Transfer Learning for Neural Network Solution of an Inverse Problem in Optical Spectroscopy with Integration of Spectroscopic Methods

Artem A. Guskov, 1,2, Igor V. Isaev, 2,3, Sergey A. Burikov, 1, 2, Tatiana A. Dolenko, 1,2, Kirill A. Laptinskiy, 2, Sergey A. Dolenko, 2

1 Faculty of Physics, Moscow State University, Moscow, Russia
2 D.V. Skobeltsyn Institute of Nuclear Physics, M.V. Lomonosov Moscow State University, Moscow, Russia
3 Kotelnikov Institute of Radioengineering and Electronics, Russian Academy of Sciences, Moscow, Russia

In this study, we consider determination of the concentrations of heavy metal ions in water by Raman scattering, IR absorption and optical absorption spectra. Getting concentration values from spectra is an inverse problem (IP), one of the ways to solve which is using machine learning (ML) methods, e.g., neural networks (NN).

Obtaining spectra of any type is a laborious and expensive process, which makes it difficult to obtain a training sample of the size sufficient for use of ML methods. On the other hand, the experimental spectra are sensitive to impurities contained in water, including the dissolved organic matter (DOM) contained in natural waters. Their presence negatively affects the quality of the solution of the IP. Moreover, DOM is specific for each water source. Therefore, it is not possible to create a universal training dataset capable of providing a stable solution on any samples.

To improve the quality of the NN solution, here we use the transfer learning approach. NN were pre-trained on a basic dataset (about 3700 patterns) containing the spectra of solutions prepared with distilled water. Then fine tuning and testing of the NN were carried out on specific datasets (200-400 patterns) containing the spectra of solutions prepared in water taken from different rivers.

We also combine transfer learning with integration of the three types of spectroscopy – simultaneous use of their spectra to solve the studied IP.

This study was performed at the expense of the grant of the Russian Science Foundation no. 19-11-00333, https://rscf.ru/en/project/19-11-00333/.