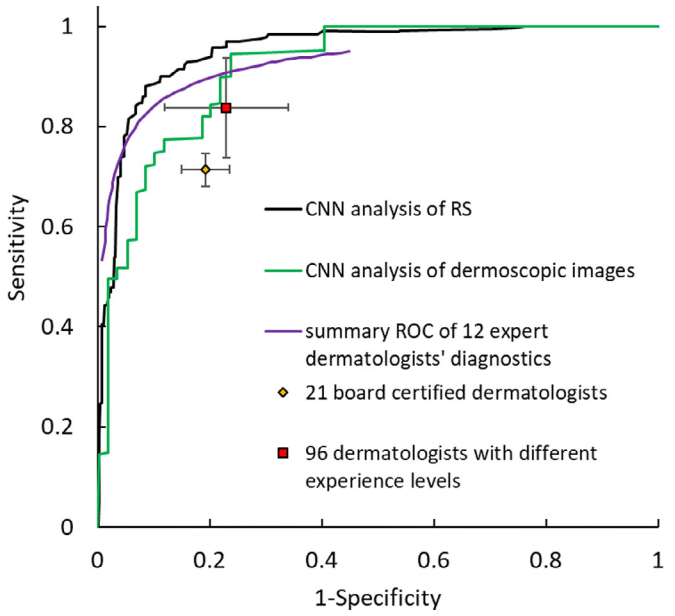

Fig. 4. The ROC curves and diagnostics performance of different medical specialists and proposed CNN analysis of Raman spectra (RS) with ROC AUC of 0.96 (0.94 – 0.97; 95% CI); CNN analysis of dermoscopic images performed in Ref. [51] to classify 60 cases of benign lesions vs 40 cases of malignant and premalignant lesions with ROC AUC of 0.918 (95% CI 0.866–0.970); summary ROC of 12 expert dermatologists' diagnostics presented in a *Cochrane review* [53]. The diagnostics performance of 21 board certified dermatologists to classify 71 malignant vs 40 benign lesions was performed in Ref. [52]. The diagnostic performance of 96 dermatologists with different experience levels: 17 beginners (<2 years of experience), 29 skilled (2 – 5 years of experience), and 40 experts (>5 years of experience) is presented in Ref. [51].

For the board certified dermatologists results (see Fig. 4), the data were adopted from the study by Esteva et al. [52]. The averaged diagnostic sensitivity and specificity of the 21 board certified dermatologists to distinguish between 71 Malignant tumors and 40 Benign tumors were 71% (66–76%; 95% CI) and 81% (77–86%; 95% CI) respectively. The summary ROC plot for in-person visual inspection evaluations based on the diagnostics by the trained dermatologists from different clinical studies was adopted from the *Cochrane review* [53]. Using the data presented in Fig. 4, the proposed CNN-based approach to the neoplasm Raman spectra analysis demonstrates performance higher rates of than those by dermatologists in general. Thus, after extensive tests, such approach can be implemented in clinical practice.

If we compare the achieved results with other neural network-based approaches of cancer detection, we can see that most efforts in this field were focused on the analysis of tumor images. Numerous approaches were suggested to compete the dermatologists and in many cases the demonstrated effectiveness of the neural network classifiers outperformed the medical specialists [54,55]. Some of the well-known studies based on pretrained Google's Inception v3 CNN showed the ROC AUCs for skin cancer detection as high as 0.91 to 0.96 [52] and 0.918 (0.866–0.970; 95% CI) [51]. At the same time, neural network analysis of Raman spectral data for skin tumor classification also demonstrates particularly high accuracy of 94.2% ± 3.1% (mean ± SD) for Malignant and Benign+Normal tissue discrimination [56]. Nevertheless, in contrast to our study, Sigurdsson et al. utilized spectral data in the 2000 to 3500 cm^{-1} region and tested their approach on a dataset of 222 samples (48 BCC, 21 MM, 89 normal skin, 23 NE, 23 SK). Possible increase in the diagnostic performance can be achieved utilizing joint analysis of skin tumor images and spectral response with hyperspectral setups [57]. However, these setups are quite costly (utilizing Raman modalities) and require further studies to prove their applicability in skin cancer classification. Another possibility to increase the analysis accuracy of CNN-based Raman spectra classification is interpretation of false classification results [52]. As the proposed algorithm highlights the Raman bands utilized in the analysis, it is possible to add or exclude specific bands from the analysis and possibly achieve better classification accuracy. This approach should be tested in future studies.

In recent decades the application of CNN in cancer detection and screening have been widely increased as there is a need to analyze complex biomedical data. Such applications included analysis of cancer images [58], nodules malignancy assessment [59], prediction of cancer type based on gene expression [60], prediction of cancer survival [61] and many others applications [62–66]. Analysis of the recent publications dedicated to CNN application in cancer detection and classification showed an opportunity to significantly increase the accuracy of performed analysis in comparison to standard statistical methods. Thus, it is not surprising, that in case of Raman spectra analysis CNN may outperform standard approaches as PLS-DA. CNN is able to catch hidden in the complex information (as mixed Raman and AF spectra) signs of specific target groups, and utilize such data in precise classification of different groups (cancer types in the considered task). At the same time, a special attention should be paid to the analysis of the obtained data during groups classification to avoid overestimation of the models and avoid utilization of noises in the models. In case of PLS-DA application such important information may be found in VIP distributions, as VIP represent weighted sum of PCs. In the current study, shape of the VIPs totally depends on the composition of Raman bands and AF peaks in the registered spectral database, and therefore we are sure, that the constructed classification models do not utilize random noises. In case of CNN implementation, we utilize well-known approach based on VIPs estimations using model-specific approach [28]. Such approach helped us to achieve better results in comparison to PLS-DA, as VIP analysis in CNN is able to highlight very weak Raman bands masked by huge AF signal. Once again the VIPs shape demonstrate (Fig. 3), that the utilized in the CNN-based classification model data is based only on specific spectral bands, and do not contain random noises. Thus, the obtained results of skin tissues CNN classification based on Raman spectra analysis are solid and stable, and may be applied in future clinical applications.

In summation, according to the obtained results, the discussed portable Raman systems in combination with CNN analysis can be applied for skin neoplasm detection to improve the accuracy of provisional and clinical diagnosis. The proposed technique demonstrates better or comparable results with respect to those by the trained medical staff. Further studies may include incorporating patients' demographics into the analysis [42] or applying more complex networks including networks ensembles [67]. The proposed system should be evaluated in large-scale multicenter trials to prove its applicability in clinical settings or to demonstrate joint applicability of dermoscopy analysis (preferably complemented with neural network-based analysis [54]) and portable Raman systems. Note, that the latest studies demonstrated that portable RS systems may compete with deeply cooled RS systems in the detection of specific components in tissues [68], so the application of portable RS setups maybe very effective in some clinical applications.

5. Conclusions

In this study we demonstrated that CNN analysis significantly outperforms PLS-DA based classification of Raman signals of skin tumors. For CNN implementation we obtained ROC AUCs of 0.96 (0.94 – 0.97; 95% CI), 0.90 (0.85–0.94; 95% CI), and 0.92 (0.87 – 0.97; 95% CI) for classifying a) malignant vs benign tumors, b)

melanomas vs pigmented tumors and c) melanomas vs seborrheic keratosis respectively. The performance of CNN-based tumor classification with the proposed portable Raman device is comparable to the accuracy of the Raman devices with deeply-cooled detectors. The accuracy of automatic analysis with the proposed CNN-based approach is higher than the accuracy provided by GPs and trainees, and is comparable or higher with respect to the accuracy of trained dermatologists. The increased accuracy of CNN implementation is a result of a more precise (compared to PLS-DA) accounting of low intensity Raman bands in the intense AF background. The classification performance can be further improved by using more complex neural network algorithms of Raman spectra analysis (e.g. with neural networks ensembles [67]) or by adding complementary information to the analysis (such as patients' demographics [42]). Important to note, that the proposed approach may be combined with other optical techniques, such as dermoscopy-based systems, for more precise tumors detection. Proposed Raman systems complemented with the CNN can be tested for their ability to detect other illness (track diabetes [69,70] or find kidney disorders [71]) and to find tumors in other organs. However, such CNN-based novel approaches need to be tested in future large multicenter trials.

Declaration of Competing Interest

The authors declare no conflict of interests.

Statement of ethical approval

The protocols of the in vivo tissue diagnostics were approved by the ethical committee of Samara State Medical University (Samara Region, Samara, Russia, protocol No 132, 29 May 2013, the clinical studies fall within *The Code of Ethics of a Doctor of Russia*, approved at the 4th conference of the *Russian Medical Association*, and within the *World Medical Association Declaration of Helsinki*). All patients were at least 18 years old. Written informed consent was acquired from all patients before the in vivo study.

Acknowledgments

This study was partly supported by RFBR grant № 20–32–90040 (L.A. Bratchenko) and joint grant by RFBR and Bulgarian NSF #KP06-Russia/19/2019 "Multivariative Raman and fluorescence diagnosis of cutaneous tumors" № 19–52–18001 (I.A. Bratchenko, Yu.A. Khristoforova, A.A. Moryatov, S.V. Kozlov, V.P. Zakharov).

References

[1] R.L. Siegel, K.D. Miller, A. Jemal, Cancer statistics, 2020, CA Cancer J. Clin. 70 (1) (2020) 7–30, doi:10.3322/caac.21590.
[2] A. Jemal, R. Siegel, J. Xu, Cancer statistics, 2010, CA Cancer J. Clin. 60 (5) (2010) 277–300, doi:10.3322/caac.20073.
[3] G. Argenziano, S. Puig, I. Zalaudek, F. Sera, R. Corona, M. Alsina, F. Barbato, C. Carrera, G. Ferrara, A. Guilabert, D. Massi, J.A. Moreno-Romero, C. Munoz-Santos, G. Petrillo, S. Segura, H.P. Soyer, R. Zanchini, J. Malvehy, Dermoscopy improves accuracy of primary care physicians to triage lesions suggestive of skin cancer, J. Clin. Oncol. 24 (12) (2016) 1877–1882, doi:10.1200/JCO.2005.05.0864.
[4] A.M. Anderson, M. Matsumoto, M.I. Saul, A.M. Secret, L.K. Ferris, Accuracy of skin cancer diagnosis by physician assistants compared with dermatologists in a large health care system, JAMA Dermatol. 154 (5) (2018) 569–573, doi:10.1001/jamadermatol.2018.0212.
[5] R.R. Alfano, Y. Pu, Optical biopsy for cancer detection, in: H. Jelinkova (Ed.), Laser For Medical Applications, Woodhead Publishing Limited, Cambridge, 2013, pp. 325–367, doi:10.1533/9780857097545.3.325.
[6] K.I. Zaytsev, K.G. Kudrin, V.E. Karasik, I.V. Reshetov, S.O. Yurchenko, In vivo terahertz spectroscopy of pigmentary skin nevi: pilot study of non-invasive early diagnosis of dysplasia, Appl. Phys. Lett. 106 (5) (2015) 053702, doi:10.1063/1.4907350.
[7] E. Borisova, T. Genova, V. Mircheva, P. Troyanova, I. Bratchenko, L. Bratchenko, Y. Khristoforova, V. Zakharov, I. Lihacova, A. Lihacovs, J. Spigulis, Multispectral fluorescence detection of pigmented cutaneous tumours, in: Proceedings of the Third International Conference Biophotonics Riga, 11585, 2020, doi:10.1117/12.2581967.
[8] I.A. Bratchenko, L.A. Bratchenko, A.A. Moryatov, Y.A. Khristoforova, D.N. Artemyev, O.O. Myakinin, A.E. Orlov, S.V. Kozlov, V.P. Zakharov, In vivo diagnosis of skin cancer with a portable Raman spectroscopic device, Exp. Dermatol. 30 (5) (2021) 652–663, doi:10.1111/exd.14301.
[9] E. Cordero, I. Latka, C. Matthäus, I.W. Schie, J. Popp, In-vivo Raman spectroscopy: from basics to applications, J. Biomed. Opt. 23 (7) (2018) 071210, doi:10.1117/1.JBO.23.7.071210.
[10] V. Colas, W. Blondel, G. Khairallah, C. Daul, M. Amouroux, Proposal for a skin layer-wise decomposition model of spatially-resolved diffuse reflectance spectra based on maximum depth photon distributions: a numerical study, Photonics 8 (2021) 444, doi:10.3390/photonics8100444.
[11] L.A. Bratchenko, Y.A. Khristoforova, A.A. Moryatov, I.A. Bratchenko, Raman spectroscopy based diagnosis of dermatofibrosarcoma protuberans: case report, Photodiagnosis Photodyn. Ther. 35 (2021) 102351, doi:10.1016/j.pdpdt.2021.102351.
[12] I.A. Bratchenko, Y.A. Khristoforova, L.A. Bratchenko, A.A. Moryatov, S.V. Kozlov, E.G. Borisova, V.P. Zakharov, Optical biopsy of amelanotic melanoma with raman and autofluorescence spectra stimulated by 785 nm laser excitation, J. Biomed. Photon. Eng. 7 (2) (2021) 020308, doi:10.18287/JBPE21.07.020308.
[13] X. Feng, M.C. Fox, J.S. Reichenberg, F.C.P.S. Lopes, Biophysical basis of skin cancer margin assessment using Raman spectroscopy, Biomed. Opt. Express 10 (1) (2019) 104–118, doi:10.1364/BOE.10.000104.
[14] T. Vo-Dinh, B. Cullum, Fluorescence Spectroscopy for Biomedical Diagnostics, in: T. Vo-Dinh (Ed.), Biomedical Photonics Handbook, CRC Press, Boca Raton, 2003, pp. 739–788, doi:10.1201/9780203008997-38.
[15] S. Wachsmann-Hogiu, T. Weeks, T. Huser, Chemical analysis in vivo and in vitro by Raman spectroscopy – from single cells to humans, Curr. Opin. Biotechnol. 20 (1) (2009) 63–73, doi:10.1016/j.copbio.2009.02.006.
[16] Y.A. Khristoforova, I.A. Bratchenko, O.O. Myakinin, D.N. Artemyev, A.A. Moryatov, A.E. Orlov, S.V. Kozlov, V.P. Zakharov, Portable spectroscopic system for in vivo skin neoplasms diagnostics by Raman and autofluorescence analysis, J. Biophotonics 12 (2019) e201800400, doi:10.1002/jbio.201800400.
[17] H. Lui, J. Zhao, D. McLean, H. Zeng, Real-time Raman Spectroscopy for In Vivo Skin Cancer Diagnosis, Cancer Res. 72 (10) (2012) 2491–2500, doi:10.1158/0008-5472.CAN-11-4061.
[18] M. Yu, H. Yan, J. Xia, L. Zhu, T. Zhang, Z. Zhu, X. Lou, G. Sun, M. Dong, Deep convolutional neural networks for tongue squamous cell carcinoma classification using Raman spectroscopy, Photodiagnosis Photodyn. Ther. 26 (2019) 430–435, doi:10.1016/j.pdpdt.2019.05.008.
[19] L.-.W. Shang, D.-.Y. Ma, J.-.J. Fu, Y.-.F. Lu, Y. Zhao, Fluorescence imaging and Raman spectroscopy applied for the accurate diagnosis of breast cancer with deep learning algorithms, Biomed. Opt. Express 11 (7) (2020) 3673–3683, doi:10.1364/BOE.394772.
[20] N. Borodinov, S. Neumayer, S.V. Kalinin, O.S. Ovchinnikova, R.K. Vasudevan, S. Jesse, Deep neural networks for understanding noisy data applied to physical property extraction in scanning probe microscopy, NPJ Comput. Mater. 5 (1) (2019) 25, doi:10.1038/s41524-019-0148-5.
[21] C.-.S. Ho, N. Jean, C.A. Hogan, L. Blackmon, S.S. Jeffrey, M. Holodniy, N. Banaei, A.A.E. Saleh, S. Ermon, J. Dionne, Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning, Nat. Commun. 10 (2019) 4927, doi:10.1038/s41467-019-12898-9.
[22] J.A. Hanley, B.J. McNeil, A method of comparing the areas under receiver operating characteristic curves derived from the same cases, Radiology 148 (3) (1983) 839–843, doi:10.1148/radiology.148.3.6878708.
[23] D.M. Haaland, E.V. Thomas, Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information, Anal. Chem. 60 (1988) 1193–1202, doi:10.1021/ac00162a020.
[24] X. Robin, N. Turck, A. Hainard, N. Tiberti, F. Lisacek, J. Sanchez, M. Müller, pROC: an open-source package for R and S+ to analyze and compare ROC curves, BMC Bioinformatics 12 (2011) 77, doi:10.1186/1471-2105-12-77.
[25] I.G. Chong, C.H. Jun, Performance of some variable selection methods when multicollinearity is present, Chemometr. Intell. Lab. Syst. 78 (1–2) (2005) 103–112, doi:10.1016/j.chemolab.2004.12.011.
[26] Kucheryavskiy S., Mdatools For R, R Package Version 0.9.4. 2019. www.mdatools.com.
[27] J.-.W. Tang, Q.-.H. Liu, X.-.C. Yin, Y.-.C. Pan, P.-.B. Wen, X. Liu, X.-.X. Kang, B. Gu, Z.-.B. Zhu, L. Wang, Comparative analysis of machine learning algorithms on surface enhanced Raman spectra of clinical staphylococcus species, Front. Microbiol 12 (2021) 696921, doi:10.3389/fmicb.2021.696921.
[28] J. Hu, Y. Zou, B. Sun, X. Yu, Z. Shang, J. Huang, S. Jin, P. Liang, Raman spectrum classification based on transfer learning by a convolutional neural network: application to pesticide detection, Spectrochimica Acta Part A 265 (2022) 120366, doi:10.1016/j.saa.2021.120366.
[29] C.-.S. Ho, N. Jean, C.A. Hogan, L. Blackmon, S.S. Jeffrey, M. Holodniy, N. Banaei, A.A.E. Saleh, S. Ermon, J. Dionne, Rapid identification of pathogenic bacteria using Raman spectroscopy and deep learning, Nat. Commun. 10 (2019) 4927, doi:10.1038/s41467-019-12898-9.
[30] H. Yan, M. Yu, J. Xia, L. Zhu, T. Zhang, Z. Zhu, G. Sun, Diverse region-based CNN for tongue squamous cell carcinoma classification with raman spectroscopy, IEEE Access 8 (2020) 127313–127328, doi:10.1109/ACCESS.2020.3006567.
[31] P. Wang, L. Guo, Y. Tian, J. Chen, S. Huang, Ce Wang, P. Bai, D. Chen, W. Zhu, H. Yang, W. Yao, J. Gao, Discrimination of blood species using Raman spec-

troscopy combined with a recurrent neural network, OSA Continuum 4 (2021) 672–687, doi:10.1364/OSAC.416351.
[32] S. Yu, X. Li, W. Lu, H. Li, Y.V. Fu, F. Liu, Analysis of Raman spectra by using deep learning methods in the identification of marine pathogens, Anal. Chem. 93 (2021) 11089–11098, doi:10.1021/acs.analchem.1c00431.
[33] M. Wu, S. Wang, S. Pan, A.C. Terentis, J. Strasswimmer, X. Zhu, Deep learning data augmentation for Raman spectroscopy cancer tissue classification, Sci. Rep. 11 (2021) 1–13, doi:10.1038/s41598-021-02687-0.
[34] K. He, X. Zhang, S. Ren, J. Sun, Proceedings of the IEEE conference on computer vision and pattern recognition, (2016) 770–778. 10.1109/CVPR.2016.90
[35] D. Falbel, J.J. Allaire, F. Chollet, Google RStudio, Yuan Tang, Wouter Van Der Bijl, Martin Studer, Sigrid Keydana, Package "Keras, Version 2.3.0.0 (2020).
[36] B.M. Greenwell, B.C. Boehmke, Variable Importance Plots – An Introduction to the vip Package, R J. 12 (1) (2020) 343–366, doi:10.32614/RJ-2020-013.
[37] Y. Date, J. Kikuchi, Application of a deep neural network to metabolomics studies and its performance in determining important variables, Anal. Chem. 90 (3) (2018) 1805–1810, doi:10.1021/acs.analchem.7b03795.
[38] M.A.R. Lagerquist, D.J. Gagne, G.E. Jergensen, K.L. Elmore, C.R. Homeyer, T. Smith, Making the black box more transparent: understanding the physical implications of machine learning, Bull. Amer. Meteor. Soc 100 (2019) 2175–2199.
[39] J.-.B. Yang, K.-.Q. Shen, C.-.J. Ong, X.-.P. Li, Feature selection for MLP neural network: the use of random permutation of probabilistic outputs, IEEE Trans. Neural Networks 20 (12) (2009) 1911–1922, doi:10.1109/TNN.2009.2032543.
[40] J. Zhao, H. Zeng, S. Kalia, H. Lui, Wavenumber selection based analysis in Raman spectroscopy improves skin cancer diagnostic specificity, Analyst 141 (2016) 1034–1043, doi:10.1039/C5AN02073E.
[41] Y.A. Khristoforova, I.A. Bratchenko, D.N. Artemyev, O.O. Myakinin, S.V. Kozlov, A.A. Moryatov, V.P. Zakharov, Method of autofluorescence diagnostics of skin neoplasms in the near infrared region, J. Biomed. Photon. Eng. 1 (3) (2015) 186–192, doi:10.18287/jbpe-2015-1-3-186.
[42] J. Zhao, H. Zeng, S. Kalia, H. Lui, Incorporating patient demographics into Raman spectroscopy algorithm improves in vivo skin cancer diagnostic specificity, Transl. Biophoton. 1 (1–2) (2019) e201900016, doi:10.1002/tbio.201900016.
[43] H.D. Heibel, L. Hooey, C.J. Cockerell, A Review of Noninvasive Techniques for Skin Cancer Detection in Dermatology, Am. J. Clin. Dermatol. 21 (2020) 513–524, doi:10.1007/s40257-020-00517-z.
[44] P. Gerami, J.P. Alsobrook, T.J. Palmer, H.S. Robin, Development of a novel noninvasive adhesive patch test for the evaluation of pigmented lesions of the skin, J. Am. Acad. Dermatol. 71 (2) (2014) 237–244, doi:10.1016/j.jaad.2014.04.042.
[45] L.K. Ferris, B. Jansen, J. Ho, K.J. Busam, K. Gross, D.D. Hansen, J.P. Alsobrook II, Z. Yao, G.L. Peck, P. Gerami, Utility of a noninvasive 2-gene molecular assay for cutaneous melanoma and effect on the decision to biopsy, JAMA Dermatol. 153 (2017) 75–80, doi:10.1001/jamadermatol.2017.0473.
[46] P. Gerami, Z. Yao, D. Polsky, B. Jansen, K. Busam, J. Ho, M. Matini, L.K. Ferris, Development and validation of a noninvasive 2-gene molecular assay for cutaneous melanoma, J. Am. Acad. Dermatol. 76 (1) (2017) 114–120 e2, doi:10.1016/j.jaad.2016.07.038.
[47] G. Monheit, A.B. Cognetta, L. Ferris, H. Rabinovitz, K. Gross, M. Martini, J.M. Grichnik, M. Mihm, V.G. Prieto, P. Googe, R. King, A. Toledano, N. Kabelev, M. Wojton, D. Gutkowicz-Krusin, The performance of MelaFind: a prospective multicenter study, Arch. Dermatol. 147 (2011) 188–194, doi:10.1001/archdermatol.2010.302.
[48] L.F. di Ruffano, Y. Takwoingi, J. Dinnes, N. Chuchu, S.E. Bayliss, C. Davenport, R.N. Matin, K. Godfrey, C. O'Sullivan, A. Gulati, S.A. Chan, A. Durack, S. O'Connell, M.D. Gardiner, J. Bamber, J.J. Deeks, H.C. Williams, Computer-assisted diagnosis techniques (dermoscopy and spectroscopy-based) for diagnosing skin cancer in adults, Cochrane Database Systematic Rev. (2018) 12, doi:10.1002/14651858.CD013186.
[49] Y. Zhang, A.J. Moy, X. Feng, H. Nguyen, K.R. Sebastian, J.S. Reichenberg, C.O. Wilke, M.K. Markey, J.W. Tunnell, Assessment of Raman spectroscopy for reducing unnecessary biopsies for melanoma screening, Molecules 25 (12) (2020) 2852.
[50] J. Dinnes, J.J. Deeks, N. Chuchu, L.F. di Ruffano, R.N. Matin, D.R. Thomson, K.Y. Wong, R.B. Aldridge, R. Abbott, M. Fawzy, S.E. Bayliss, M.J. Grainge, Y. Takwoingi, C. Davenport, K. Godfrey, F.M. Walter, H.C. Williams, Dermoscopy, with and without visual inspection, for diagnosing melanoma in adults, Cochrane Database Systematic Rev. (2018) 12, doi:10.1002/14651858.CD011902.pub2.
[51] H.A. Haenssle, C. Fink, F. Toberer, J. Winkler, W. Stolz, T. Deinlein, R. Hofmann-Wellenhof, A. Lallas, S. Emmert, T. Buhl, M. Zutt, A. Blum, M.S. Abassi, L. Thomas, I. Tromme, P. Tschandl, A. Enk, A. Rosenberger, Man against machine reloaded: performance of a market-approved convolutional neural network in classifying a broad spectrum of skin lesions in comparison with 96 dermatologists working under less artificial conditions, Ann. Oncol. 31 (1) (2020) 137–143, doi:10.1016/j.annonc.2019.10.013.
[52] A. Esteva, B. Kuprel, R.A. Novoa, J. Ko, S.M. Swetter, H.M. Blau, S. Thrun, Dermatologist-level classification of skin cancer with deep neural networks, Nature 542 (2017) 115–118, doi:10.1038/nature21056.
[53] J.J. Deeks, M.J. Grainge, N. Chuchu, L.F. di Ruffano, R.N. Matin, D.R. Thompson, K.Y. Wong, R.B. Aldridge, R. Abbott, M. Fawzy, S.E. Bayliss, Y. Takwoingi, C. Davenport, K. Godfrey, F.M. Walter, H.C. Williams, Visual inspection for diagnosing cutaneous melanoma in adults, Cochrane Database Systematic Rev. (2018) 12, doi:10.1002/14651858.CD013194.
[54] M. Goyal, T. Knackstedt, S. Yan, S. Hasanpour, Artificial intelligence-based image classification methods for diagnosis of skin cancer: challenges and opportunities, Comput. Biol. Med. 127 (2020) 104065, doi:10.1016/j.compbiomed.2020.104065.
[55] N. Zhang, Y.-.X. Cai, Y.-.Y. Tian, X.-.L. Wang, B. Badami, Skin cancer diagnosis based on optimized convolutional neural network, Artif Intell Med 102 (2020) 101756, doi:10.1016/j.artmed.2019.101756.
[56] S. Sigurdsson, P.A. Philipsen, L.K. Hansen, J. Larsenm, M. Gniadecka, H.C. Wulf, Detection of skin cancer by classification of Raman spectra, IEEE Trans. Biomed. Eng. 51 (10) (2004) 1784–1793, doi:10.1109/TBME.2004.831538.
[57] J. Räsänen, M. Salmivuori, I. Pölönen, M. Grönroos, N. Neittaanmäki, Hyperspectral Imaging Reveals Spectral Differences and Can Distinguish Malignant Melanoma from Pigmented Basal Cell Carcinomas, A Pilot Study, Acta Derm. Venereol. 101 (2021) 00405, doi:10.2340/00015555-3755.
[58] D. Hoar, P.Q. Lee, A. Guida, S. Patterson, C.V. Bowen, J. Merrimen, C. Wang, R. Rendon, S.D. Beyea, S.E. Clarke, Combined transfer learning and test-time augmentation improves convolutional neural network-based semantic segmentation of prostate cancer from multi-parametric MR images, Comput. Methods Programs Biomed. 210 (2021) 106375, doi:10.1016/j.cmpb.2021.106375.
[59] I. Bonavita, X. Rafael-Palou, M. Ceresa, G. Piella, V. Ribas, M.A.G. Ballester, Integration of convolutional neural networks for pulmonary nodule malignancy assessment in a lung cancer classification pipeline, Comput. Methods Programs Biomed. 185 (2020) 105172, doi:10.1016/j.cmpb.2019.105172.
[60] M. Mostavi, Y.-.C. Chiu, Y. Huang, Y. Chen, Convolutional neural network models for cancer type prediction based on gene expression, BMC Med. Genomics 13 (2020) 44, doi:10.1186/s12920-020-0677-2.
[61] L.-.G. Guillermo, J.M. Jerez, L. Franco, F.J. Veredas, Transfer learning with convolutional neural networks for cancer survival prediction using gene-expression data, PLoS One 15 (3) (2020) e0230536, doi:10.1371/journal.pone.0230536.
[62] S. Ding, Z. Wu, Y. Zheng, Z. Liu, X. Yang, X. Yang, G. Yuan, J. Xie, Deep attention branch networks for skin lesion classification, Comput. Methods Programs Biomed. 212 (2021) 106447, doi:10.1016/j.cmpb.2021.106447.
[63] M. Goyal, T. Knackstedt, S. Yan, S. Hassanpour, Artificial intelligence-based image classification methods for diagnosis of skin cancer: challenges and opportunities, Comput. Biol. Med. 127 (2020) 104065, doi:10.1016/j.compbiomed.2020.104065.
[64] A. Pacheco, R.A. Krohling, The impact of patient clinical information on automated skin cancer detection, Comput. Biol. Med. 116 (2020) 103545, doi:10.1016/j.compbiomed.2019.103545.
[65] N. Alfed, F. Khelifi, Bagged textural and color features for melanoma skin cancer detection in dermoscopic and standard images, Expert Syst. Appl. 90 (2017) 101–110, doi:10.1016/j.eswa.2017.08.010.
[66] R. Kothari, V. Jones, D. Mena, V.B. Reyes, Y. Shon, J.P. Smith, D. Schmolze, P.D. Cha, L. Lai, Y. Fong, M.C. Storrie-Lombardi, Raman spectroscopy and artificial intelligence to predict the Bayesian probability of breast cancer, Sci. Rep. 11 (2021) 6482, doi:10.1038/s41598-021-85758-6.
[67] R. Arnaout, L. Curran, Y. Zhao, J.C. Levine, E. Chinn, A.J. Moon-Grady, An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease, Nat. Med. 27 (2021) 882–891, doi:10.1038/s41591-021-01342-5.
[68] A.M. Fales, I.K. Ilev, T.J. Pfefer, Evaluation of standardized performance test methods for biomedical Raman spectroscopy, J. Biomed. Opt. 27 (7) (2021) 074705, doi:10.1117/1.JBO.27.7.074705.
[69] E. Guevara, J.C. Torres-Galván, M.G. Ramírez-Elías, C. Luevano-Contreras, F.J. González, Use of Raman spectroscopy to screen diabetes mellitus with machine learning tools, Biomed. Opt. Express 9 (10) (2018) 4998–5010, doi:10.1364/BOE.9.004998.
[70] F.R. Paolillo, V.S. Mattos, A.O. de Oliveira, F.E.G. Guimarães, V.S. Bagnato, J.C. de Castro Neto, Noninvasive assessments of skin glycated proteins by fluorescence and Raman techniques in diabetics and nondiabetics, J. Biophotonics, 12(1) (2019) e201800162. doi:10.1002/jbio.201800162.
[71] L.A. Bratchenko, I.A. Bratchenko, Y.A. Khristoforova, D.N. Artemyev, D.Y. Konovalova, P.A. Lebedev, V.P. Zakharov, Raman spectroscopy of human skin for kidney failure detection, J. Biophotonics 14 (2) (2021) e202000360, doi:10.1002/jbio.202000360.