Analyzing the serum of hemodialysis patients by means of the combination of SERS and machine learning

Lyudmila A. Bratchenko[1], Sahar Z. Al-Sammarraie[1], Elena N. Tupicova[1], Daria Y. Konovalova[2], Peter A. Lebedev[2], Valery P. Zakharov[1], Ivan A. Bratchenko[1]*

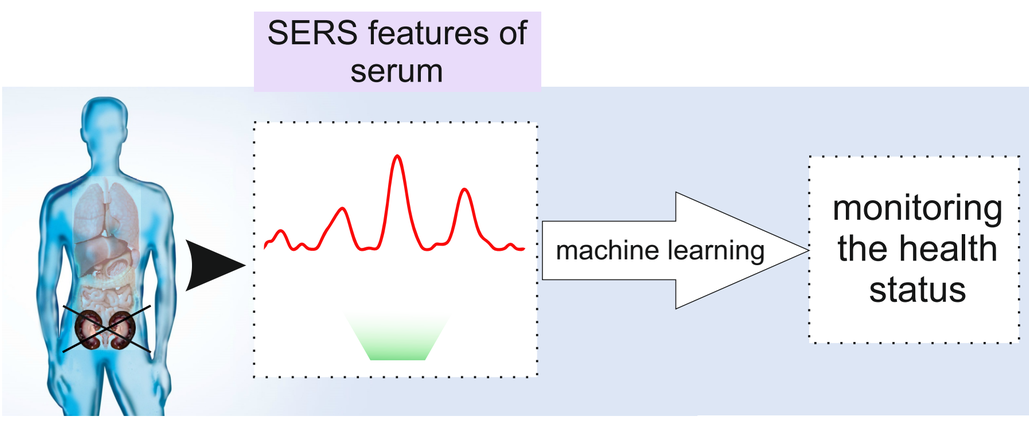
1 Samara University, Samara, Russia

2 Samara State Medical University, Samara, Russia

[*]e-mail:shamina94@inbox.ru

The study was supported by a grant from the Russian Science Foundation No. 21-75-10097, https://rscf.ru/project/21-75-10097/

1. Introduction



The aim of this work is the analysis of the spectral features of the serum in hemodialysis patients with end-stage chronic kidney disease (CKD) using a combination of surface-enhanced Raman spectroscopy (SERS) and machine learning methods.

2. Materials and methods

- Silver structures based on dried silver colloid are utilized to achieve surface enhancement of Raman scattering in the near infrared range. A silver colloid was obtained by reduction from an aqueous solution of silver nitrate with sodium citrate at the temperature of 95 °C for 20 minutes.

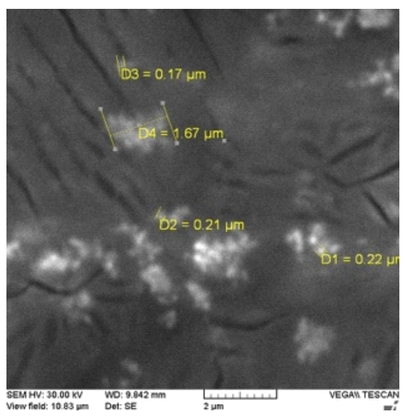


Fig. 1 Silver substrate for SERS analysis

In this study, the in vitro analysis of human serum was performed for 136 subjects, including 78 subjects of 1-3a stages of CKD depending on the glomerular filtration rate and 58 hemodialysis patients with end-stage CKD.

Multivariate analysis based on projection on the latent structures (PLS) and convolution neural network (CNN) of the processed spectral characteristics of serum was performed to accomplish the following tasks:

1. "hemodialysis end-stage CKD" group vs "1-3a stages of CKD" group: discriminating the hemodialysis patients with end-stage of CKD and the patients with stages 1-3a of CKD;
2. regression between the serum spectral characteristics and the levels of creatinine and urea for the whole dataset.

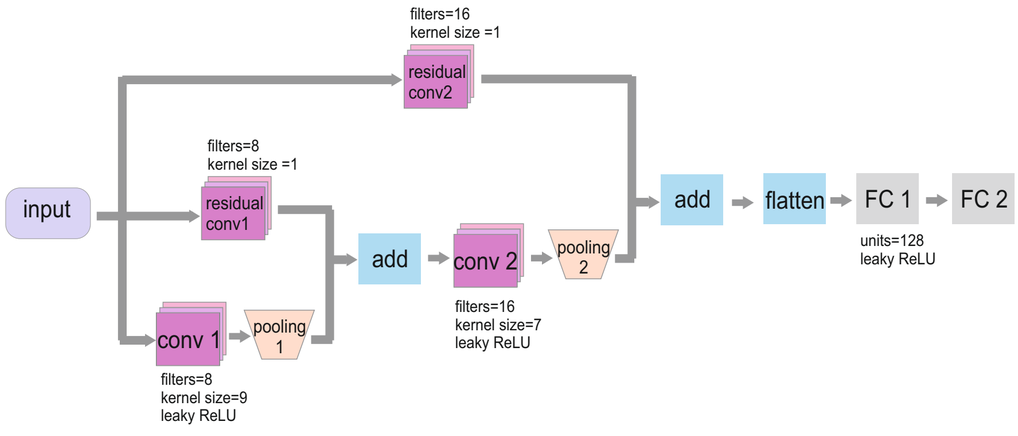


*Fig. 2 Schematic diagram of the architecture of one-dimensional CNN for the recognition of Raman spectra and serum classification, where: **input** - the input layer ; **conv** - the one-dimensional convolutional layer; **pooling** - the max pooling layer; **residual conv** - the one-dimensional convolutional layer; **add** - the layer that adds residual tensor and convolutional base; **flatten** - a flatten layer; **FC** - a fully connected layer*

3. Spectral characteristics of human serum

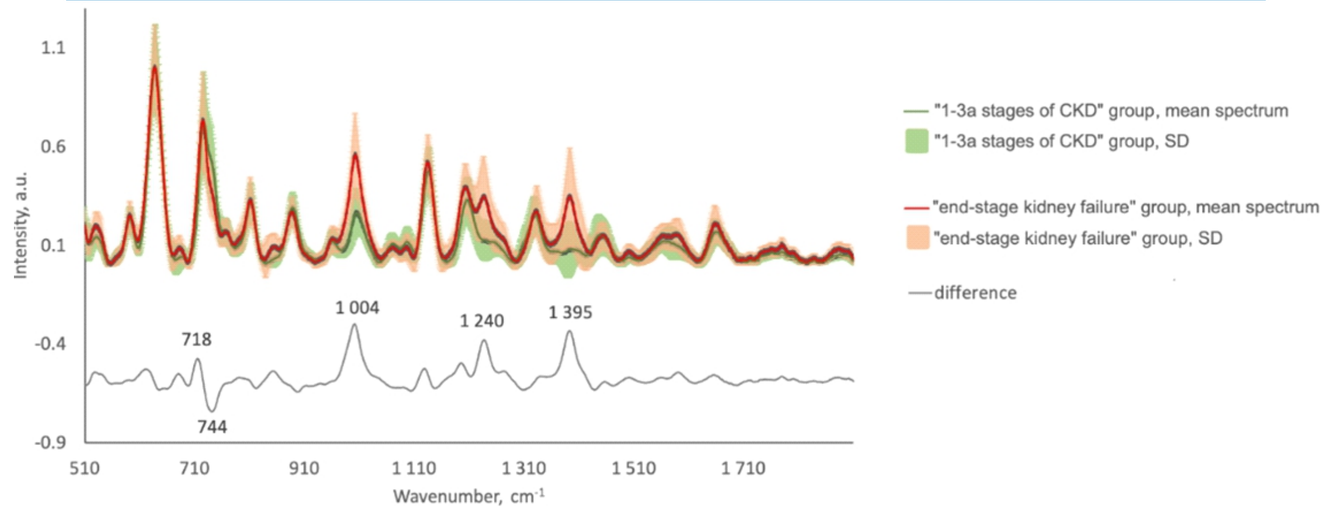


Fig. 3 Mean SERS spectra of serum for the group of patients with 1-3a stages of CKD and the group of hemodialysis patients.

For classification of subjects by the presence/absence of end-stage kidney failure, the PLS solution is slightly inferior to the solution based on CNN. For detecting the target subjects using a deep learning solution, the specificity at 0.95, the sensitivity at 0.92, and the accuracy at 0.94 are sufficient for clinical use. When constructing the models, the importance of predictors in accomplishing the classification task was assessed by means of the distribution of variable importance (VIP) in the constructed model.

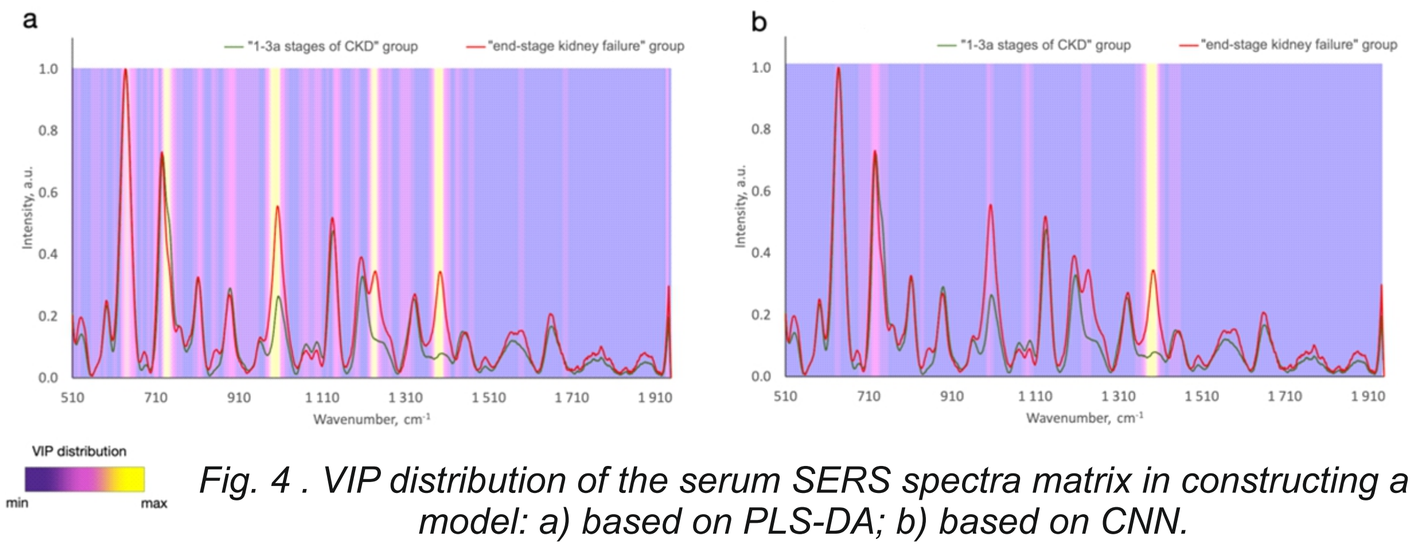


Fig. 4 . VIP distribution of the serum SERS spectra matrix in constructing a model: a) based on PLS-DA; b) based on CNN.

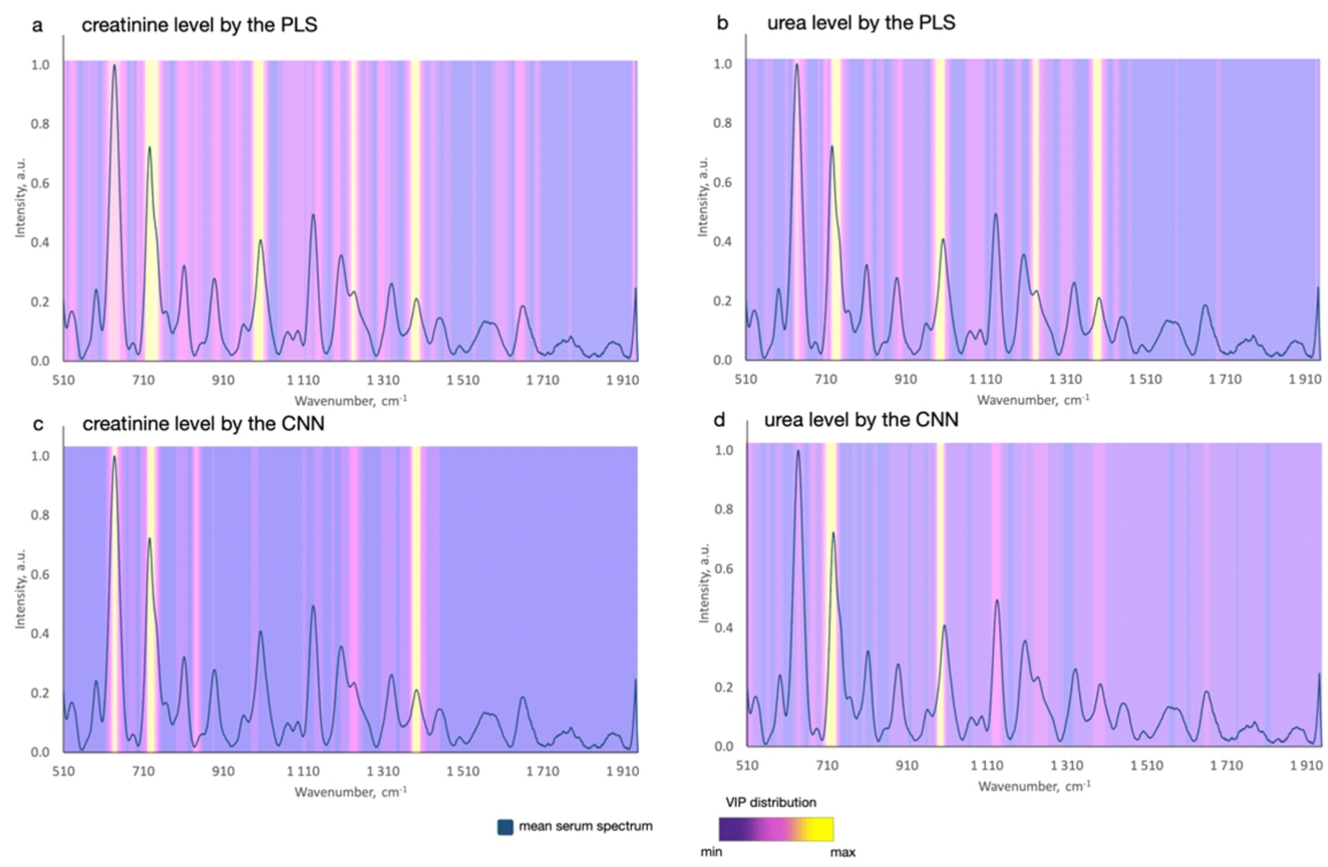


Fig. 5 VIP-distribution of the matrix of serum SERS spectra: a) regression of the creatinine level by the PLS method (R^2 = 0.74); b) regression of the urea level by the PLS method (R^2 = 0.76); c) regression of the creatinine level by the CNN method (R^2 = 0.81); and d) regression of the urea level by the CNN method (R^2 = 0.82).

Conclusion

The spectral bands for identifying dialysis patients by variable importance distribution were determined: the 641 cm^{-1}, 724 cm^{-1}, 1094 cm^{-1} and 1393 cm^{-1} bands are associated with the degree of kidney function inhibition; and the 1001 cm^{-1} band is able to demonstrate the distinctive features of hemodialysis patients with end-stage CKD. In general, the reported approach may form the basis for monitoring the health status of dialysis patients and find application in studying other pathological conditions of the human body.