

Enhancing Antibiotics Detection in Raman Spectra with Deep Generative Models

Diogo Cachetas¹, Ensieh Iranmehr¹, Ana Vieira¹, Miguel Rocha^{2,3}, Begoña Espiña¹, Laura Rodríguez-Lorenzo¹

¹International Iberian Nanotechnology Laboratory (INL), Av. Mestre José Veiga, Braga, Portugal

²Centre of Biological Engineering, University of Minho, Braga, Portugal

³LABBELS - Associate Laboratory, Braga Guimarães, Portugal

diogo.cachetas@inl.int

Antibiotics are recognized as contaminants of emerging concern (CECs) due to their substantial impact on various environmental matrices, with far-reaching consequences for both human and animal health [1]. In alignment with the Third Watch List recommendations, the detection of sulfonamide residues in the environment provides a critical opportunity to track the distribution of these key antibiotic contaminants in aquatic systems [2].

Surface-enhanced Raman spectroscopy (SERS) is a rapid and highly sensitive method for detecting trace levels of antibiotic residues, including sulfonamides [3]. However, as SERS increasingly relies on data-driven approaches, conventional methods for spectral data analysis face several limitations. Deep learning (DL) algorithms have demonstrated remarkable potential in extracting meaningful information from spectral data, but their performance heavily relies on large, labeled datasets for model training [4]. This dependence creates a significant challenge due to the limited available high-quality spectra and data privacy concerns.

To address these challenges, our research focuses on developing a deep-learning generative model trained with available SERS spectra, to generate high-quality synthetic data (**Figure 1**). This innovative approach aims to enhance the detection of sulfonamides in aquatic environments, enhancing DL models trained with these augmented datasets.

Our findings show that the generated spectra are informative and high-quality (**Figure 2**). Preliminary results also suggest that the generated data could also significantly improve the classification accuracy of sulfonamides with SERS spectra in comparison to traditional classifiers trained solely on original spectra data, such as Convolutional Neural Networks (CNNs) and Multilayer Perceptrons (MLPs). These results highlight the potential of generative models to overcome data limitations and advance the application of SERS in environmental monitoring.

References

[1] Y. Ben *et al.*, “Efficient detection and assessment of human exposure to trace antibiotic residues in drinking water,” *Water Res.*, vol. 175, p. 115699, May 2020, doi: 10.1016/j.watres.2020.115699.

[2] European Commission, “Commission Implementing Decision (EU) 2020/1161 of 4 August 2020 establishing a watch list of substances for Union-wide monitoring in the field of water policy pursuant to Directive 2008/105/EC of the European Parliament and of the Council,” 2020.

[3] M. Luo *et al.*, “Selective surface enhanced Raman detection and effective photocatalytic degradation of sulfonamides antibiotic based on a flexible three-dimensional chitosan/carbon nitride/silver substrate,” *J. Hazard. Mater.*, vol. 459, p. 132131, Oct. 2023, doi: 10.1016/j.jhazmat.2023.132131.

[4] Y. Qi *et al.*, “Recent Progresses in Machine Learning Assisted Raman Spectroscopy,” *Adv. Opt. Mater.*, vol. 11, no. 14, Jul. 2023, doi: 10.1002/adom.202203104.

Figures

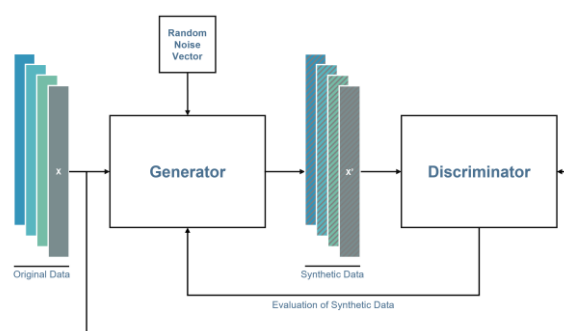


Figure 1. Representation of the Deep Generative architecture.

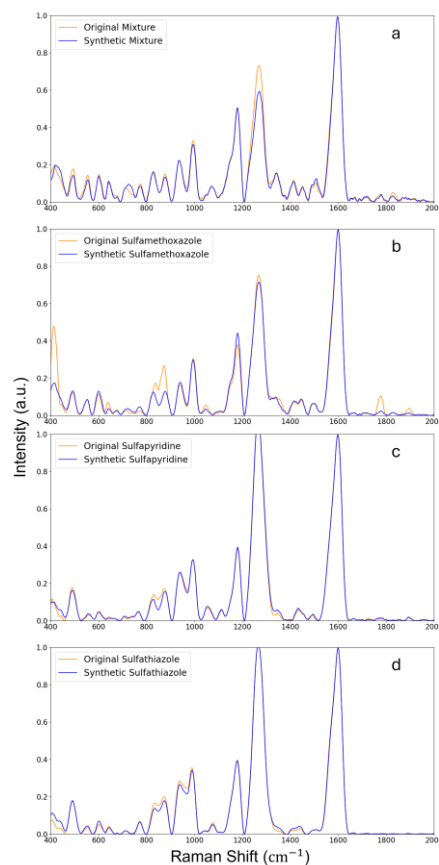


Figure 2. Average spectra for original data (orange) and synthetic data (blue): **a** mixture; **b** sulfamethoxazole; **c** sulfapyridine; **d** sulfathiazole.