LIBPhys | Projecto de Tese de Doutoramento | Engenharia Física

Improvement of Raman spectra analysis using Machine Learning algorithms

Objectives:

Raman spectroscopy is a non-destructive spectroscopic technique typically used to determine vibrational modes of molecules. It has proven to be a valuable tool for the evaluation of both the organic and the inorganic composition of materials and it is used in a wide range of applications: biomedicine, materials analysis, cultural heritage, environmental assessment. However, the accuracy of the data interpretation is dependent of the reliability of the spectra evaluation and of the significance of the data acquired in the measurements. In this thesis we propose the implementation of a novel spectra evaluation approach, using Machine Learning algorithms based on Deep Learning´s Convolutional Neural Networks. This tool will then be implemented in a confocal Raman microscope situated at LIBPhys and used profusely in biomedical and cultural heritage applications.

The main motivation for the use of this computational resource is the efficient spectra evaluation for outcome prediction with a limited database.

This network will be trained with Raman spectra obtained from different materials and samples:

- Human tissues: Bone and teeth and organ tissues- these samples are composed of both organic and inorganic compounds and the quick assessment of their presence/abundancy could be paramount in the diagnosis of diseases.
- Pigments: the analysis of pigments is a very useful tool in dating artworks and conservation assessment and Raman is also a powerful tool for the analysis of both organic and inorganic pigments.
- Corrosion products: corrosion is a constraint in the preservation of metal artifacts over time. Although metal analysis is not efficient using Raman

microscopy, the evaluation of the corrosion products (oxides, sulfides) might give important information on the assessment of conservation strategies.

Framework:

Raman spectroscopy has demonstrated, repeatedly in several research studies, to be a very suitable technique to characterize both organic and inorganic superficial components, since it is possible to acquire accurate and precise spectroscopic information on present minerals through the observation of the characteristic energies of their vibrational modes. This method is becoming progressively important especially for its high biochemical specificity, low water sensitivity, simplicity, capability to work in the near-infrared region and remote analysis potential by means of fiber-optics.

Taks and timeline:

The proposal is divided in 4 main tasks that will occur in concurrently:

Task 1 (12 months) – Raman spectra evaluation using function-based fitting – Samples gathered and selected for the proposal will be evaluated using previously tested methodologies based on Raman analysis in LIBPhys-NOVA facilities, using existing benchtop Raman microscope and the inbuilt software (Labspec) for background removal using linear function and peak fitting using Voigt profile approach.

Task 2 (18 months) - Spectra evaluation of XploRA spectra using Machine Learning algorithms - Data acquired in the previous task will be used as input to train the Machine Learning (ML) algorithm for Raman spectra evaluation. The input of the machine will be the spectra, with all its rich information, and the output will approach the demineralization estimation made by the standard method of depolarization ratio determination using frequentist analysis. The expectation is that the developed system will be a more robust estimator, as it will not be prone to fatigue or subjective evaluation.

In addition to the aforementioned main goal, and because the overall project aims at producing a dedicated Raman spectrometer, we will search for the minimal set of input data, entering the ML predictor, that achieves a predictive power that is deemed sufficient for deployment of the technique. That will be performed in the design of the ML architectures to be tested, by pruning such structures, as well as by understanding the networks that are found to be successful in the prediction task.

Task 3 (6 months) – Creation of a database for the acquired materials.

Task 4 (12 months) – Case study applications – the developed spectra evaluation approach will be used in real case studies ongoing at the research group.

Location:
FCT-NOVA

Candidate profile
Master Biomedical Engineering, Master Physics Engineering, Master Physics.

Orientador

- Nome: Sofia Pessanha
- Instituição: FCT-NOVA
- Email: sofia.pessanha@fct.unl.pt
- ORCID ID: 0000-0002-6039-2494

Co-Orientador

- Nome: Ricardo Vigário
- Instituição: FCT-NOVA
- Email: r.vigario@fct.unl.pt
- ORCID ID: 0000-0003-0950-6035