

UNIVERSIDADE DE SÃO PAULO  
INSTITUTO DE FÍSICA DE SÃO CARLOS

ALFREDO ANTONIO ALENCAR EXPOSITO DE QUEIROZ

Espectroscopia Raman e redes neurais artificiais na identificação  
de minerais dos grupos do pirocloro e da microlita

São Carlos  
2019



ALFREDO ANTONIO ALENCAR EXPOSITO DE QUEIROZ

Espectroscopia Raman e redes neurais artificiais na identificação  
de minerais dos grupos do pirocloro e da microlita

Dissertação apresentada ao Programa de Pós-  
Graduação em Física do Instituto de Física de  
São Carlos da Universidade de São Paulo, para  
obtenção do título de Mestre em Ciências.

Área de concentração: Física Aplicada  
Opção: Física Computacional  
Orientador: Prof. Dr. Marcelo Barbosa de  
Andrade

Versão corrigida  
(Versão original disponível na Unidade que aloja o Programa)

São Carlos  
2019

AUTORIZO A REPRODUÇÃO E DIVULGAÇÃO TOTAL OU PARCIAL DESTA  
TRABALHO, POR QUALQUER MEIO CONVENCIONAL OU ELETRÔNICO PARA  
FINS DE ESTUDO E PESQUISA, DESDE QUE CITADA A FONTE.

Queiroz, Alfredo Antonio Alencar Exposito de  
Espectroscopia Raman e redes neurais artificiais na  
identificação de minerais dos grupos do pirocloro e da  
microlita / Alfredo Antonio Alencar Exposito de Queiroz;  
orientador Marcelo Barbosa de Andrade - versão corrigida -  
- São Carlos, 2019.  
101 p.

Dissertação (Mestrado - Programa de Pós-Graduação em  
Física Aplicada Computacional) -- Instituto de Física de  
São Carlos, Universidade de São Paulo, 2019.

1. Redes neurais artificiais. 2. Aprendizado de  
máquina. 3. Espectroscopia Raman. 4. Identificação. 5.  
Supergrupo do pirocloro . I. Andrade, Marcelo Barbosa de,  
orient. II. Título.

Aos meus pais.



## **AGRADECIMENTOS**

Aos meus pais (Alvaro e Mari) por todo o incentivo, carinho e apoio antes e durante este caminho.

Ao Prof. Dr. Marcelo Barbosa, pela atenção, motivação, orientação e paciência durante o processo.

Ao Prof. Dr. Javier Ellena agradeço pela oportunidade, pela infraestrutura do laboratório e pelas discussões envolvendo partes do trabalho.

Agradeço ao Museu de Geociências da USP e a Andrea Ferreira do Museu das Minas e do Metal – MM Gerdau por disponibilizar algumas das amostras utilizadas neste trabalho.

Aos meus membros do grupo e amigos: ao Prof. Dr. Carlos Simone, José Augusto, Matheus, Juan, Gerson, Natália, Leonardo, Adriele, Cairo, Adilson, Leandro, Luan, Felipe, Raphael, Nicolas, Natali, José, Marcos e Noel pelos bons momentos, conversas e comemorações.

Ao Instituto de Física de São Carlos – USP, por proporcionar a realização do curso de mestrado.

Agradeço a Secretaria da pós-graduação e a Biblioteca por todas as orientações durante o mestrado.

Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) – processo 131650/2017-3, pela concessão de bolsa de mestrado.

A todos que contribuíram para este trabalho.





“Saber muito não lhe torna inteligente. A inteligência se traduz na forma que você recolhe, julga, maneja e, sobretudo, onde e como aplica a informação.”

**Carl Sagan**



## RESUMO

QUEIROZ, A. A. A. E. **Espectroscopia Raman e redes neurais artificiais na identificação de minerais dos grupos do pirocloro e da microlita**. 2019. 101 p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019.

A atividade de mineração está diretamente relacionada com o desenvolvimento socioeconômico de um país, exercendo uma grande influencia nas exportações de matéria-prima, produção de energia e infraestrutura. O Brasil possui as maiores jazidas de minerais do supergrupo do pirocloro e isso gera bilhões de dólares na exportação de nióbio e produtos associados para o país. Esta situação ilustra a importância de desenvolvimento de técnicas de identificação e caracterização dos minerais com o objetivo de aperfeiçoar os processos de beneficiamento, otimizar o aproveitamento de recursos naturais e reduzir os impactos ambientais do procedimento de extração. A espectroscopia Raman é uma técnica adequada para a caracterização de minerais, uma vez que não é necessário o preparo prévio da amostra para a análise e não é destrutiva. Neste trabalho é realizada a identificação de minerais dos grupos da microlita e pirocloro pertencentes ao supergrupo do pirocloro, utilizando redes neurais artificiais para classificar os espectros Raman característicos coletados em três regiões de análise: 100–1400  $\text{cm}^{-1}$  (região de bandas Raman características), 1200–1800  $\text{cm}^{-1}$  (região das bandas de deformação angular de  $\text{H}_2\text{O}$ ) e 2800–4000  $\text{cm}^{-1}$  (região das bandas de estiramento de OH). Os espectros coletados foram tratados por dois métodos de correção de linha de base (Mínimos Quadrados Assimétricos e Vizinhos Próximos) antes de serem utilizados nas primeiras etapas de aprendizagem das redes neurais. Outras técnicas de caracterização foram utilizadas (difração de raios-X por monocristal e microsonda eletrônica) para confirmar a identificação das amostras padrão deste trabalho que foram usadas como conjunto de dados de treinamento das redes neurais. As redes estruturadas neste trabalho apresentaram bom desempenho, obtendo acurácia de 0,88 com valor da função custo de 0,52, em comparação com os valores da literatura.

Palavras-Chaves: Redes neurais artificiais. Aprendizado de máquina. Espectroscopia Raman. Identificação. Supergrupo do pirocloro.



## Abstract

QUEIROZ, A. A. A. E. **Raman spectroscopy and artificial neural networks in the identification of minerals of pyrochlore and microlite groups**. 2019. 101 p. Dissertação (Mestrado em Ciências) – Instituto de Física de São Carlos, Universidade de São Paulo, São Carlos, 2019.

The mining activity is directly related to the socioeconomic development of a country, since it highly influences commodity exports, energy production and infrastructure, for example. Brazil has largest deposits of pyrochlore supergroup minerals worldwide, which generates billions of dollars in exports of niobium and associated products for the country. This scenario illustrates the importance of techniques for the identification and characterization of minerals towards improving beneficiation processes, optimizing the use of natural resources and reducing environmental impacts of the extraction procedure. Raman spectroscopy is an excellent technique for the characterization of minerals, since it requires no pre-preparation of the samples for analyses and is not destructive. This dissertation addresses the identification of minerals of microlite and pyrochlore groups from the pyrochlore supergroup by artificial neural networks for the classification of characteristic Raman spectra collected in three regions of analysis, namely: 100-1400  $\text{cm}^{-1}$  (characteristic Raman bands), 1200-1800  $\text{cm}^{-1}$  (region of  $\text{H}_2\text{O}$  bending) and 2800-4000  $\text{cm}^{-1}$  (OH stretching bands). The spectra were treated by two methods of baseline correction (Asymmetric Least Squares and Near Neighbors) prior to use in the first stages of neural network learning. Single-crystal X-ray and electron microprobe were also applied towards confirming the identification of standard samples used as a training dataset in networks. The networks structured showed good performance and achieved 0.88 accuracy at a 0.52 cost function in comparison to other networks reported in the literature.

Keywords: Artificial neural networks. Machine learnig. Raman spectroscopy. Identification. Pyrochlore supergroup.



## Lista de Figuras

Figura 1 - Principais países com atividade de mineração.....	20
Figura 2 - Estruturas de espécie mineral do grupo da microlita.....	21
Figura 3 - Da esquerda para a direita, os físicos Adolf Gustav Smekal, Chandrasekhara Venkata Raman, Kariamanickam Srinivasa Krishnan, Grigory Samuilovich Landsberg e Leonid Isaakovich Mandelstam. ....	25
Figura 4 – Diagrama do espalhamento elástico (Rayleigh) e inelástico (Raman). ....	26
Figura 5 - Esquema do espalhamento Raman. ....	28
Figura 6 - Resumo histórico das redes neurais artificiais até a década de 1990.....	30
Figura 7 - Ilustração da estrutura básica de um neurônio biológico (a) e sua visualização ao microscópio eletrônico de varredura (b). ....	31
Figura 8 - Modelo de um neurônio artificial. ....	31
Figura 9 - Ilustração de uma rede de múltiplas camadas.....	32
Figura 10 - Diagrama dos fluxos de dados presentes em uma rede PMC, as linhas contínuas representam o <i>feedforward</i> e as tracejadas o <i>backpropagation</i> . ....	33
Figura 11 – (a) Espectrômetro Raman da Horiba, modelo LabRam HR UV-Visível-NIR (200-2100 nm) do Centro de Caracterização de Espécies Minerais (CCEM) do Instituto de Física de São Carlos (IFSC). É mostrado em (b): Entrada dos lasers de 532 nm (EL <sub>2</sub> ), 785 nm (EL <sub>2</sub> ) e 633 nm (EL <sub>1</sub> ). F = filtro, E = espelho, L = lente, LA = Laser incidente na amostra, LE = Laser espalhado pela amostra e G = indica a direção do feixe para a grade de difração que encaminhará o sinal para o detector CCD. Em (c) é ilustrado o caminho ótico do feixe até o detector. ....	39
Figura 12 - Espectro Raman bruto na região de 50 a 1400 cm <sup>-1</sup> da amostra MI_AN4 excitadas pelos lasers 532 (verde), 633 (vermelho) e 785 (preto) nm. ....	45
Figura 13 - Espectros de Intensidade (U.A.) em função do Deslocamento Raman (cm <sup>-1</sup> ) sem correção (em vermelho) e com as linhas de base calculadas (Implementação de ALS e VP em verde e preto, respectivamente). ....	47
Figura 14 - Espectros Raman da espécie PY1_AN1, na região de 100-200 cm <sup>-1</sup> , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	48
Figura 15 - Espectros Raman da espécie PY_AN02, na região de 100-200 cm <sup>-1</sup> , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	49
Figura 16 - Espectros Raman da espécie PY_AN3, na região de 100-200 cm <sup>-1</sup> , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	49
Figura 17 - Espectros Raman da espécie MI_AN4, na região de 100-200 cm <sup>-1</sup> , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	50

Figura 18 - Espectros Raman da espécie FLUMICRO_AN05, na região de 100-200 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	51
Figura 19 - Espectros Raman da espécie MICR_AN06, na região de 100-200 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	51
Figura 20 - Espectros Raman da espécie PY1_AN1, na região de 200-500 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	52
Figura 21 - Espectros Raman da espécie PY_AN02, na região de 200-500 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	52
Figura 22 - Espectros Raman da espécie PY_AN3, na região de 200-500 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	53
Figura 23 - Espectros Raman da espécie MI_AN4, na região de 200-500 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	53
Figura 24 - Espectros Raman da espécie FLUMICRO_AN05, na região de 200-500 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	54
Figura 25 - Espectros Raman da espécie MICR_AN06, na região de 200-500 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	54
Figura 26 - Espectros Raman da espécie PY1_AN1, na região de 500-800 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	55
Figura 27 - Espectros Raman da espécie PY_AN02, na região de 500-800 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	56
Figura 28 - Espectros Raman da espécie PY_AN3, na região de 500-800 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	56
Figura 29 - Espectros Raman da espécie MI_AN4, na região de 500-800 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	57
Figura 30 - Espectros Raman da espécie FLUMICRO_AN05, na região de 500-800 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	57
Figura 31 - Espectros Raman da espécie MICR_AN06, na região de 500-800 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	58
Figura 32 - Espectros Raman da espécie PY1_AN1, na região de 800-1400 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	59
Figura 33 - Espectros Raman da espécie PY_AN02, na região de 800-1400 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	59
Figura 34 - Espectros Raman da espécie PY_AN3, na região de 800-1400 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	60
Figura 35 - Espectros Raman da espécie MI_AN4, na região de 800-1400 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	60



Figura 36 - Espectros Raman da espécie FLUMICRO_AN05, na região de 800-1400 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. (a) Utilizando os três lasers. (b) Sem a medida com o laser 785 nm. ....	61
Figura 37 - Espectros Raman da espécie MICR_AN06, na região de 800-1400 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	62
Figura 38 - Espectros Raman da espécie MI_AN4, na região de 1400-1800 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	63
Figura 39 - Espectros Raman da espécie PY1_AN1, na região de 2800-4000 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Em (a) Esquerda por ALS e direita por VP. Em (b), espectros corrigidos por ALS utilizando os lasers 532 nm (em verde) e 633 nm (em vermelho). ....	64
Figura 40 - Espectros Raman da espécie PY_AN02, na região de 2800-4000 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	65
Figura 41 - Espectros Raman da espécie PY_AN3, na região de 2800-4000 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	65
Figura 42 - Espectros Raman da espécie MI_AN4, na região de 2800-4000 $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. ....	66
Figura 43 - Acurácia e função custo em função da variação do número de camadas ocultas para o conjunto da região de 100 – 1400 $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b). ....	68
Figura 44 - Acurácia e função custo em função da variação do número de camadas ocultas para o conjunto da região de 1200 – 1800 $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b). ....	68
Figura 45 - Acurácia e função custo em função da variação do número de camadas ocultas para o conjunto da região de 2800 – 4000 $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b). ....	69
Figura 46 - Acurácia e função custo em função da variação do número de neurônios na última camada oculta para o conjunto da região de 100 – 1400 $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b). ....	69
Figura 47 - Acurácia e função custo em função da variação do número de neurônios na última camada oculta para o conjunto da região de 1200 – 1800 $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b). ....	69
Figura 48 - Acurácia e função custo em função da variação do número de neurônios na última camada oculta para o conjunto da região de 2800 – 4000 $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b). ....	70
Figura 49 - Variação da função custo em relação à taxa de aprendizagem (TA): (a) TA = 0.01 e (b) TA = 1 ....	71
Figura 50 - Acurácia em relação ao número de épocas, indicação presença de sobreajuste ....	71
Figura 51 – Exemplos dos resultados da rede com conjunto de testes na região de 100-1400 $\text{cm}^{-1}$ , processado por ALS. Representada em azul a probabilidade da amostra ser de	

- cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada e (b) caso de erro de identificação da classe, com acerto do grupo da microlita. .... 72
- Figura 52 - Exemplos dos resultados da rede com conjunto de testes na região de 100-1400  $\text{cm}^{-1}$ , processado por VP. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo da microlita e (c) caso de erro na identificação de espécie e grupo. .... 73
- Figura 53 - Exemplos dos resultados da rede com conjunto de testes na região de 1200-1800  $\text{cm}^{-1}$ , processado por ALS. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo da microlita e (c) caso de erro na identificação de espécie e grupo. .... 73
- Figura 54 - Exemplos dos resultados da rede com conjunto de testes na região de 1200-1800  $\text{cm}^{-1}$ , processado por VP. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo do pirocloro e (c) caso de erro na identificação de espécie e grupo. .... 74
- Figura 55 - Exemplos dos resultados da rede com conjunto de testes na região de 2800-4000  $\text{cm}^{-1}$ , processado por ALS. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo do pirocloro e (c) caso de erro na identificação de espécie e grupo. .... 74
- Figura 56 - Exemplos dos resultados da rede com conjunto de testes na região de 2800-4000  $\text{cm}^{-1}$ , processado por VP. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo do pirocloro e (c) caso de erro na identificação de espécie e grupo. .... 75

## SUMÁRIO

1	INTRODUÇÃO.....	19
1.1	MOTIVAÇÕES PARA O TRABALHO.....	19
1.2	OBJETIVOS DO TRABALHO.....	22
1.2.1	OBJETIVO GERAL.....	22
1.2.2	OBJETIVOS ESPECÍFICOS.....	22
1.3	MOTIVAÇÕES PARA UTILIZAÇÃO DE UMA RNA NA IDENTIFICAÇÃO DE MINERAIS.....	23
1.4	ESTRUTURA DO TRABALHO.....	23
2	FUNDAMENTOS DE ESPECTROSCOPIA RAMAN E REDES NEURAIS.....	25
2.1	A TEORIA DO ESPALHAMENTO RAMAN.....	25
2.1.1	O MOMENTO DE DIPOLO INDUZIDO.....	26
2.1.2	DIAGRAMA DO PROCESSO DE INTERAÇÃO ENTRE FÓTON-ELÉTRON.....	27
2.2	REDES NEURAIS ARTIFICIAIS.....	28
2.2.1	BREVE HISTÓRICO.....	28
2.2.2	MODELO DE UM NEURÔNIO ARTIFICIAL.....	30
2.2.3	REDE NEURAL PERCEPTRON DE MULTIPLAS CAMADAS.....	32
3	MATERIAIS E MÉTODOS.....	35
3.1	AMOSTRAS UTILIZADAS.....	35
3.2	O ESPECTRÔMETRO RAMAN.....	37
3.3	TRATAMENTO COMPUTACIONAL DOS DADOS.....	40
3.4	MONTAGEM TOPOLOGICA DA REDE NEURAL ARTIFICIAL.....	41
4	RESULTADOS E DISCUSSÃO.....	45
4.1	CARACTERÍSTICAS DO ESPECTRO RAMAN.....	45
4.1.1	REMOÇÃO DO RUÍDO DE FUNDO.....	46
4.2	AMOSTRAS DE TREINAMENTO.....	47
4.2.1	VIBRAÇÃO DA REDE.....	48
4.2.2	DEFORMAÇÃO ANGULAR.....	51
4.2.3	ESTIRAMENTO SIMÉTRICO E ANTISSIMÉTRICO.....	55
4.2.4	SOBRETONS E COMBINAÇÕES DE BANDAS.....	58
4.2.5	DEFORMAÇÃO ANGULAR DE H <sub>2</sub> O.....	62
4.2.6	ESTIRAMENTO DAS LIGAÇÕES O-H DA MOLÉCULA DE ÁGUA.....	63
4.3	APRENDIZADO DA REDE.....	66
4.3.1	TREINAMENTO E TESTES.....	66
4.3.2	VALIDAÇÃO.....	75
5	CONCLUSÃO.....	77
5.1	CONCLUSÕES FINAIS.....	77

5.2	PERSPECTIVAS FUTURAS.....	78
	REFERÊNCIAS .....	81
	APÊNDICE A – SAÍDA DAS REDES NEURAS ARTIFICIAIS PARA O CONJUNTO DE VALIDAÇÃO.....	87
	REGIÃO DE 100-1400 $\text{cm}^{-1}$ .....	87
	REGIÃO DE 1200-1800 $\text{cm}^{-1}$ .....	95
	REGIÃO DE 2800-4000 $\text{cm}^{-1}$ .....	99

# 1 INTRODUÇÃO

## 1.1 MOTIVAÇÕES PARA O TRABALHO

Os minerais desempenham um papel tão importante na evolução humana que divide a pré-história humana em função dos tipos de minerais utilizados tais como idade da pedra, do bronze e do ferro.<sup>1</sup>

Entre os indicadores sociais de desenvolvimento de um país, a atividade de mineração desempenha um papel social importante por fortalecer a pesquisa científica e a inovação nacional, isto promove a diversificação industrial e a agregação de valor às commodities. Existe, portanto uma relação direta entre desenvolvimento socioeconômico e o consumo de bens minerais.<sup>2</sup> Muitos países como Estados Unidos, Austrália e Canadá devem seu desenvolvimento econômico e social às atividades de mineração, tendo grande impacto na produção de infraestrutura, energia e maquinário. O carvão continua sendo explorado, principalmente nos Estados Unidos, Rússia e China, e é responsável pela geração de um quarto da energia do mundo.<sup>3</sup> Os países mais importantes em termos de produção minérios são mostrados na Figura 1. Muitos desses países são atualmente economias emergentes sendo que as maiores jazidas minerais encontram-se nos países em desenvolvimento.<sup>4</sup> Por exemplo, atualmente o Brasil é o detentor das maiores reservas de nióbio, que rende bilhões de dólares por ano em exportação para o país.<sup>5</sup>

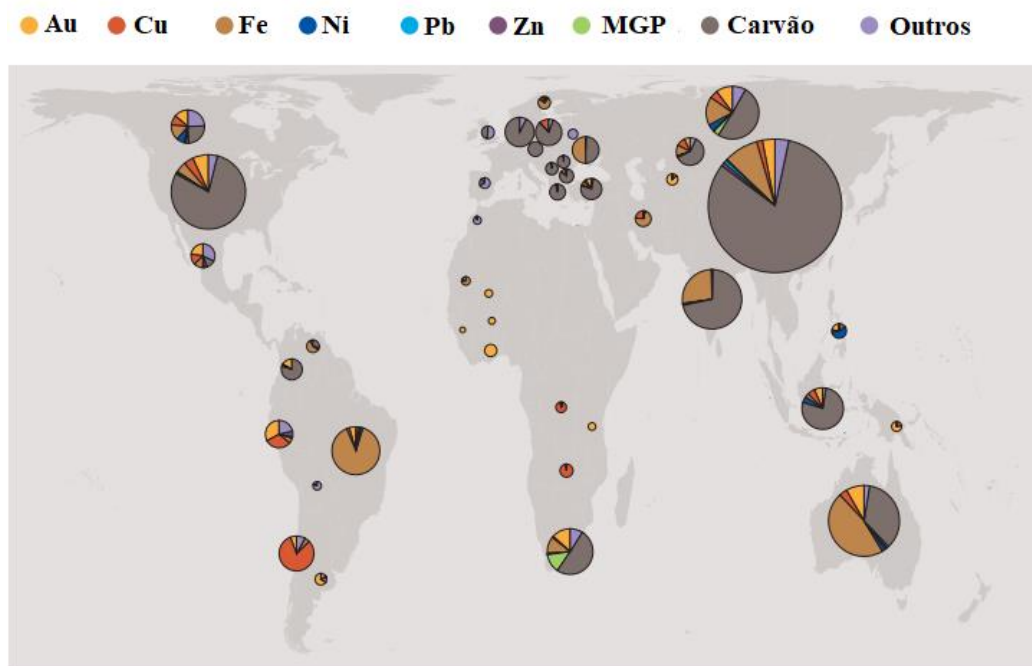


Figura 1 - Principais países com atividade de mineração. \*MGP: metais do grupo da platina, grupo formado pelos elementos platina (Pt), paládio (Pd), ródio (Rd), rutênio (Rh), irídio (Ir) e ósmio(Os).

Fonte: Adaptada de INTERNATIONAL ...<sup>4</sup>

A formação dos minerais na Terra é resultado de uma série de fenômenos físicos e químicos que ocorrem praticamente em todas as eras geológicas do planeta através dos processos magmáticos, metamórficos e sedimentários. Um mineral é definido como um sólido que ocorre naturalmente com arranjo atômico altamente ordenado e composição química homogênea (não necessariamente fixa).<sup>6</sup> Os minerais do supergrupo do pirocloro são representados pela fórmula geral  $A_2B_2X_6Y$ . O sítio cristalográfico A pode hospedar os íons  $Na^{1+}$ ,  $Ca^{2+}$ ,  $Ag^{1+}$ ,  $Mn^{2+}$ ,  $Sr^{2+}$ ,  $Ba^{2+}$ ,  $Fe^{2+}$ ,  $Pb^{2+}$ ,  $Sn^{2+}$ ,  $Sb^{3+}$ ,  $Bi^{3+}$ , Y, Ce (e outros elementos de terras raras), Sc, U, Th, □ (vacância) ou  $H_2O$ , com raio iônico de 0,9 até 1,7 Å. O B pode hospedar os íons  $Nb^{5+}$ ,  $Ta^{5+}$ ,  $Ti^{4+}$ ,  $Sb^{5+}$ ,  $W^{6+}$ ,  $Mg^{2+}$ ,  $Al^{3+}$  e em menor quantidade  $V^{5+}$ ,  $Sn^{4+}$ , Zr, Hf,  $Fe^{3+}$ , e  $Si^{4+}$ , com raio iônico entre 0,4 até 0,8 Å. No sítio X os íons  $O^{2-}$ ,  $OH^{1-}$  e  $F^{1-}$ . Por último, o sítio Y hospeda os íons  $OH^{1-}$ ,  $F^{1-}$ ,  $O^{2-}$  e pode haver vacância e cátions monovalentes, como o íon  $K^{1+}$ .<sup>7,8</sup>

Para determinar o grupo de uma espécie mineral dentro do supergrupo do pirocloro, primeiramente se analisa a valência dominante do sítio B e depois o elemento dominante desta valência.<sup>7</sup> Por exemplo, neste trabalho analisamos minerais de valência 5+ dominante no sítio B. O grupo do pirocloro e da microlita tem como elementos

dominantes  $\text{Nb}^{5+}$  e  $\text{Ta}^{5+}$  respectivamente. Como exemplo, a Figura 2 ilustra a estrutura de minerais do grupo da microlita.

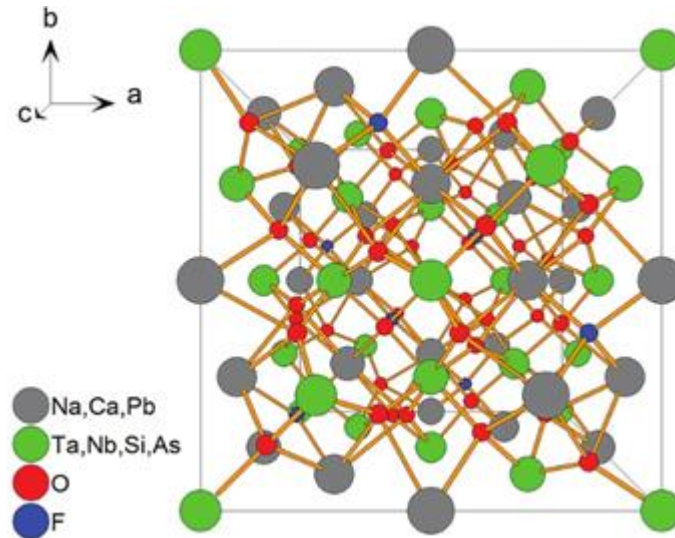


Figura 2 - Estruturas de espécie mineral do grupo da microlita.

Fonte: ANDRADE et al.<sup>9</sup>

O Nb possui grandes aplicações na indústria metalúrgica, principalmente para estabilizar os aços inoxidáveis contra a corrosão.<sup>10</sup> A partir de 1970, o Nb passou a ser utilizado em diversas aplicações avançadas, como a produção de superligas resistentes à temperaturas elevadas para a produção de turbinas.<sup>11</sup> No século XXI, o Brasil detém mais de 80% do mercado mundial, evidenciando que os principais países industrializados dependem do nióbio brasileiro para suprir suas indústrias siderúrgicas.

A economia desenvolvida ao redor dos minerais do grupo do pirocloro e da microlita ilustra bem a importância do desenvolvimento de técnicas de identificação e caracterização dos minerais visando ao aproveitamento máximo dos recursos naturais de nosso país. Ao mesmo tempo, a caracterização dos minerais permite se conhecer de antemão o valor econômico do minério em pesquisa, sua demanda no mercado nacional/internacional além de proporcionar uma significativa redução nos impactos ambientais do processo de mineração.

Tendo em vista os pontos apresentados anteriormente, esse trabalho propõe a identificação mineralógica dos minerais do supergrupo do pirocloro (grupo pirocloro e microlita) utilizando a espectroscopia Raman aliada a uma rede neural artificial, concentrando-se especialmente nas fases minerais presentes. A utilização de redes neurais

artificiais na identificação de minerais, a partir da interação com a radiação eletromagnética na região do visível ou raios X, é relativamente recente. Isto incentiva a utilizar estes recursos de inteligência artificial em conjunto com a espectroscopia Raman.<sup>12-18</sup> Os dados obtidos a partir da espectroscopia Raman e da base de dados RRUFF<sup>19</sup> foram então interpretados com o emprego de uma rede neural artificial.

## **1.2 OBJETIVOS DO TRABALHO**

### **1.2.1 OBJETIVO GERAL**

Verificar a utilização de Redes Neurais Artificiais (RNA) como técnica de reconhecimento de padrões e auxílio na identificação ou classificação de minerais do supergrupo do pirocloro (grupos pirocloro e microlita) mediante a análise dos seus espectros ópticos obtidos por aparelho de espectroscopia Raman.

### **1.2.2 OBJETIVOS ESPECÍFICOS**

1. Aplicar uma rede neural, utilizando uma rede *Perceptron* com Múltiplas Camadas (PMC) como ferramenta de identificação de algumas espécies minerais do supergrupo do pirocloro (grupo pirocloro e microlita);
2. Treinar a rede PMC utilizando a aprendizagem supervisionada (isto é, os resultados de saída são previamente conhecidos), com o fluxo de informação iniciando pela camada de entrada e seguindo em direção à camada de saída, com as alterações dos pesos da rede pelo método de retropropagação do erro.
3. Realizar a correção da linha de base dos espectros Raman coletados por dois métodos (Mínimos Quadrados Assimétricos e Vizinhos Próximos)
4. Analisar e comparar o desempenho da rede proposta para classificação de espécies de minerais do supergrupo do pirocloro (grupo pirocloro e microlita) mediante espectros de espalhamento Raman corrigidos, sem filtragem estatística. Nesse caso, a RNA é alimentada diretamente com os espectros Raman obtidos experimentalmente.



### **1.3 MOTIVAÇÕES PARA UTILIZAÇÃO DE UMA RNA NA IDENTIFICAÇÃO DE MINERAIS**

Uma RNA é um modelo computacional programado para mimetizar a estrutura neural do funcionamento do cérebro humano a partir de uma estrutura interligada de neurônios produzidos artificialmente por um algoritmo, que funcionam como vias para a transferência de dados. A principal característica de uma RNA é a sua capacidade de aprendizagem abstrata, adquirindo experiência própria e aperfeiçoando seu desempenho. Nesse sentido, uma rede de neurônios artificiais pode ser treinada através de arquiteturas próprias, baseando-se em dados de espectros Raman experimental para obter um modelo genérico com o objetivo de identificar padrões nesses espectros.

Criar uma RNA para a classificação de espécies minerais a partir do espectro Raman pode apresentar as seguintes vantagens:

- O método introduz significativa rapidez na identificação de espécies minerais por especialistas em laboratório ou em campo utilizando-se equipamentos Raman portáteis;
- Pode indicar ao especialista uma direção, dentre as várias espécies a serem analisadas, para a identificação da amostra, visto que a rede pode não ser capaz de fornecer uma resposta exata, mas próxima da verdadeira;
- Permite o reconhecimento de similaridades entre os padrões de minerais e as amostras analisadas;
- A rede desenvolvida pode ser executada em computadores portáteis em conjunto com o Raman, cuja utilização não necessita de tratamento da amostra, possui rapidez na obtenção de resultados além da facilidade de transporte do instrumental, ou seja, pode ser utilizada em campo junto a um computador portátil para análise dos resultados.

### **1.4 ESTRUTURA DO TRABALHO**

Este trabalho é composto por 5 capítulos. No primeiro capítulo é apresentada a importância dos minerais do supergrupo do pirocloro (grupo pirocloro e microlita) na economia. São apresentados também os objetivos geral e específico da dissertação bem

como as principais motivações para a criação de uma rede neural artificial para a identificação de minerais.

No segundo capítulo são apresentados os fundamentos da espectroscopia Raman e das redes neurais artificiais assim como as principais aplicações dessas tecnologias apresentadas pela literatura especializada.

O capítulo 3 descreve os minerais estudados nesse trabalho, suas características físico-químicas e a metodologia utilizada na aquisição dos espectros Raman e seu tratamento para alimentação da RNA.

O capítulo 4 apresenta a análise dos resultados obtidos pela RNA.

O capítulo 5 apresenta as principais conclusões obtidas através da análise do capítulo 4 e as sugestões para trabalhos futuros.

## 2 FUNDAMENTOS DE ESPECTROSCOPIA RAMAN E REDES NEURAIAS

Neste capítulo é apresentada uma síntese de trabalhos envolvendo os conceitos do espalhamento Raman e os fundamentos das redes neurais artificiais, inicialmente de forma global e, posteriormente, salientando o seu emprego na área de caracterização de minerais.

### 2.1 A TEORIA DO ESPALHAMENTO RAMAN

A teoria do comportamento do espalhamento Raman, utilizando os princípios da mecânica quântica, foi primeiramente investigada por A. G. Smekal em 1923. No entanto, o fenômeno só foi observado experimentalmente em 1928 por K. S. Krishnan e C. V. Raman (recebendo o prêmio Nobel em 1930) e, de forma independente, por G. Landsberg e L. Mandelstam.<sup>20-23</sup>



Figura 3 - Da esquerda para a direita, os físicos Adolf Gustav Smekal, Chandrasekhara Venkata Raman, Kariamanickam Srinivasa Krishnan, Grigory Samuilovich Landsberg e Leonid Isaakovich Mandelstam.

Fonte: Elaborada pelo autor.

Quando fótons provenientes de uma fonte de radiação eletromagnética monocromática incidem sobre uma amostra, poderão ocorrer três eventos. No primeiro evento, os elétrons da amostra ganham energia e saem de seu estado fundamental, atingindo um estado vibracional virtual e depois decaem para o seu estado fundamental original com emissão de fótons com a mesma frequência dos fótons incidentes (espalhamento Rayleigh). No segundo, os elétrons ganham energia e saem de seu estado fundamental, atingem um estado vibracional virtual, mas ao decair permanecem em um estado fundamental superior, emitindo fótons com frequência menor do que a radiação incidente (espalhamento Raman Stokes). No terceiro, os elétrons situados em uma camada vibracional superior, atingem um estado vibracional virtual e decaem para um estado

vibracional fundamental abaixo do seu estado original, emitindo fótons com frequência superior a da radiação eletromagnética incidente (espalhamento Raman anti-Stokes). A Figura 4 ilustra os três eventos associados à interação entre a radiação eletromagnética e a matéria.

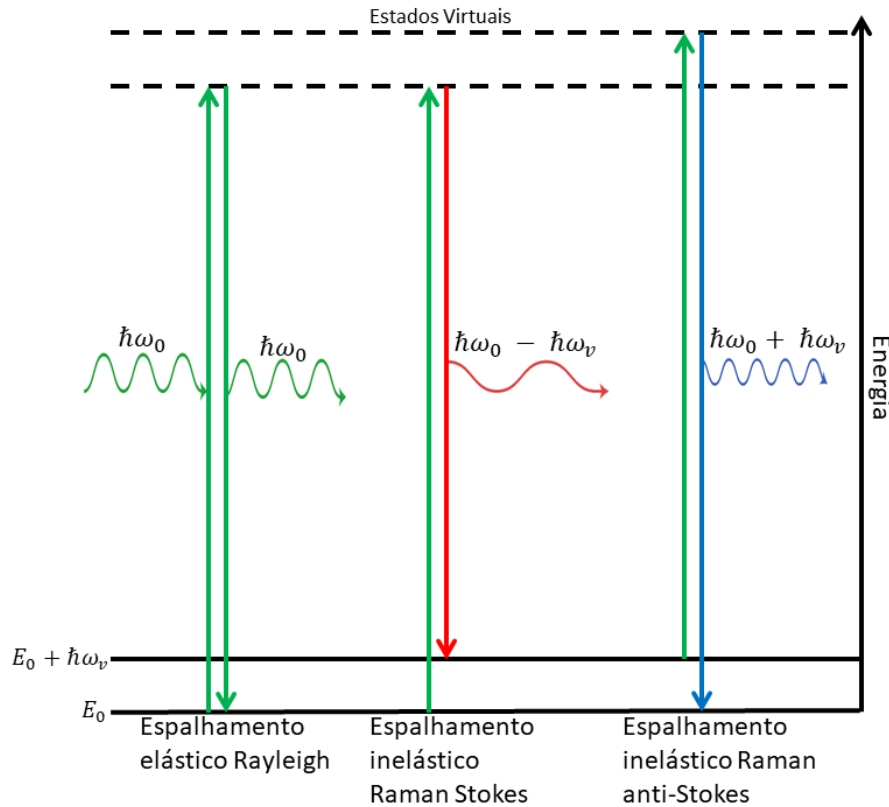


Figura 4 – Diagrama do espalhamento elástico (Rayleigh) e inelástico (Raman).

Fonte: Elaborada pelo autor.

Outro modo de se compreender o efeito Raman é através da análise do momento de dipolo induzido e pela interação fóton-elétron.

### 2.1.1 O MOMENTO DE DIPOLO INDUZIDO

O espalhamento resultante da interação da radiação eletromagnética monocromática com a matéria pode ser elástico (espalhamento Rayleigh) ou inelástico (espalhamento Raman). Neste último, ocorre a variação do momento de dipolo induzido na amostra pelo campo elétrico da radiação incidente e há uma mudança do estado vibracional da amostra.

O vetor de momento de dipolo induzido,  $\vec{P}$ , pode ser escrito como:

$$\vec{P} = \alpha\vec{E} \quad (1)$$

no qual  $E$  é o campo elétrico da radiação incidente e  $\alpha$  é a polarizabilidade da molécula. Ao expandir  $\alpha$  em série de Taylor em função da coordenada generalizada  $u$  se tem que<sup>24</sup>:

$$\alpha = \alpha_0 + \left(\frac{\partial\alpha}{\partial u}\right)_0 u + \dots \quad (2)$$

Supondo que o deslocamento  $\vec{u}$  da vibração e o campo elétrico  $\vec{E}$  são descritos em função das frequências de vibração  $\omega_v$  e da radiação incidente  $\omega_0$ , além dos vetores de onda  $\vec{k}_i$  e  $\vec{q}$  como:

$$\vec{u}(\vec{r}, t) = \vec{u}(\vec{q}, \omega_v) \cos(\vec{q} \cdot \vec{r} - \omega_v t) \text{ e } \vec{E}(\vec{r}, t) = \vec{E}(\vec{k}_i, \omega_0) \cos(\vec{k}_i \cdot \vec{r} - \omega_0 t) \quad (3)$$

e lembrando que  $\cos(a) \cos(b) = \frac{1}{2} [\cos(a+b) + \cos(a-b)]$ , o vetor de momento de dipolo induzido poderá ser escrito como:

$$\begin{aligned} \vec{P} = & \alpha_0 \vec{E}(\vec{k}_i, \omega_0) \cos(\vec{k}_i \cdot \vec{r} - \omega_0 t) \\ & + \frac{1}{2} \left(\frac{\partial\alpha}{\partial u}\right)_0 \vec{u}(\vec{q}, \omega_v) \vec{E}(\vec{k}_i, \omega_0) \{ \cos[(\vec{k}_i + \vec{q}) \cdot \vec{r} - (\omega_0 + \omega_v)t] \\ & + \cos[(\vec{k}_i - \vec{q}) \cdot \vec{r} - (\omega_0 - \omega_v)t] \} \end{aligned} \quad (4)$$

Os termos da equação 4 estão relacionados com os modos de espalhamento elástico (primeiro termo à direita) e inelástico (segundo termo), sendo este último dependente da existência da: variação da polarizabilidade em torno da posição de equilíbrio<sup>24</sup>, conservação de energia ( $\hbar\omega_e = \hbar\omega_0 \pm \hbar\omega_v$ ) e momento ( $k_e = k_0 \pm k_v$ ) entre a radiação incidente e a radiação espalhada (aqui simbolizado pelo subíndice “e”). Estas radiações espalhadas têm como frequências  $\omega_0 - \omega_v$  (espalhamento Raman Stokes) e  $\omega_0 + \omega_v$  (espalhamento Raman anti-Stokes).<sup>20,21</sup>

### 2.1.2 DIAGRAMA DO PROCESSO DE INTERAÇÃO ENTRE FÓTON-ELÉTRON

A Figura 5 ilustra um diagrama do processo de espalhamento Raman. Nesse caso, o fóton incidente com frequência  $\omega_0$  interage com o elétron (interação elétron-radiação, ER) que o transfere para o estado virtual de maior energia, criando um par elétron-buraco. Isto pode criar ou destruir fônons com frequência  $\omega_v$ , (interação elétron-rede, EL). Logo em seguida, o elétron se recombina com o buraco pela interação elétron-radiação e emite um fóton com frequência  $\omega_e$  (interação elétron-radiação) menor ou maior que  $\omega_0$ .<sup>20,25</sup>

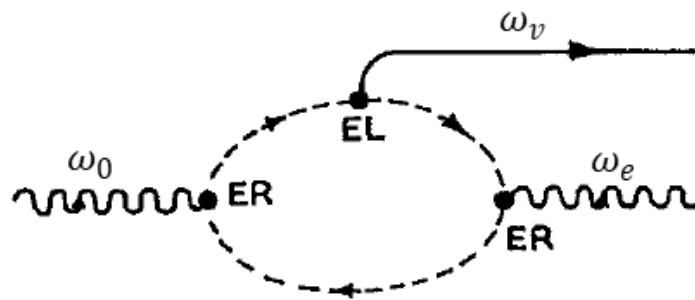


Figura 5 - Esquema do espalhamento Raman.

Fonte: adaptada de LOUDON.<sup>26</sup>

Utilizando-se a espectrometria Raman é possível obter-se espectros com informações sobre a composição química e estrutural de uma amostra de mineral. Porém, a interpretação dos dados obtidos pelo espectrômetro é complexa uma vez que uma determinada espécie pode conter diferentes tipos de impurezas, o que dificulta a identificação de características específicas para uma determinada espécie mineralógica. Neste trabalho, com o intuito de acelerar o processo de identificação de um mineral, utilizou-se um sistema embasado em Redes Neurais Artificiais para a classificação de espécies, em particular, do grupo do pirocloro e microlita; mediante a análise dos espectros Raman desses minerais.

## 2.2 REDES NEURAIAS ARTIFICIAIS

A rede neural artificial é um processador paralelamente distribuído, formado de pequenas unidades de processamento simples, com capacidade para armazenar conhecimento experimental e torná-lo disponível para uso e se assemelha ao cérebro humano por tais características.<sup>27</sup>

Essa característica de mapear os sistemas e aprender com os comportamentos do meio em que foi inserida, faz com que esta ferramenta se torne muito atrativa, principalmente na aplicação de classificação de padrões, na qual tem como objetivo associar um padrão de entrada para uma das classes previamente definidas.<sup>28</sup>

### 2.2.1 BREVE HISTÓRICO

Em 1943, McCulloch e Pitts publicam o primeiro artigo sobre redes neurais que descreve um cálculo lógico das redes neurais que unificava os estudos de neurofisiologia e

da lógica matemática. Seis anos mais tarde, veio o livro de Hebb “*The organization of Behavior*” propondo que conforme o organismo aprende em suas tarefas, as conexões do cérebro são alteradas e, nisso, são definidos os agrupamentos neurais.<sup>29</sup>

Após 15 anos do trabalho de McCulloch e Pitts, Rosenblatt propõe uma nova abordagem para o problema de reconhecimento de padrões em seu trabalho sobre *perceptron*, um método inovador de aprendizagem supervisionada que é capaz de reconhecer padrões simples. Em 1960, Widrow e Hoff introduzem o algoritmo do mínimo quadrado médio e a formulação do *Adaline* (elemento linear adaptativo), mais tarde, Widrow e seus estudantes propõem uma das primeiras redes neurais com múltiplos elementos adaptativos, isto é, múltiplos *Adaline* (Madaline).<sup>27,28</sup>

Desde os trabalhos de McCulloch e Pitts até fim da década de 1960, os pesquisadores ficaram incentivados pela possibilidade de que as redes neurais poderiam realizar qualquer coisa. Mas em 1969, Minsky e Papert publicam o livro “*Perceptrons – an introduction to computational geometry*”, demonstraram matematicamente a existência de limites fundamentais para aquilo que os *perceptrons* e *Adaline* poderiam calcular,<sup>27</sup> como a limitação de aprender a relação dos sinais de entrada e saída de funções simples como o ou-exclusivo.<sup>28</sup> Após o livro de Minsky e Papert, o número de trabalhos na área de redes neurais diminuiu junto com a redução drástica de grande parte do financiamento, sendo um período de Inverno da inteligência artificial.<sup>30</sup>

Apenas no final da década de 1980, as redes neurais artificiais voltaram a ter grande destaque por um conjunto de fatores, entre eles: as novas descobertas sobre o sistema nervoso biológico, o desenvolvimento de computadores com maior capacidade de processamento e o livro de Rumelhart, Hinton e Williams (“*Parallel distributed processing*” publicado em 1986). Neste livro, é desenvolvido o algoritmo *backpropagation* que permite ajustar os pesos sinápticos em uma rede de mais de uma camada e resolve o problema apontado por Minsky e Papert.<sup>28</sup>

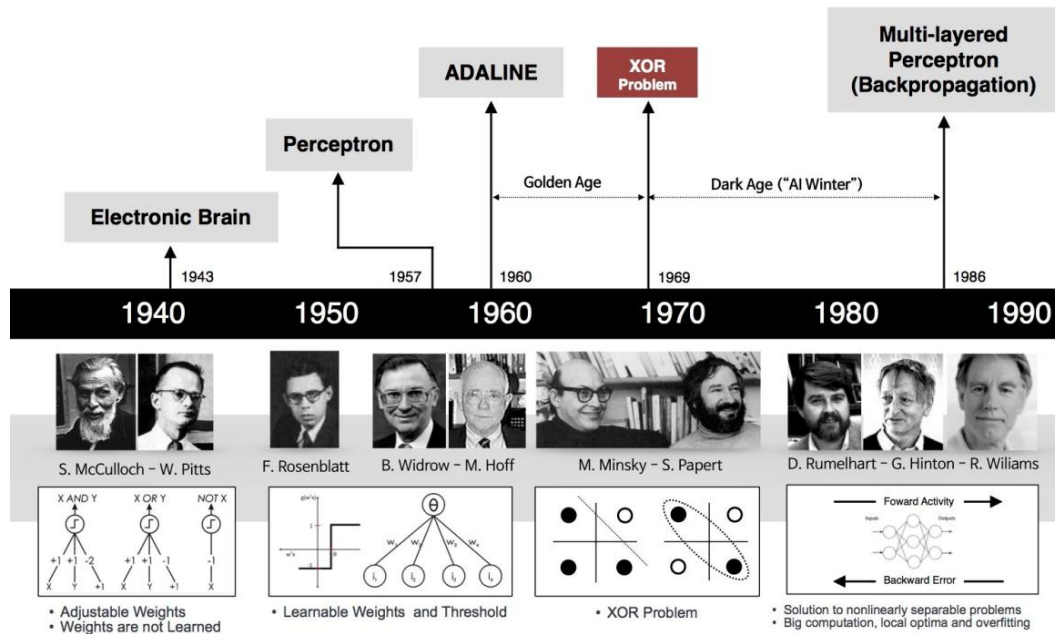


Figura 6 - Resumo histórico das redes neurais artificiais até a década de 1990.

Fonte: Adaptada de DEEP ...<sup>30</sup>

## 2.2.2 MODELO DE UM NEURÔNIO ARTIFICIAL

De modo geral, o neurônio biológico pode ser dividido em três partes; dendritos, corpo celular e axônio.<sup>31</sup> Na primeira parte, os dendritos captam estímulos provindos de outros neurônios ou do meio externo, em seguida, esses estímulos são direcionados ao corpo celular. Aqui, essas informações são processadas para produzir um potencial de ativação que será conduzido para o axônio. Nesta última etapa, o axônio conduz os impulsos elétricos para outros neurônios ou para o tecido muscular.<sup>28,31</sup> A Figura 7 ilustra a estrutura básica de um neurônio e sua visualização no microscópio eletrônico de varredura.



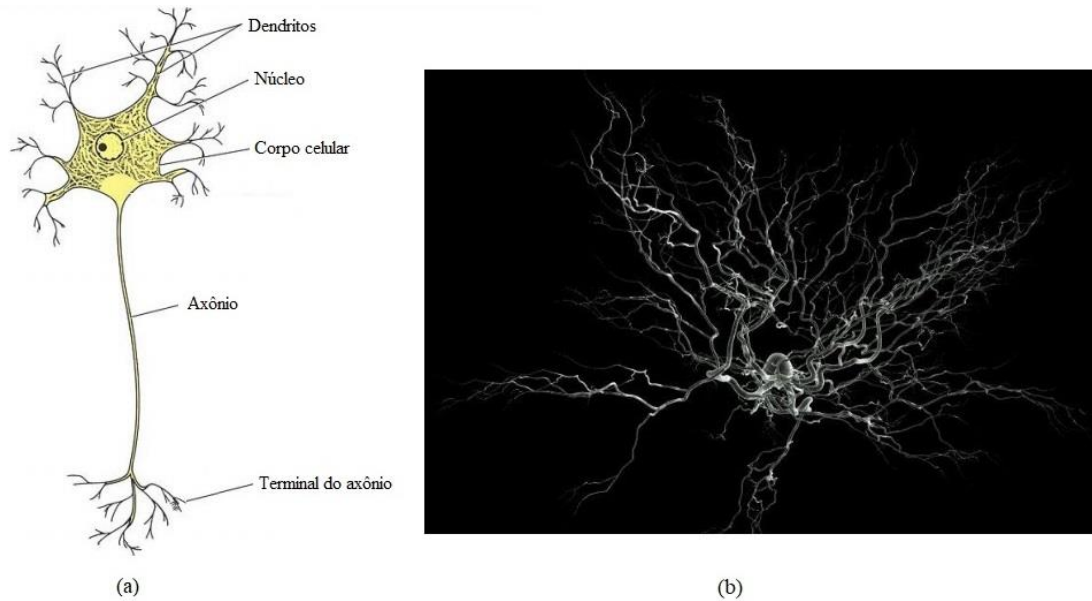


Figura 7 - Ilustração da estrutura básica de um neurônio biológico (a) e sua visualização ao microscópio eletrônico de varredura (b).

Fonte: Adaptada de SNELL,<sup>31</sup> ROUGIER.<sup>32</sup>

A Figura 8, ilustra como é o modelo do neurônio artificial baseado no sistema nervoso. Como em um neurônio biológico, as informações do meio formam os sinais de entrada. Os sinais entram no neurônio pelas sinapses e cada sinapse contém um peso próprio. Em seguida, no potencial de ativação, os sinais são ponderados e é aplicado o limiar de ativação, este especifica o intervalo apropriado para gerar o disparo ou sinal de saída do neurônio. Para limitar a amplitude do sinal de saída, é utilizada a função de ativação, possibilitando com que a saída permaneça dentro de um intervalo de valores predefinidos.<sup>27,28</sup>

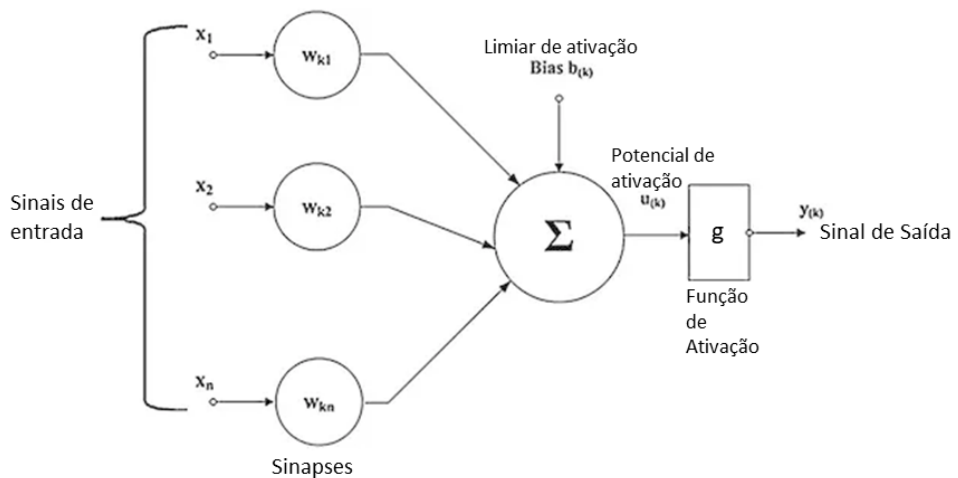


Figura 8 - Modelo de um neurônio artificial.

Fonte: Adaptada de DEEP ...<sup>30</sup>

### 2.2.3 REDE NEURAL PERCEPTRON DE MULTIPLAS CAMADAS

A Figura 9 ilustra uma rede neural de múltiplas camadas. A camada de entrada está posicionada mais à esquerda. Há outras duas camadas escondidas ou ocultas contendo cinco a três nós/neurônios em cada e a camada de saída mais direita com um neurônio. Nas redes *Perceptron* de múltiplas camadas (PMC), há pelo menos uma camada oculta.<sup>33-34</sup> O funcionamento é semelhante ao sistema nervoso. Os sinais de entrada entram na rede pela camada de entrada. As camadas seguintes, camadas ocultas, identificam as principais características dos sinais de entrada. Na última camada, camada de saída, é produzida a resposta final da rede. Este fluxo de informações segue em direção à saída (*feedforward*).

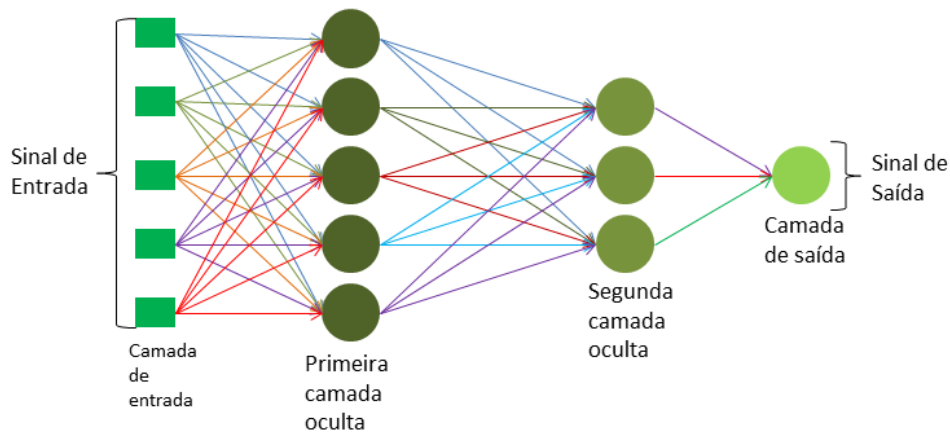


Figura 9 - Ilustração de uma rede de múltiplas camadas.

Fonte: Elaborada pelo autor

Para que uma rede neural artificial consiga entender e responder as principais características de um sistema na qual foi inserida é necessário dividir o processo em duas etapas: a etapa de treinamento e a etapa de testes. Na primeira etapa, ocorre à aprendizagem supervisionada, aqui são calculados desvios dos valores de saída gerados pela rede com as respectivas saídas desejadas (valores já conhecidos anteriormente). Os valores dos pesos sinápticos e limiares da camada de saída são alterados ao aplicar o método do gradiente descendente nesses desvios. Em seguida, os desvios se propagam para as camadas anteriores, alterando os pesos e limiares dessas camadas. Esta mudança na última camada para as camadas anteriores é caracterizada pelo método de retropropagação do erro (*backpropagation*). Isto possibilita a automática alteração da rede para cada interação, implicando a diminuição da soma dos desvios encontrados.<sup>28,35-36</sup> A Figura 10 ilustra o sentido dos fluxos em uma parcela de neurônios da rede. Na etapa de

teste, é observado se os valores de resposta da rede (ao utilizar outro conjunto de amostras não presentes na etapa anterior) correspondem ao comportamento do sistema.

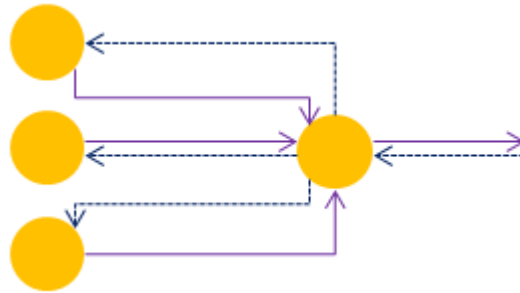


Figura 10 - Diagrama dos fluxos de dados presentes em uma rede PMC, as linhas contínuas representam o *feedforward* e as tracejadas o *backpropagation*.

Fonte: Adaptada de HAYKIN.<sup>27</sup>

A utilização de RNA tem se mostrado útil quando a relação de interdependência entre as variáveis dependentes e independentes é desconhecida ou muito complexa. Nas últimas décadas, pode ser observado um grande número de trabalhos utilizando a RNA para aplicações em áreas como controle de processos da indústria química, reconhecimento de imagens médicas, aperfeiçoamento dos sistemas de transporte e análise de fármacos e medicamentos.<sup>37-41</sup>

Após mais de duas décadas, um número considerável de trabalhos tem demonstrado a viabilidade da utilização das RNA's na interpretação qualitativa e quantitativa de espectros na região do infravermelho (IR).<sup>42-50</sup> Entretanto, a espectroscopia IR apresenta algumas limitações que podem restringir a sua utilização na caracterização de minerais. A espectroscopia IR baseia-se na absorção de radiação, enquanto que a espectroscopia Raman baseia-se no espalhamento.<sup>51</sup>

A caracterização de minerais por IR requer a necessidade do preparo da amostra, muitas vezes através da sua diluição em brometo de potássio (KBr) seguido do preparo de uma pastilha. Este procedimento é um fator limitante da técnica, uma vez que destrói a amostra.<sup>52</sup> Também deve ser ressaltado, que o infravermelho tem sido utilizado principalmente na identificação de água presente na estrutura da amostra.<sup>7,53</sup>

Nesse caso, espectrômetros Raman associados a microscópios permitem a identificação de componentes a partir de alguns micrômetros em misturas heterogêneas, como é o caso dos minerais; enquanto que nos espectrômetros IR convencionais a área mínima investigada é de algumas dezenas de micrômetros.<sup>54</sup>

A literatura sobre a utilização da espectroscopia Raman na identificação de espécies mineralógicas, terrestres ou extraterrestres também é vasta.<sup>55-61</sup> Entretanto, sua aplicação na caracterização de minerais do supergrupo do pirocloro parece ser ainda ser inicial com relação à identificação de minerais brasileiros.<sup>7,62-63</sup> Até o presente momento não foi identificado trabalhos sobre a utilização de redes neurais artificiais como ferramenta na interpretação quantitativa de espectros Raman de minerais do supergrupo do pirocloro, o que reforça a motivação para o desenvolvimento desse trabalho.

### 3 MATERIAIS E MÉTODOS

#### 3.1 AMOSTRAS UTILIZADAS

As amostras dos minerais estudados e de referência (que serão utilizadas na etapa de treinamento nas RNA) neste trabalho (Tabela 1 e Tabela 2). As amostras fornecidas não sofreram nenhum tipo de tratamento prévio às análises espectroscópicas.

Os parâmetros cristalográficos das amostras foram determinados anteriormente através do uso dos difratômetros de raios X por monocristal com radiação  $\text{MoK}\alpha$  (0,71073 Å): Bruker X8 APEX2 CCD (Universidade do Arizona) e Enraf-Nonius Kappa-CCD (IFSC-USP).

Tabela 1 - Características das amostras dos minerais utilizados nesse trabalho.

<b>Código</b>	<b>Mineral</b>	<b>Ocorrência</b>	<b>Parâmetros de Cella</b>
PY1_AN1	fluornatropirocloro	Ilha dos Açores, Portugal	a = 10.4240(9) Å, b = 10.4240(9) Å, c = 10.4240(9) Å Volume = 1132.68(9) Å <sup>3</sup>
PY_AN02	fluorcalciopirocloro	Vishevogorsk, Ural, Rússia	a = 10.4240(9) Å, b = 10.4240(9) Å, c = 10.4240(9) Å Volume = 1132.68(9) Å <sup>3</sup>
PY_AN3	fluorcalciopirocloro	Mina St Lawrence Columbium, Canadá	a = 10.453(3) Å, b = 10.453(3) Å, c = 10.453(3) Å Volume: 1142.3(2) Å <sup>3</sup>
MI_AN4	fluorcalciomicrolita	Lavra do Ipê, Marilac, Minas Gerais	a = 10.4045(19) Å, b = 10.4045(19) Å, c = 10.4045(19) Å Volume = 1126.32(36) Å <sup>3</sup>
FLUMICRO _AN05	fluornatromicrolita	Mina Quixaba, Frei Martinho, Paraíba	a = 10.4361(1) Å, b = 10.4361(1) Å, c = 10.4361(1) Å Volume = 1136,62(3) Å <sup>3</sup>
MICR_AN06	fluorcalciomicrolita	Mina Urubu, Itinga, Minas Gerais	a = 10.4203(4) Å, b = 10.4203(4) Å, c = 10.4203(4) Å Volume = 1131.46(13) Å <sup>3</sup>

Fonte: Elaborada pelo autor

Tabela 2 – Fórmula química das amostras dos minerais utilizados nesse trabalho\*, de grupo espacial Fd $\bar{3}m$ .

<b>Código</b>	PY1_AN1
<b>Fórmula Química</b>	$(Na_{0,92}Ca_{0,72}Ce_{0,12}\square_{0,06}Nd_{0,05}U_{0,05}La_{0,04}Pr_{0,02}Y_{0,01}Th_{0,01})_{\Sigma=2}(Nb_{1,57}Ti_{0,35}Ta_{0,06}Zr_{0,02})_{\Sigma=2}(O_{5,96}OH_{0,04})_{\Sigma=6}(F_{0,80}OH_{0,20})_{\Sigma=1}$
<b>Código</b>	PY_AN02
<b>Fórmula Química</b>	$(Ca_{1,04}Na_{0,80}\square_{0,07}Ce_{0,04}La_{0,02}Sr_{0,02}Nd_{0,01})_{\Sigma=2}(Nb_{1,78}Ti_{0,22})_{\Sigma=2}(O_{5,92}OH_{0,08})_{\Sigma=6}(F_{0,91}OH_{0,09})_{\Sigma=1}$
<b>Código</b>	PY_AN3
<b>Fórmula Química</b>	$(Ca_{1,17}Na_{0,49}Ce_{0,19}Mn_{0,03}La_{0,03}Pr_{0,01}Th_{0,01})(Nb_{1,55}Ti_{0,31}Fe_{0,09}Zr_{0,04}Ta_{0,01})O_6(F_{0,63}OH_{0,37})$
<b>Código</b>	ML_AN4
<b>Fórmula Química</b>	$(Ca_{1,50}Na_{0,35}\square_{0,09}Mn_{0,04}Sn_{0,01}Fe_{0,01})_{\Sigma=2}(Ta_{1,77}Nb_{0,15}Ti_{0,07}Si_{0,01})_{\Sigma=2}(O_{6,00})_{\Sigma=6}(F_{0,57}O_{0,41}OH_{0,02})_{\Sigma=1}$
<b>Código</b>	FLUMICRO_AN05
<b>Fórmula Química</b>	$(Na_{1,02}Ca_{0,70}Bi_{0,18})_{\Sigma=1,90}Ta_2O_6(F_{0,93}O_{0,02})_{\Sigma=0,95}$
<b>Código</b>	MICR_AN06
<b>Fórmula Química</b>	$(Ca_{0,98}Na_{0,92})_{\Sigma=1,9}(Ta_{1,88}Nb_{0,11}Sn_{0,01})_{\Sigma=2}O_6(F_{0,86}OH_{0,01})_{\Sigma=0,87}$

\* A representação de vacância na estrutura cristalina é dada por ( $\square$ ).

Fonte: Elaborada pelo autor

A composição química de cada amostra foi determinada previamente em colaboração com o Lunar and Planetary Laboratory da Universidade do Arizona. As amostras foram analisadas por meio de microsonda eletrônica CAMECA SX100 acoplada a um sistema de dispersão de comprimento de onda (Wavelength Dispersion System - WDS) e feixe com diâmetro da ordem de 1  $\mu\text{m}$ .

A partir dos dados de microsonda, as fórmulas químicas das espécies minerais foram obtidas com o auxílio do programa MinCalc (Mineralogical Calculations) e são apresentadas na Tabela 2.

Adicionalmente, alguns dados de microsonda eletrônica e de Raman originários do projeto RRUFF<sup>19</sup> também foram acrescentados, amostras: PY\_AN3 (número RRUFF R110198), FLUMICRO\_AN05 (R060589) e MICRO\_AN06 (R050625).

### 3.2 O ESPECTRÔMETRO RAMAN

O espectrômetro Raman utilizado nesse trabalho foi o HORIBA LabRam HR UV-Visível-NIR (200-2100 nm) do laboratório multiusuários (EMU-LabRam) que faz parte Centro de Caracterização de Espécies Minerais (CCEM) do Instituto de Física de São Carlos (IFSC). O equipamento possui distancia focal de 800 mm, resolução de  $1\text{cm}^{-1}/\text{pixel}$  e detector CCD resfriado termoeletricamente (Peltier).<sup>64</sup> O espectrômetro está equipado com três lasers para excitação das amostras: 633 nm (He-Ne) e dois lasers de diodo, 532 nm e 785 nm.

A coleta dos espectros característicos foi realizada com duas grades de difração (600 e 1800 linhas/mm) com resolução espacial de aproximadamente 1  $\mu\text{m}$  e com resolução espectral de  $1\text{cm}^{-1}$ . A Figura 11 mostra o espectrômetro utilizado nesse trabalho e o esquema do caminho ótico do feixe de laser desde a fonte, sua interação com a amostra e até a sua detecção.

As amostras são posicionadas no estágio motorizado (presente na Figura 11 (a)) e suas superfícies são visualizadas pelo microscópio óptico com uma objetiva de 10 vezes. Após encontrar a região mais plana e sem grandes irregularidades, é utilizada a objetiva de 50 vezes de trabalho a longa distância para ampliar a visualização dessa região. Verifica se não há mais alguma irregularidade na superfície e realiza-se a medida. Antes de realizar a análise, é selecionado o laser 532 nm com tempo de exposição de 1 s e altera-se sua potência de 0,01 até 100%. Ao observar a intensidade dos picos, da ordem de alguns milhares de contagens, verifica-se a razão sinal ruído. Se não houver degradação da

amostra, procede-se a coleta do espectro. As mesmas etapas são seguidas para os demais lasers.

Na Figura 11 (b) e (c), é ilustrado o caminho percorrido do feixe de laser e da radiação espalhada pela amostra. O feixe de laser atravessa o polarizador e o filtro de absorção de intensidade. O feixe resultante passar por um filtro do tipo *edge* e completamente refletido em direção à amostra. A radiação espalhada pela amostra sofre retroespalhamento, em  $180^\circ$ . Por fim, a radiação é encaminhada para uma grade de difração que a dispersa e a envia para um dispositivo de carga acoplada (detector CCD) que registra a intensidade (Figura 11 (c)).



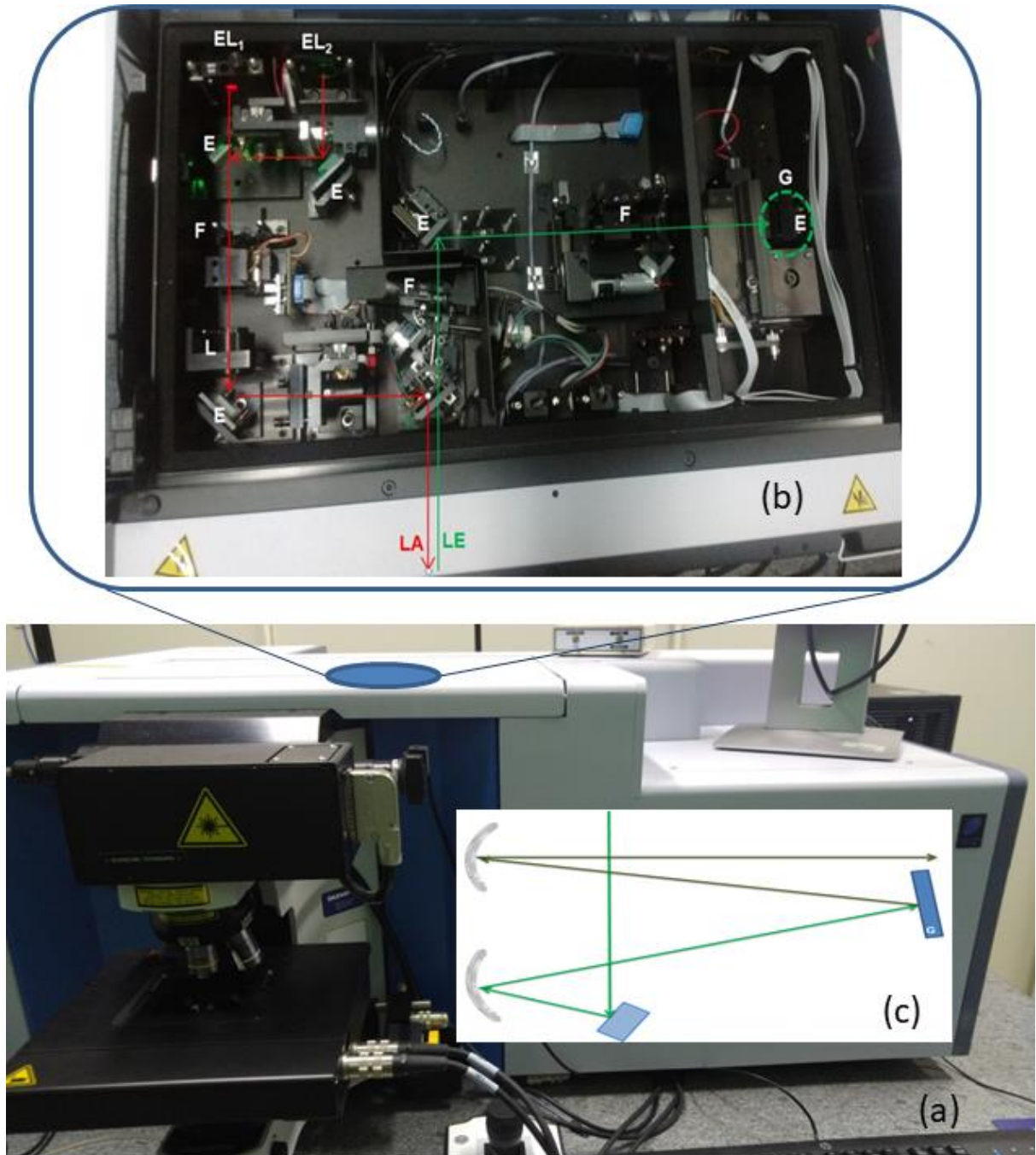


Figura 11 – (a) Espectrômetro Raman da Horiba, modelo LabRam HR UV-Visível-NIR (200-2100 nm) do Centro de Caracterização de Espécies Minerais (CEM) do Instituto de Física de São Carlos (IFSC). É mostrado em (b): Entrada dos lasers de 532 nm (EL<sub>2</sub>), 785 nm (EL<sub>2</sub>) e 633 nm (EL<sub>1</sub>). F = filtro, E = espelho, L = lente, LA = Laser incidente na amostra, LE = Laser espalhado pela amostra e G = indica a direção do feixe para a grade de difração que encaminhará o sinal para o detector CCD. Em (c) é ilustrado o caminho óptico do feixe até o detector.

Fonte: Elaborada pelo autor.

### 3.3 TRATAMENTO COMPUTACIONAL DOS DADOS

Os dados do espalhamento Raman foram coletados utilizando o equipamento LabRAM HR UV-Vísible-NIR da Horiba. Para diferenciar o sinal do espalhamento Raman de efeitos, como fluorescência, foram utilizados os lasers 532, 633 e 785 nm.<sup>65</sup> Assim, se uma determinada banda se repetir, significa que independe do comprimento de onda do laser incidente e pode ser considerada uma típica banda Raman.

Os espectros foram divididos em três grupos de regiões: o grupo de 100-1400 cm<sup>-1</sup> (principal região de impressão digital, “*fingerprint*”, dos minerais), de 1200-1800 cm<sup>-1</sup> (contendo a região da banda de deformação de H<sub>2</sub>O<sup>66-67</sup>) e de 2800-4000 cm<sup>-1</sup> (região das bandas de estiramento de H<sub>2</sub>O e OH<sup>68-70</sup>). Alguns espectros de minerais dos grupos do pirocloro e microlita, presentes da base de dados do projeto RRUFF, foram acrescentados ao trabalho. O projeto RRUFF contém uma base de dados de minerais com ênfase na disponibilização de mais de 14 mil espectros Raman de alta qualidade, além de incluir análises por: microsonda eletrônica, difração de raios X de pó e por monocristal.<sup>19,71</sup>

Em cada espectro, foi realizada a correção da linha de base se baseando em dois métodos: Mínimos Quadrados Assimétricos (“Asymmetric Least Squares” – ALS) e por Vizinhos Próximos (VP).<sup>72-73</sup> Na implementação do ALS, a linha de base foi calculada até que diferença entre as duas últimas interações fosse menor que 0,001. Após as correções, foi utilizada a função Savitzky-Golay para suavizar os espectros. Cada um destes passos foi implementado usando a linguagem python 2.7 e a função “*savgol\_filter*” da biblioteca *scipy*.<sup>74</sup>

Na implementação do método ALS, se busca minimizar o resíduo ( $S$ ) da diferença entre os elementos do espectro medido ( $y_i$ ) com os elementos da linha de base ( $z_i$ ) presentes na equação 5:

$$S = \sum_i \omega_i (y_i - z_i)^2 + \lambda \sum_i (z_i - 2z_{i-1} + z_{i-2})^2 \quad (5)$$

Os parâmetros  $\omega_i$  e  $\lambda$  representam o peso e o grau de suavização respectivamente. O primeiro altera seu valor conforme as seguintes condições: se  $y_i > z_i$  então  $\omega_i = p$  ( $p$ , parâmetro de assimetria), caso contrário  $\omega_i = 1 - p$ . O parâmetro  $\lambda$  pode assumir valores entre  $10^2$  a  $10^9$  e  $p$  valores entre 0.001 a 0.1. Isso faz com que as regiões de pico apresentem peso menor do que as outras regiões.<sup>72</sup>

Na implementação do método por VP, equação 6<sup>73</sup>:

$$\phi(x) = \frac{[f(x_i) - f(x_{i-1})]}{x_i - x_{i-1}} \left[ x - \frac{x_i + x_{i-1}}{2} \right] + \frac{f(x_i) + f(x_{i-1})}{2} \quad (6)$$

Inicialmente, as posições dos mínimos locais são encontradas no espectro medido  $f(x)$ , formando os elementos  $x_{i-1}$  e  $x_i$ . Após isso, a linha de base  $\phi(x)$  é calculada, na qual  $x$  é um dos elementos do espectro medido entre os pontos de mínimos locais ( $x_{i-1}$  e  $x_i$ ).

Como os espectros Raman processados foram medidos com diferentes lasers e grades de difração, o número total de pontos é diferente em cada espectro. Uma vez que, estes espectros são os valores de entrada da rede, é necessário padronizar essa quantidade de pontos. Para isso, cada espectro, em seus respectivos grupos de regiões, teve a média de seus pontos calculada a cada  $0.5 \text{ cm}^{-1}$ . Como última etapa neste processamento, os espectros foram normalizados conforme a equação 7, para que a grandeza da intensidade não influencie no aprendizado da rede neural e sim o comportamento espectral da amostra:

$$Y = \frac{y - y_{min}}{y_{max} - y_{min}} \quad (7)$$

sendo  $Y$  o valor do espectro normalizado,  $y_{min}$  e  $y_{max}$  os valores mínimo e máximo do espectro  $y$  a ser normalizado entre zero e um, respectivamente.

### 3.4 MONTAGEM TOPOLÓGICA DA REDE NEURAL ARTIFICIAL

Para estruturar as redes neurais, foram utilizados os softwares *Tensorflow* e *Keras*. *Tensorflow* é uma biblioteca de código aberto para computação numérica e aprendizado de máquina, desenvolvida pela *Google Brain team*.<sup>75-76</sup> *Keras* é uma biblioteca de alto nível, escrita em python, focada na formação de rede de redes neurais e desenvolvida inicialmente pela ONEIROS (*Open-ended Neuro-Electronic Intelligent Robot Operating System*).<sup>77</sup>

Como os espectros Raman normalizados são os valores de entrada da rede, o número de neurônios na camada de entrada segue o mesmo número de pontos nos espectros, na ordem de alguns milhares.

A camada de saída, formada por cinco neurônios, se divide em cinco classes (fluornatropirocloro – classe 0, fluorcalciopirocloro – 1, fluorcalciomicrolita – 2, fluornatromicrolita – 3, outros – 4). A classe “outros” se refere a qualquer outra espécie de

mineral diferente das outras classes. Para fazer essa classificação em cinco classes, a camada de saída possui cinco neurônios e a função *softmax* como função de ativação. A função *softmax*, equação 8, gera a probabilidade de que o espectro de entrada seja uma das classes definidas com valores entre 0 e 1:

$$\sigma(\mathbf{z})_j = \frac{e^{z_j}}{\sum_{i=1}^J e^{z_i}} \quad (8)$$

para  $j = 1, \dots, J$ ,  $J$  o número de classes,  $z_j$  é a pontuação (“*score*”) de ser uma das classes.<sup>30,75,78</sup>

Nas camadas ocultas, utilizou-se três funções de ativação (logística, tangente hiperbólica e retificadora ou ReLU – Rectified Linear Unit, uma de cada vez) e variou-se de forma crescente o número de camadas ocultas, com dez neurônios em cada camada, observando com quantas camadas ocultas se obtinha maior acurácia e menor valor da função custo. Após isso, variou-se o número de neurônios na última camada oculta, procurando também observar o caso de maior acurácia e menor função custo.

A função logística, equação 9 a), possui um formato similar da função degrau e varia entre zero e um. A tangente hiperbólica, 9 b), é semelhante à função anterior e assume valores entre -1 e 1. Na última função de ativação, ReLU – 9 c), não satura e facilita o processamento por não ativar todos os neurônios ao mesmo tempo.<sup>28,30,75</sup>

$$f(x) = \frac{1}{1 + e^{-x}} \quad (9 \text{ a})$$

$$f(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (9 \text{ b})$$

$$f(x) = \max(\mathbf{0}, x) \quad (9 \text{ c})$$

Durante a etapa de treinamento, foi utilizado o algoritmo Gradiente Descendente Estocástico (*Stochastic Gradient Descent* – SGD) e, como função custo, a entropia cruzada (*cross-entropy*, equação 10); que mede quanto um conjunto de probabilidades das classes estimadas coincide com as classes desejadas e converge rapidamente a aprendizagem da rede<sup>30,75</sup>:

$$H = -\frac{1}{n} \sum_{i=1}^n \sum_{j=1}^J y_j^{(i)} \log(\sigma(\mathbf{z})_j^{(i)}) \quad (10)$$

sendo  $n$  o número total de instâncias,  $J$  é o número total de classes,  $y_j^{(i)}$  igual a 1 se a classe desejada da instância  $i$  é  $j$  ou zero caso contrário,  $\sigma(z)_j^{(i)}$  valor da função *softmax* da instância  $i$ .

Os espectros processados das amostras dos grupos do pirocloro e microlita (PY1\_AN1, PY\_AN02, PY\_AN3, MI\_AN4, FLUMICRO\_AN05, MICR\_AN06) e outros minerais (calcita – R150075, fluorcalcioromeíta – R120140, hidrokenoralstonita – R060012, kenoplumbomicrolita – R130729, muscovita – R040104 e rutilo – R060493) formaram o conjunto de minerais para a etapa de treinamento.

Para testar a rede durante o período de treinamento, o segundo conjunto de dados (conjunto de teste) foi formado por espectros presentes no projeto RRUFF. As espécies utilizadas foram: fluorcalciopirocloro (R060288, R061003, R110198), fluornatromicrolita (R130021), fluorcalcio microlita (R080146, R120127), calcita (R040070, R050048), quartzo (R040031) e rutilo (R050031).

Um terceiro conjunto de espectros foi utilizado para realizar a validação das redes, ao verificar se as redes estão realizando a correta classificação dos espectros. Este conjunto foi constituído por espécies nas quais, até o momento, se sabe apenas os grupos a que pertencem (microlita ou pirocloro) e minerais (calcita, quartzo e rutilo) de outras bases de dados (“Laboratoy of Photoinduced Effects Vibrational and X-ray Spectroscopies”- Universidade de Parma <sup>79</sup> e “Handbook of Raman Spectra for geology”- Laboratório de Geologia de Lyon <sup>80</sup>).



## 4 RESULTADOS E DISCUSSÃO

### 4.1 CARACTERÍSTICAS DO ESPECTRO RAMAN

Depois de realizar a coleta dos espectros de espalhamento das amostras, foi observado que algumas bandas se repetiram independente do comprimento de onda do laser utilizado (como exemplificado na Figura 12, em verde e vermelho), isso leva a considerar que esses sinais são bandas Raman. Há casos em que bandas não se repetem para os outros lasers, como algumas bandas exibidas em preto na Figura 12, estes sinais podem ser formados por efeitos de fluorescência das impurezas da amostra. As impurezas também podem alterar o formato ou gerar novas bandas nos espectros.

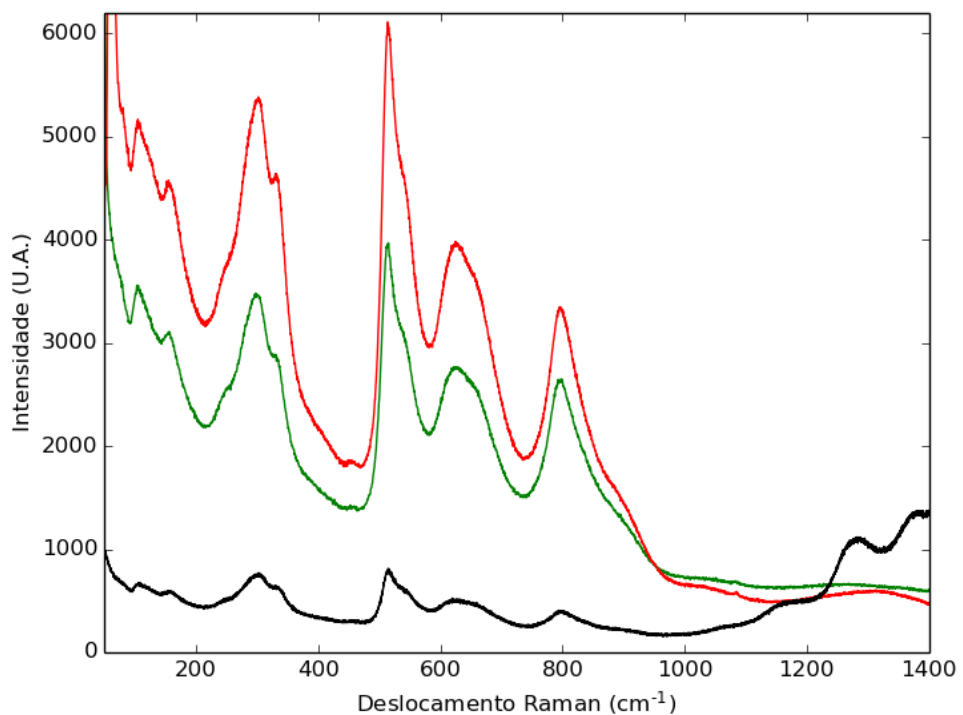


Figura 12 - Espectro Raman bruto na região de 50 a 1400  $\text{cm}^{-1}$  da amostra MI\_AN4 excitadas pelos lasers 532 (verde), 633 (vermelho) e 785 (preto) nm.

Fonte: Elaborada pelo autor.

Para conseguir separar os sinais Raman dos outros efeitos e distinguir quais as bandas características, foi realizada a remoção do ruído de fundo (“background”) pela correção da linha de base se baseando nos métodos Mínimos Quadrados Assimétricos (ALS) e Vizinhos Próximos (VP).

#### 4.1.1 REMOÇÃO DO RUÍDO DE FUNDO

Para realizar a remoção do ruído de fundo e o aprendizado das redes neurais artificiais, foi utilizado o computador pessoal com processador AMD Ryzen 7 1700 e 16GB de memória RAM. A Figura 13 exemplifica as linhas de base calculadas para os espectros deste trabalho. Quando há uma variação abrupta no espectro ou a presença de elevado ruído (como nos gráficos em (a), (b), (c), (e) e (f)), a linha de base pela ALS implementada se distancia do espectro o que resulta em formação de bandas não existentes originalmente. Dependendo muito da quantidade de vizinhos próximos e do ruído no espectro, como em (c), (d) e (e), em VP pode ocorrer à remoção em excesso do *background*. Em ambas as correções, há casos que ocorre a perda de pontos do espectro medido, só que para a linha de base por VP apresenta uma maior perda de informação, como visto em (b), (d), (e) e (f).

Quando comparado o tempo de processamento, VP foi mais rápido que ALS, uma vez que o tempo gasto para executar a remoção por VP foi de menos de 1 segundo enquanto por ALS o tempo variava entre 5 minutos (nos casos quando se tinha menor ruído e também menos pontos no espectro) até 3 horas (para os espectros de muitos pontos ou quando a linha de base não atingisse a diferença entre as interações).



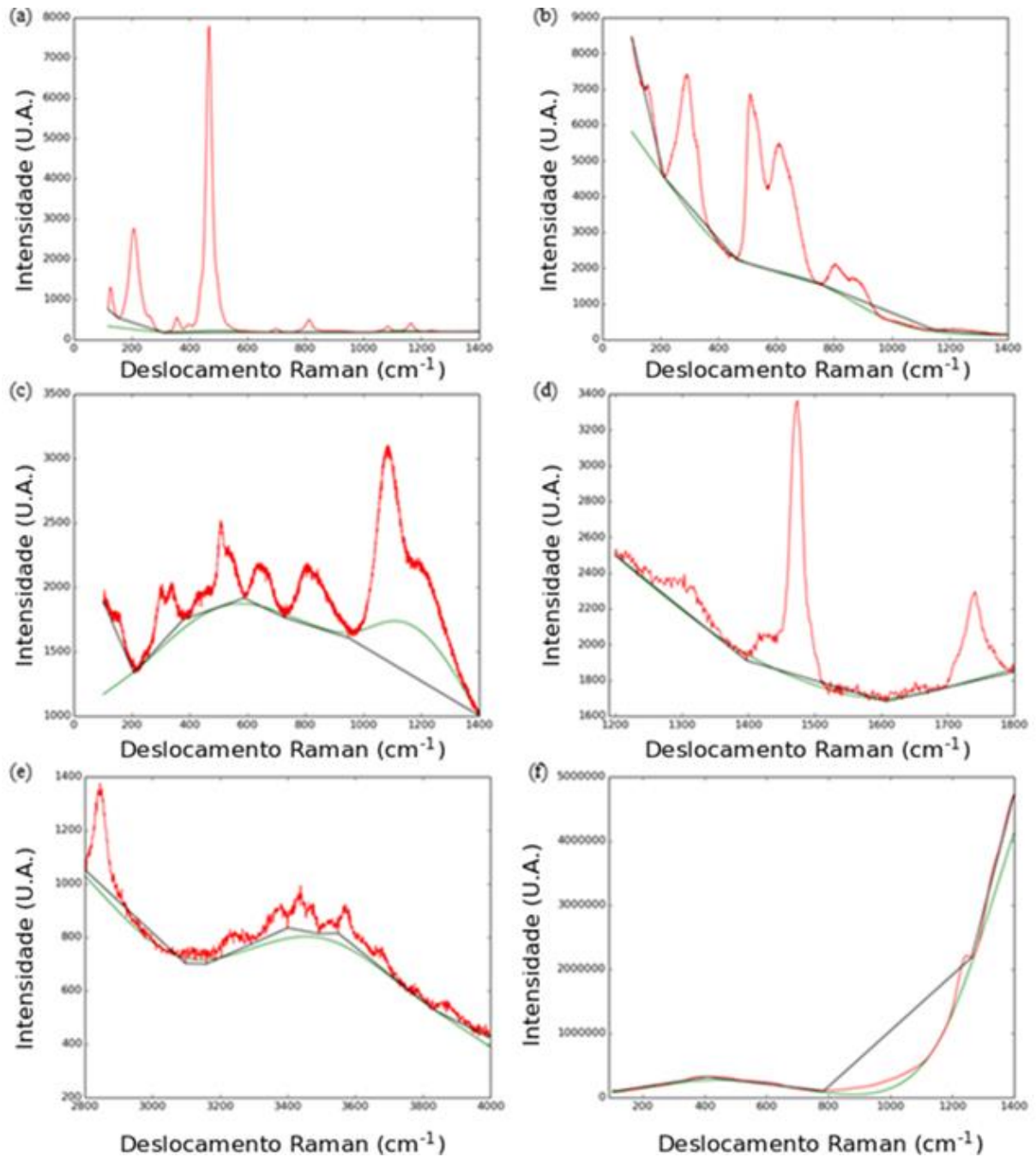


Figura 13 - Espectros de Intensidade (U.A.) em função do Deslocamento Raman ( $\text{cm}^{-1}$ ) sem correção (em vermelho) e com as linhas de base calculadas (Implementação de ALS e VP em verde e preto, respectivamente).

Fonte: Elaborada pelo autor.

## 4.2 AMOSTRAS DE TREINAMENTO

A análise dos espectros Raman do supergrupo do pirocloro se divide em seis regiões: vibrações da rede ( $100\text{--}200\text{ cm}^{-1}$ ), deformação angular ( $200\text{--}500\text{ cm}^{-1}$ ), estiramento

simétrico e antissimétrico ( $500\text{--}800\text{ cm}^{-1}$ ), sobretons ( $800\text{--}1400\text{ cm}^{-1}$ ), deformação angular de  $\text{H}_2\text{O}$  ( $1400\text{--}1800\text{ cm}^{-1}$ ) e estiramento de OH ( $2800\text{--}4000\text{ cm}^{-1}$ ).<sup>62,81-83</sup>

A determinação teórica das bandas Raman prevista por análise de fator de grupo e cálculos *ab initio* para amostras do supergrupo do pirocloro apresenta seis modos vibracionais.<sup>84-87</sup> No entanto, como visto nas figuras 14-37, há mais de seis modos vibracionais obtidos experimentalmente.

Nos gráficos, a cor verde representa o espectro medido com o laser 532 nm, vermelho o 633 nm e preto 785 nm.

#### 4.2.1 VIBRAÇÃO DA REDE

As bandas abaixo de  $200\text{ cm}^{-1}$  são complexas e provavelmente correspondem as vibrações de rede. Nos espectros coletados para o grupo do pirocloro (PY1\_AN1 e PY\_AN02), foram identificadas duas bandas, em  $130$  e  $170\text{ cm}^{-1}$  nas Figura 14 e 15. Para PY\_AN3 na Figura 16, a linha de base por VP removeu grande parte do espectro, e isto associado com o ruído ainda presente dificulta a identificação das duas ou demais bandas.

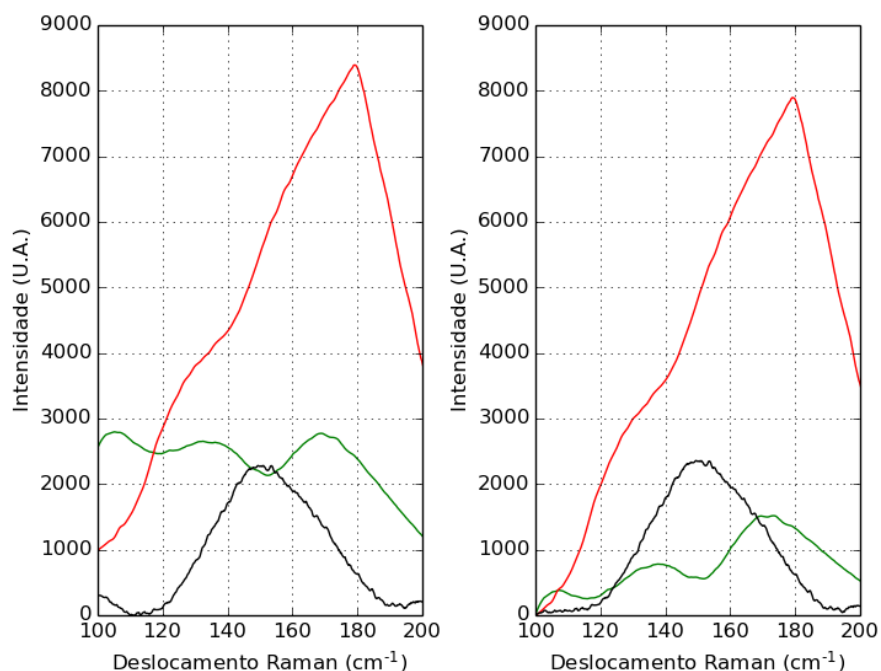


Figura 14 - Espectros Raman da espécie PY1\_AN1, na região de  $100\text{--}200\text{ cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

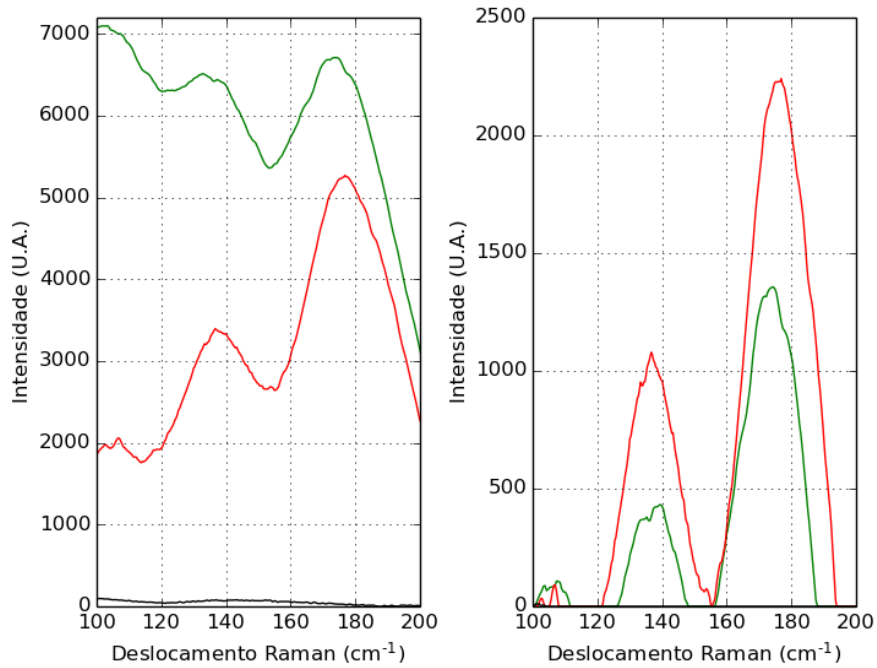


Figura 15 - Espectros Raman da espécie PY\_AN02, na região de 100-200  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

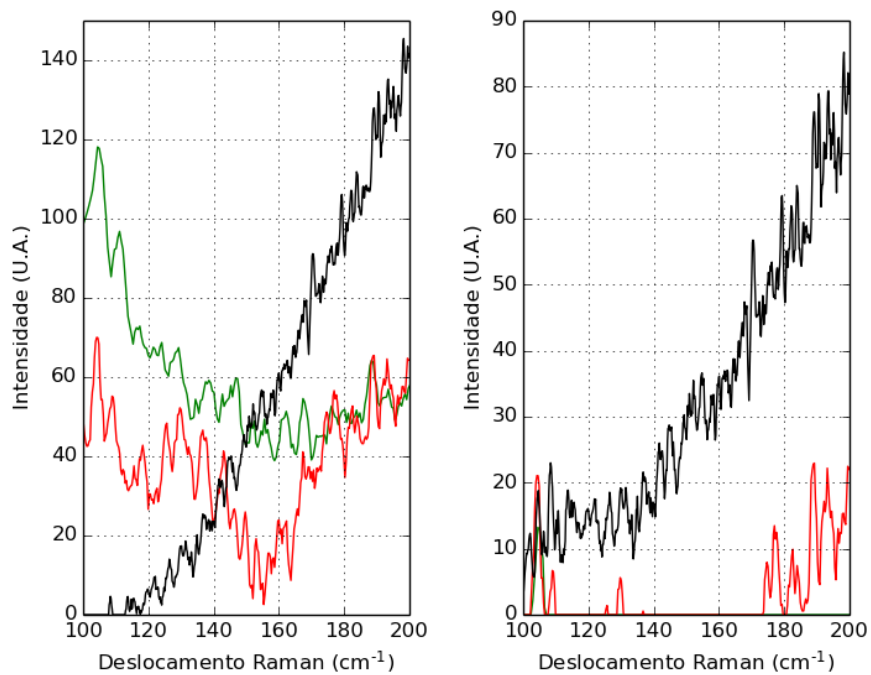


Figura 16 - Espectros Raman da espécie PY\_AN3, na região de 100-200  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

Já para as espécies do grupo da microlita (MI\_AN4, FLUMICRO\_AN05 e MICR\_AN06), os espectros exibem duas bandas, próximas de 100 e de 160  $\text{cm}^{-1}$ , como

visto nas Figura 17, Figura 18 e 19. É importante observar que as bandas da espécie MI\_AN4 se repetem para os três lasers utilizados, isto indica que são bandas Raman.

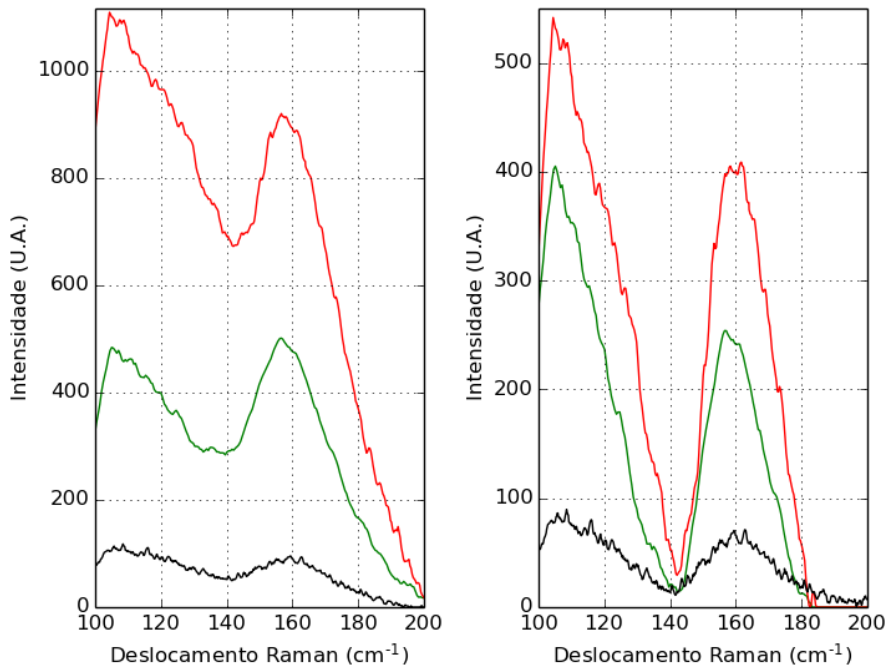


Figura 17 - Espectros Raman da espécie MI\_AN4, na região de 100-200  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

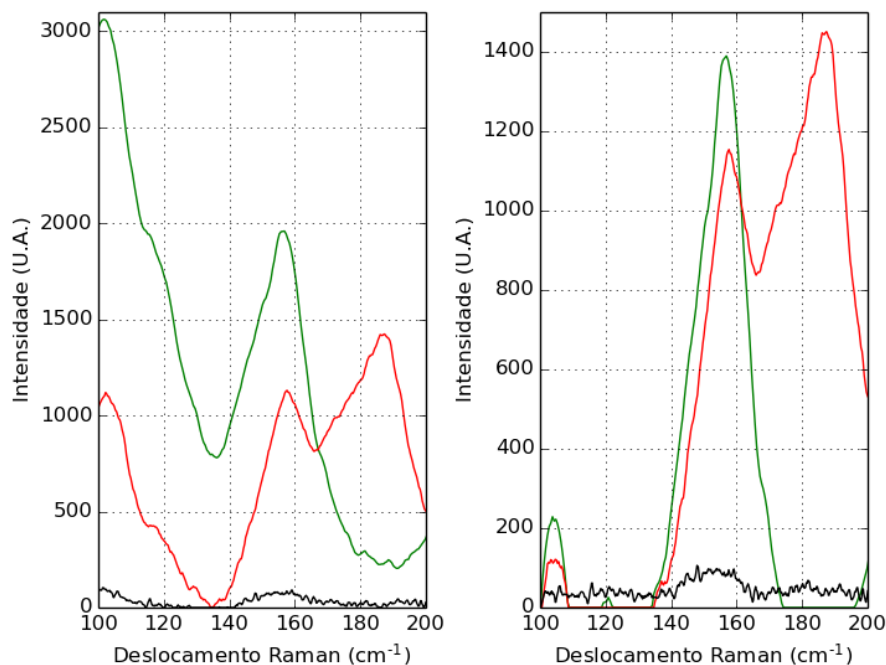


Figura 18 - Espectros Raman da espécie FLUMICRO\_AN05, na região de 100-200  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

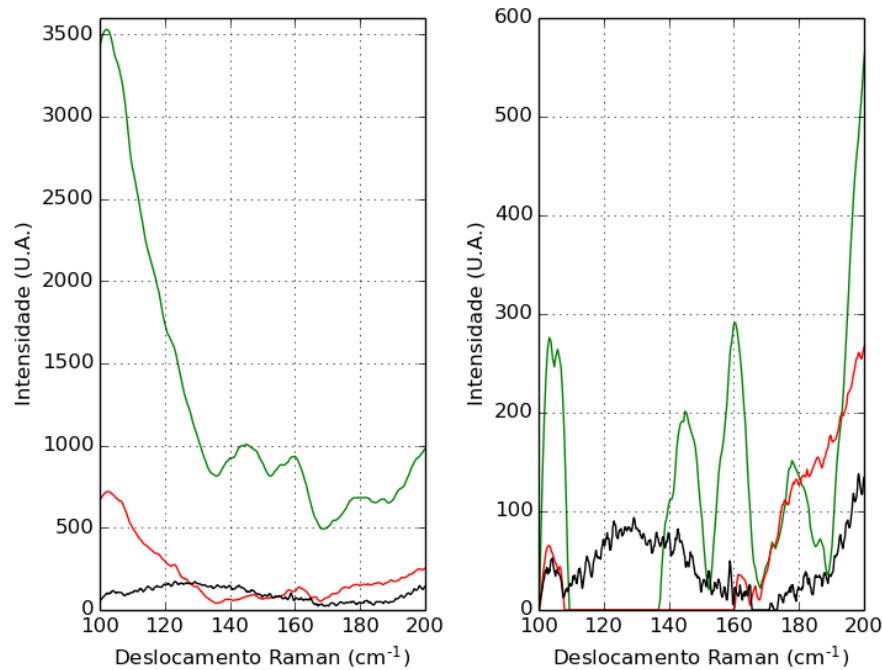


Figura 19 - Espectros Raman da espécie MICR\_AN06, na região de 100-200  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

#### 4.2.2 DEFORMAÇÃO ANGULAR

Nesta região, se tem a deformação angular de O-Nb-O no grupo do pirocloro e de O-Ta-O no grupo da microlita. Ambos os grupos possuem no mínimo duas bandas e os efeitos de fluorescência observados com o laser 785 nm diminuíram e, inclusive, isto ajudou a identificar bandas Raman, Figura 23 e Figura 25. Houve também perdas de bandas pela remoção de ruído de fundo por VP, como é visto nas figuras 20 e 22.

As espécies do grupo do pirocloro apresentam bandas próximas de 280, 370 e 420  $\text{cm}^{-1}$  (figuras 20, 21 e 22). Na amostra PY\_AN3, figura 22, observa-se uma larga banda em torno de 280  $\text{cm}^{-1}$ , o que dificulta a identificação de outras bandas. Isto provavelmente ocorreu pela presença de impurezas na amostra o que agravou o efeito de ruído de fundo na coleta. O mesmo não ocorreu para as outras amostras PY1\_AN1 e PY\_AN02, onde foi possível observar as bandas 280, 370 e 420  $\text{cm}^{-1}$ .

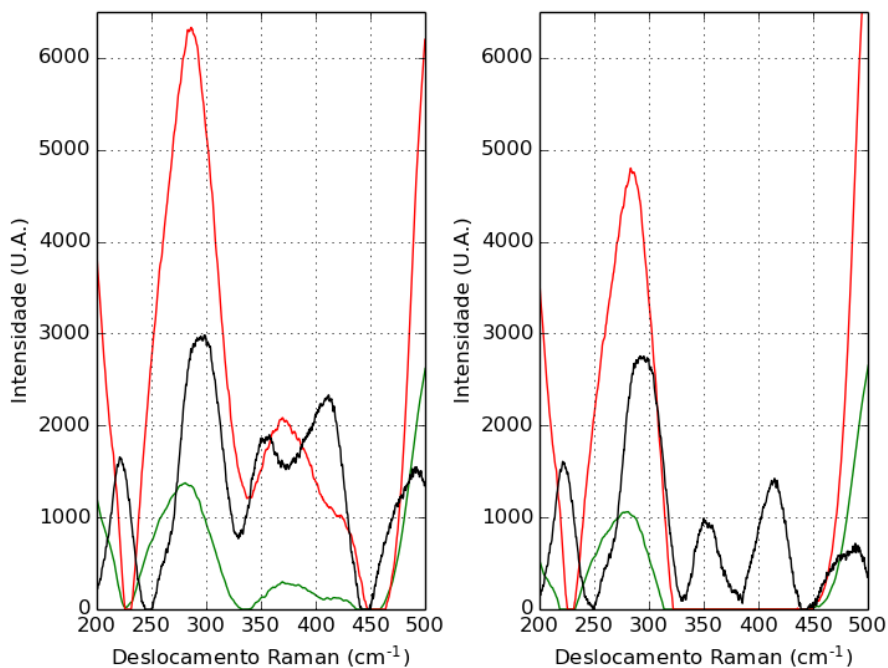


Figura 20 - Espectros Raman da espécie PY1\_AN1, na região de 200-500  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

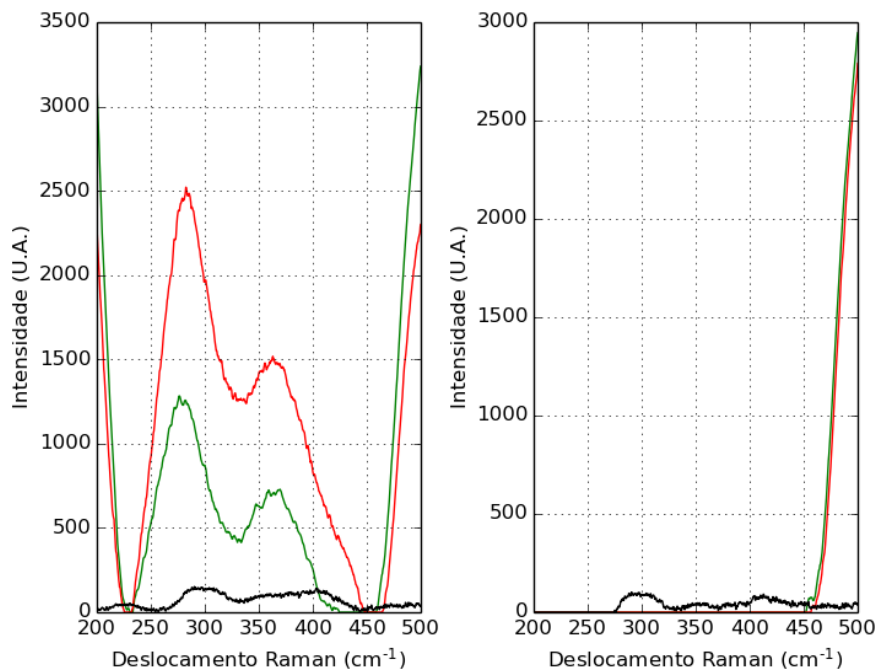


Figura 21 - Espectros Raman da espécie PY\_AN02, na região de 200-500  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

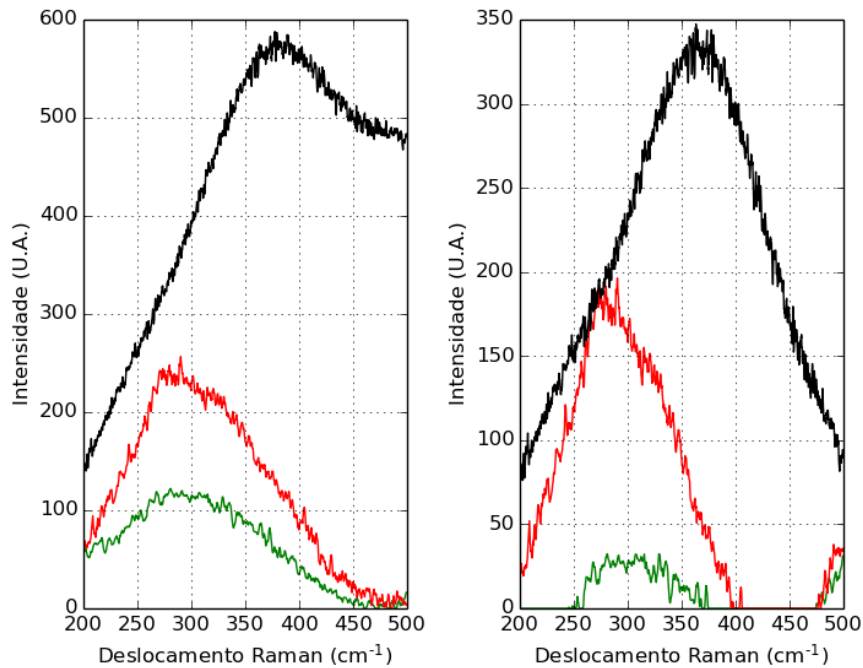


Figura 22 - Espectros Raman da espécie PY\_AN3, na região de 200-500  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

As espécies de microlita apresentam quatro bandas: 250, 300, 340 e 450  $\text{cm}^{-1}$  (figuras 23, 24 e 25). Por exemplo, em 250  $\text{cm}^{-1}$  é possível que tenhamos uma região de estiramento de A-Y combinada com deformação angular X-A-Y.<sup>81</sup>

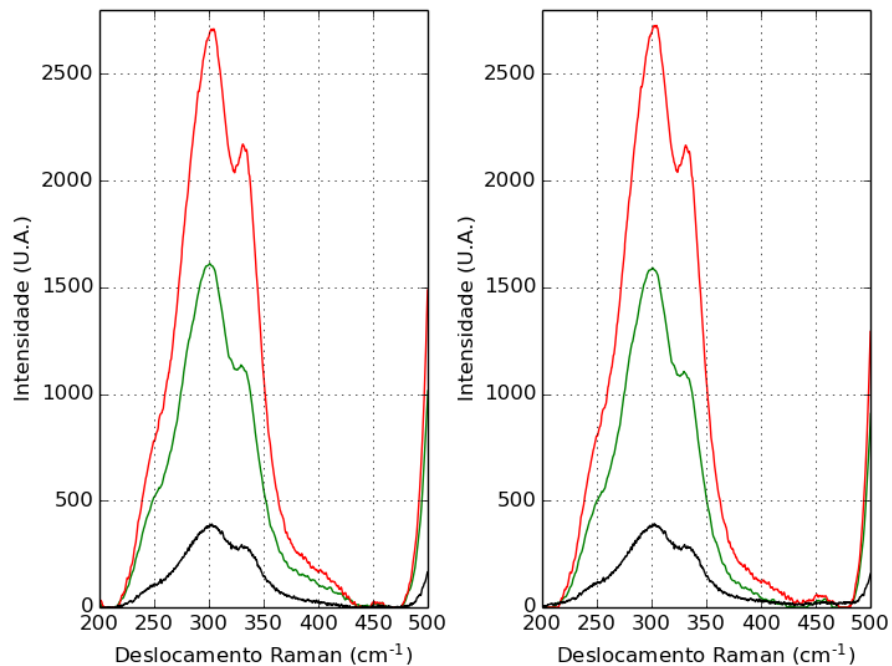


Figura 23 - Espectros Raman da espécie MI\_AN4, na região de 200-500  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

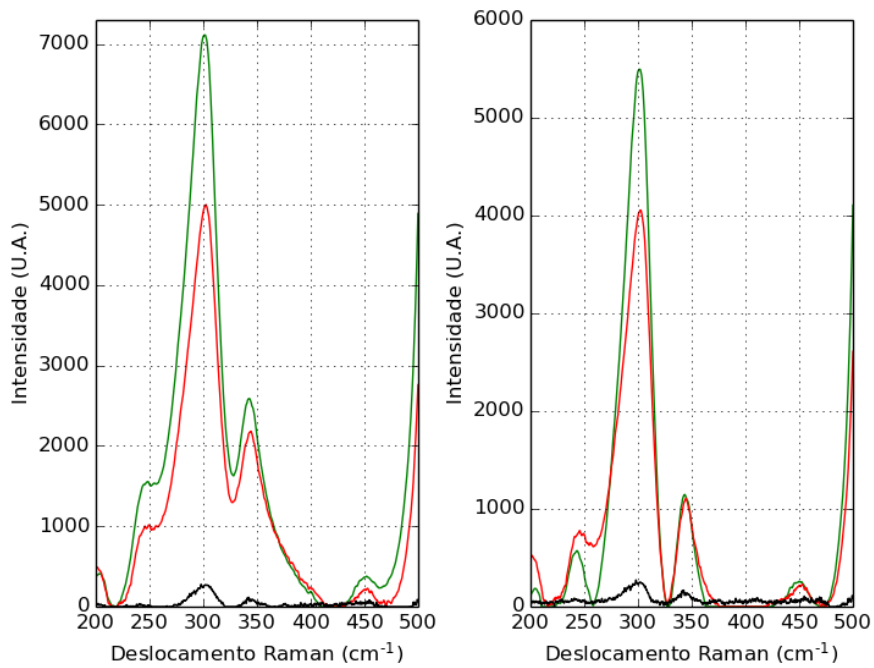


Figura 24 - Espectros Raman da espécie FLUMICRO\_AN05, na região de 200-500  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

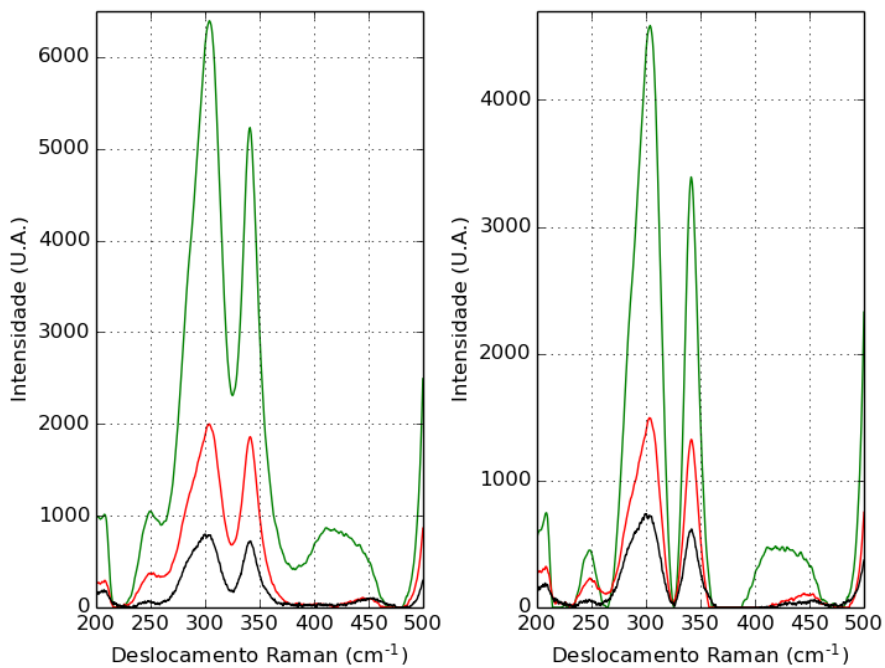


Figura 25 - Espectros Raman da espécie MICR\_AN06, na região de 200-500  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.



### 4.2.3 ESTIRAMENTO SIMÉTRICO E ANTISSIMÉTRICO

O estiramento de Nb-O e Ta-O ocorre na região entre 500 e 800  $\text{cm}^{-1}$ . A maioria das amostras neste trabalho, apresentam duas bandas próximas de: 540 e 630  $\text{cm}^{-1}$  (para PY1\_AN1 e PY\_AN2) e 520 e 630  $\text{cm}^{-1}$  (para MI\_AN4, FLUMICRO\_AN05 e MICR\_AN06), como visto nas figuras 26, 27, 29, 30 e 31. Em PY\_AN3, na figura 28, observa-se apenas uma banda próxima de 630  $\text{cm}^{-1}$  devido ao grande ruído de fundo. A figura 30 da amostra FLUMICRO\_AN05 observa-se uma terceira banda em 660  $\text{cm}^{-1}$ , que também pode ser atribuída ao estiramento.<sup>81</sup>

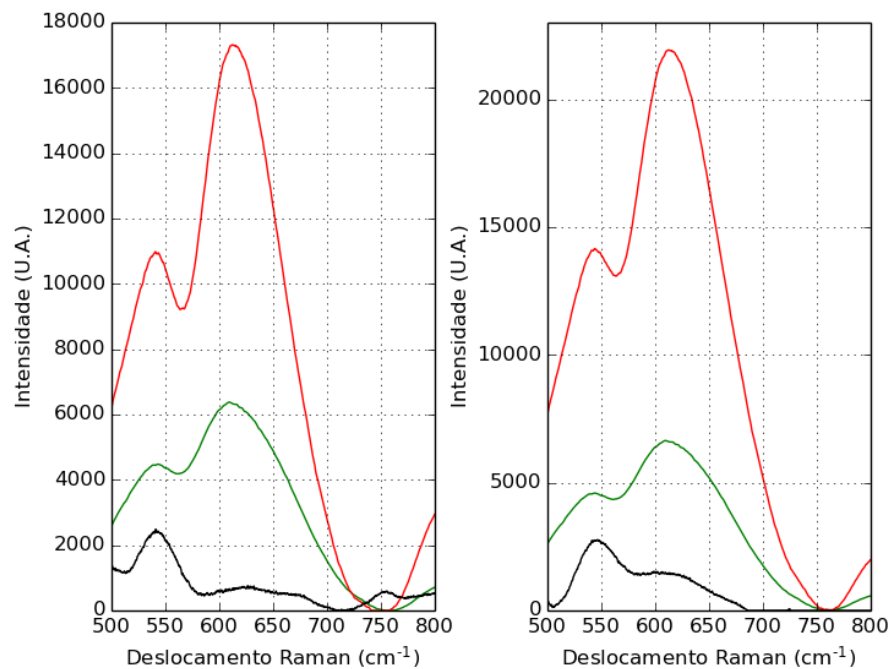


Figura 26- Espectros Raman da espécie PY1\_AN1, na região de 500-800  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

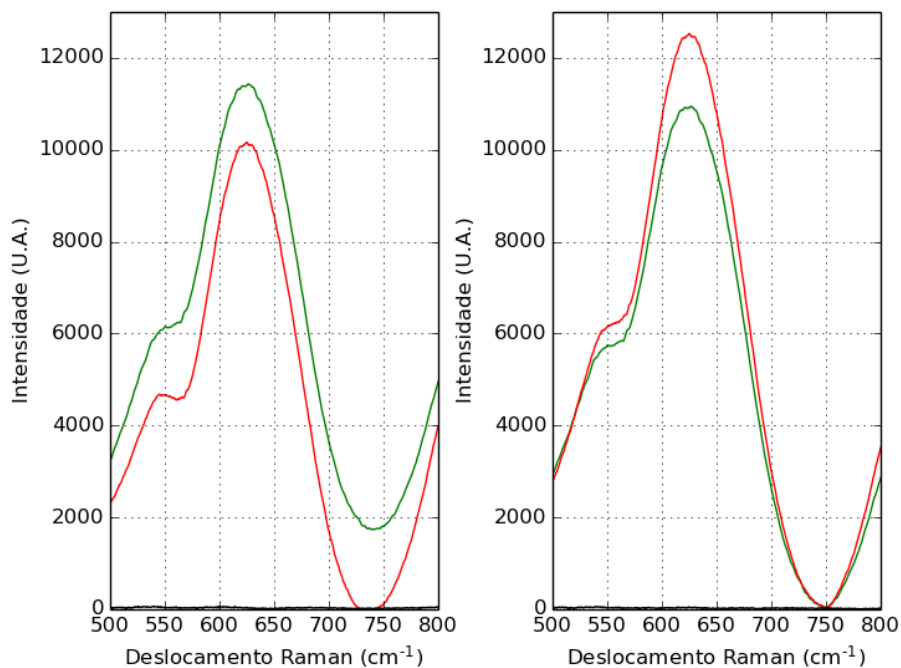


Figura 27- Espectros Raman da espécie PY\_AN02, na região de 500-800  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

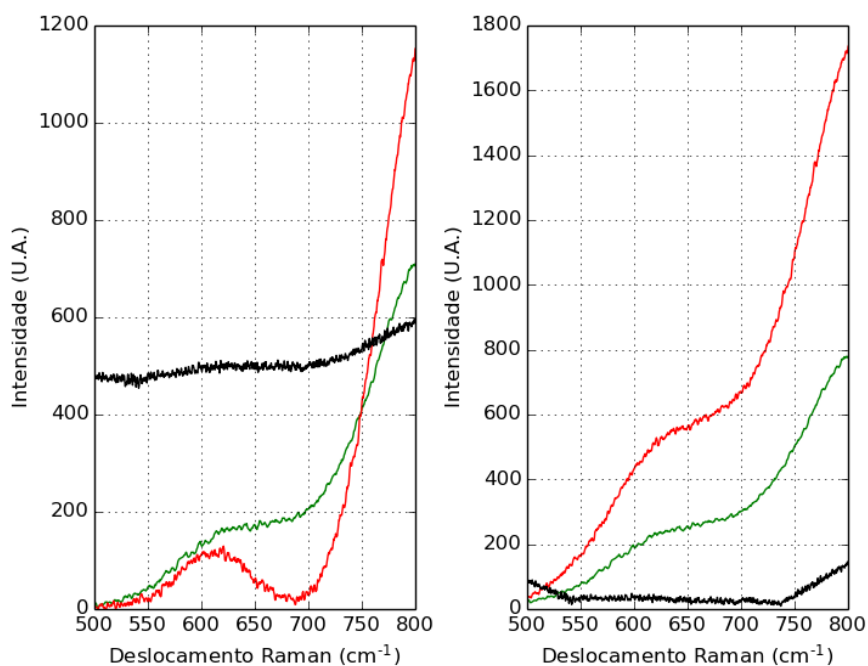


Figura 28- Espectros Raman da espécie PY\_AN3, na região de 500-800  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

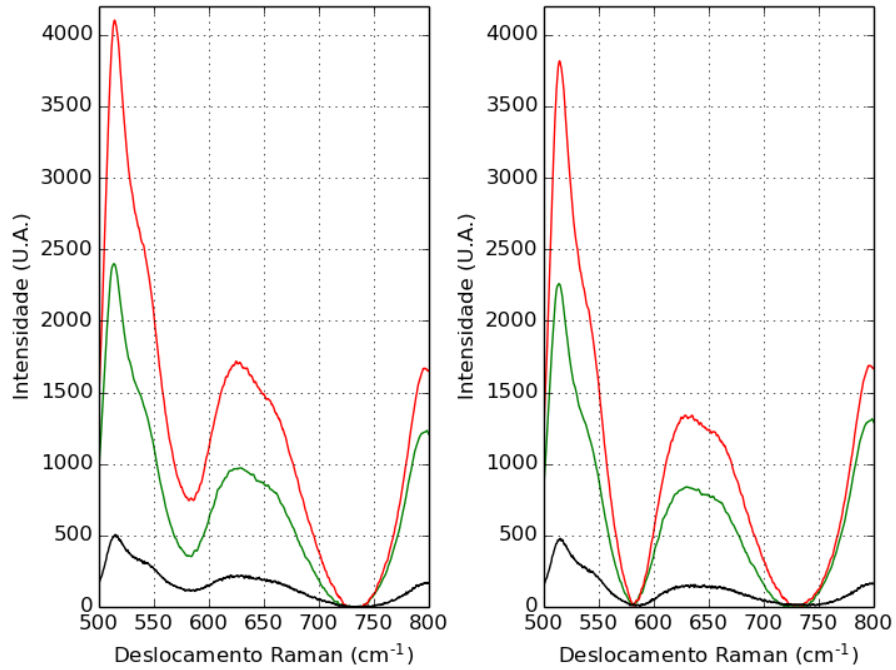


Figura 29- Espectros Raman da espécie MI\_AN4, na região de 500-800  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

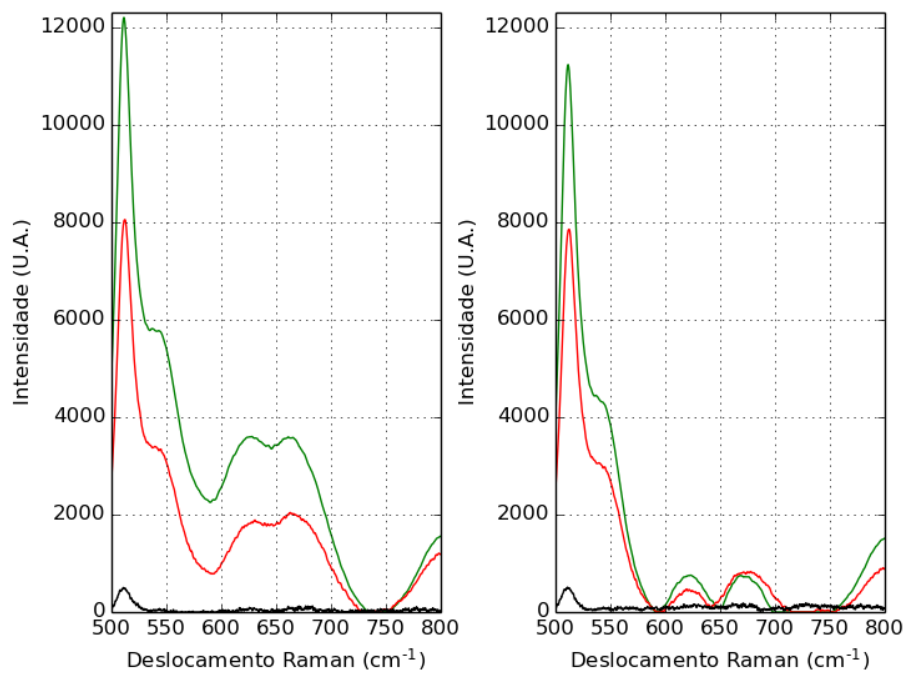


Figura 30- Espectros Raman da espécie FLUMICRO\_AN05, na região de 500-800  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

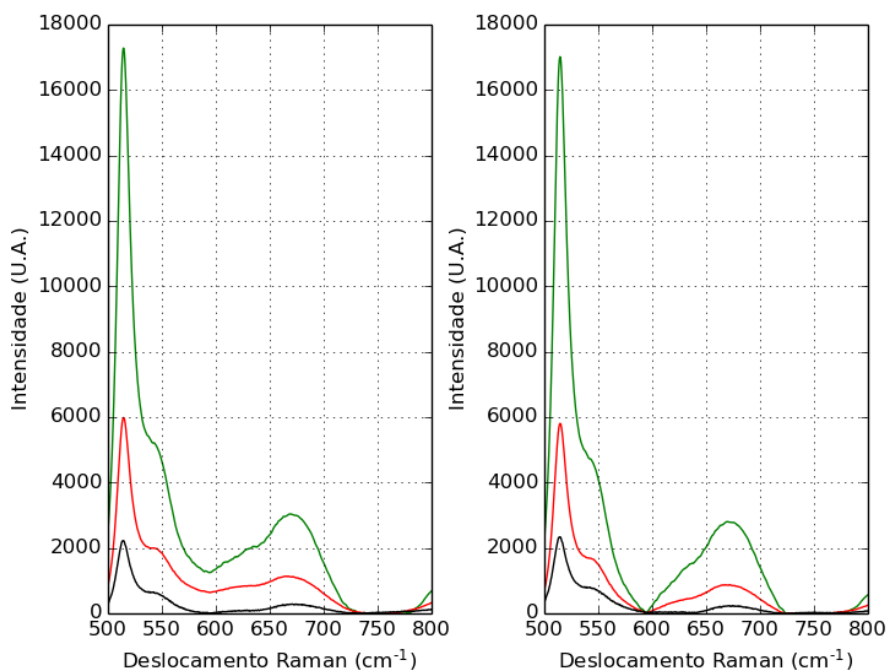


Figura 31- Espectros Raman da espécie MICR\_AN06, na região de 500-800  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

#### 4.2.4 SOBRETONS E COMBINAÇÕES DE BANDAS

Na região de 800 até 1400  $\text{cm}^{-1}$  é observado bandas que podem ser sobretons (“*overtones*”) ou combinações de bandas. Mas há casos, como as bandas em 810  $\text{cm}^{-1}$ ,<sup>83</sup> que podem ser também classificadas como de estiramento. Nesta região é difícil ter a determinação de qual tipo de modo vibracional ocorre, até mesmo para cálculos de predições teóricas de bandas Raman na literatura.<sup>88</sup> Nas figuras abaixo se tem mais exemplos de perda de informação com a remoção de ruído por VP, além de efeitos de fluorescência com o lasers 785 nm.

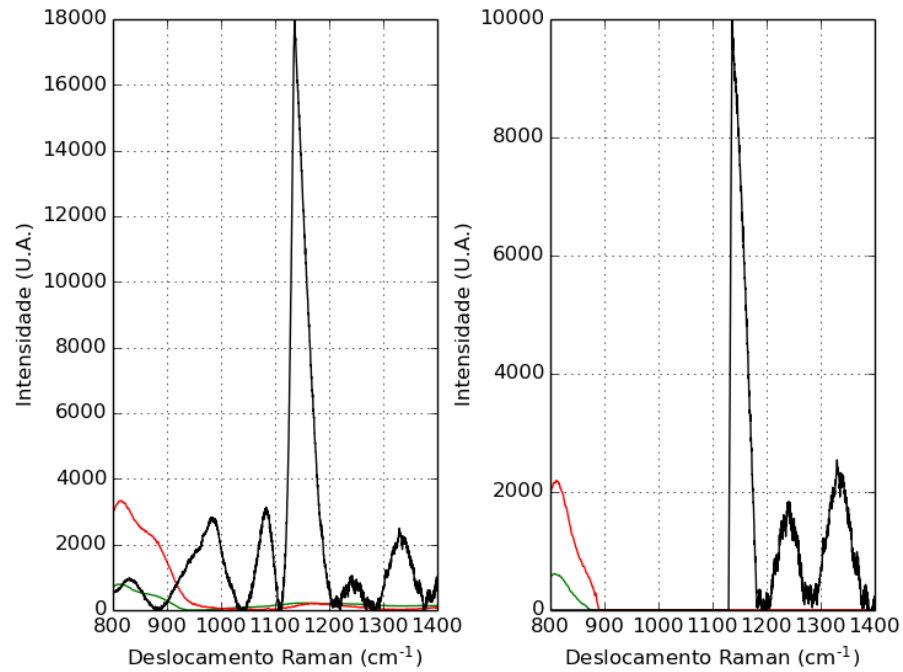


Figura 32 - Espectros Raman da espécie PY1\_AN1, na região de 800-1400  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

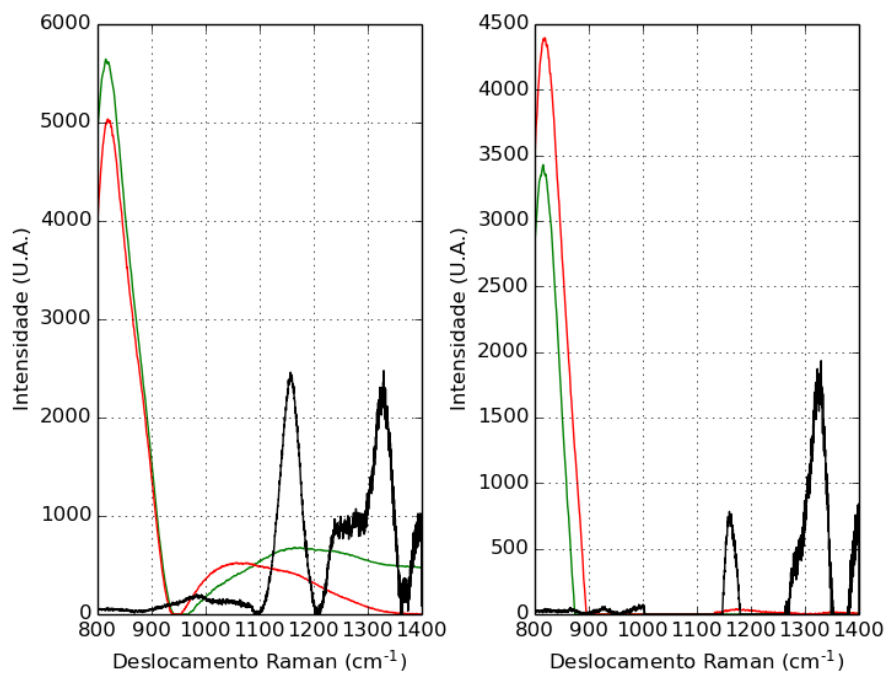


Figura 33 - Espectros Raman da espécie PY\_AN02, na região de 800-1400  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

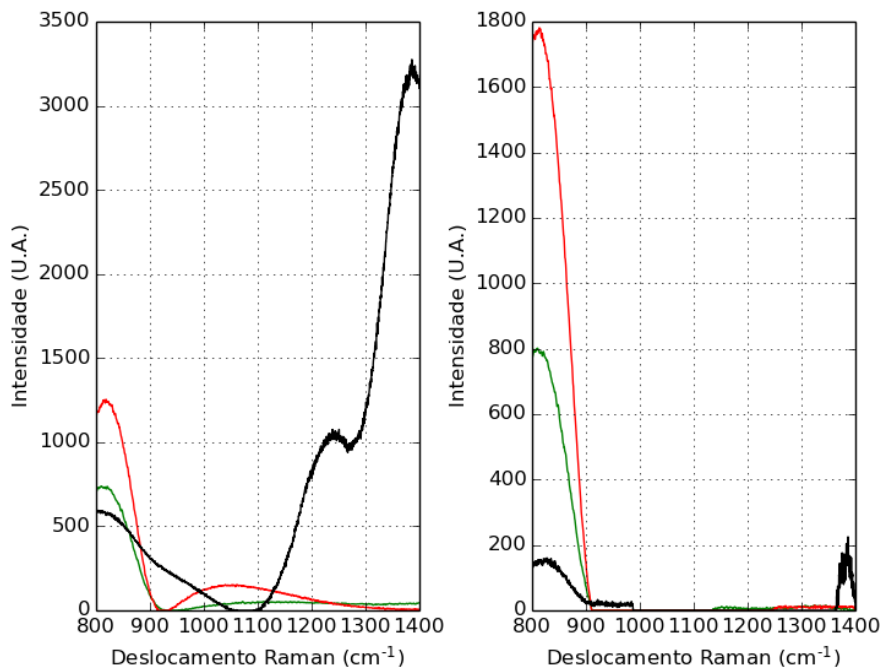


Figura 34 - Espectros Raman da espécie PY\_AN3, na região de 800-1400  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

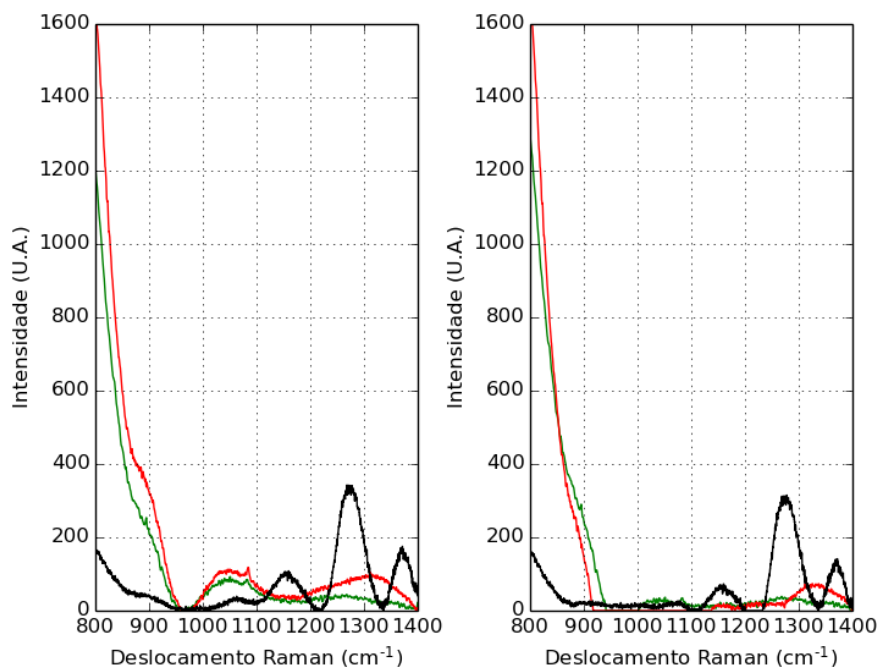


Figura 35 - Espectros Raman da espécie MI\_AN4, na região de 800-1400  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

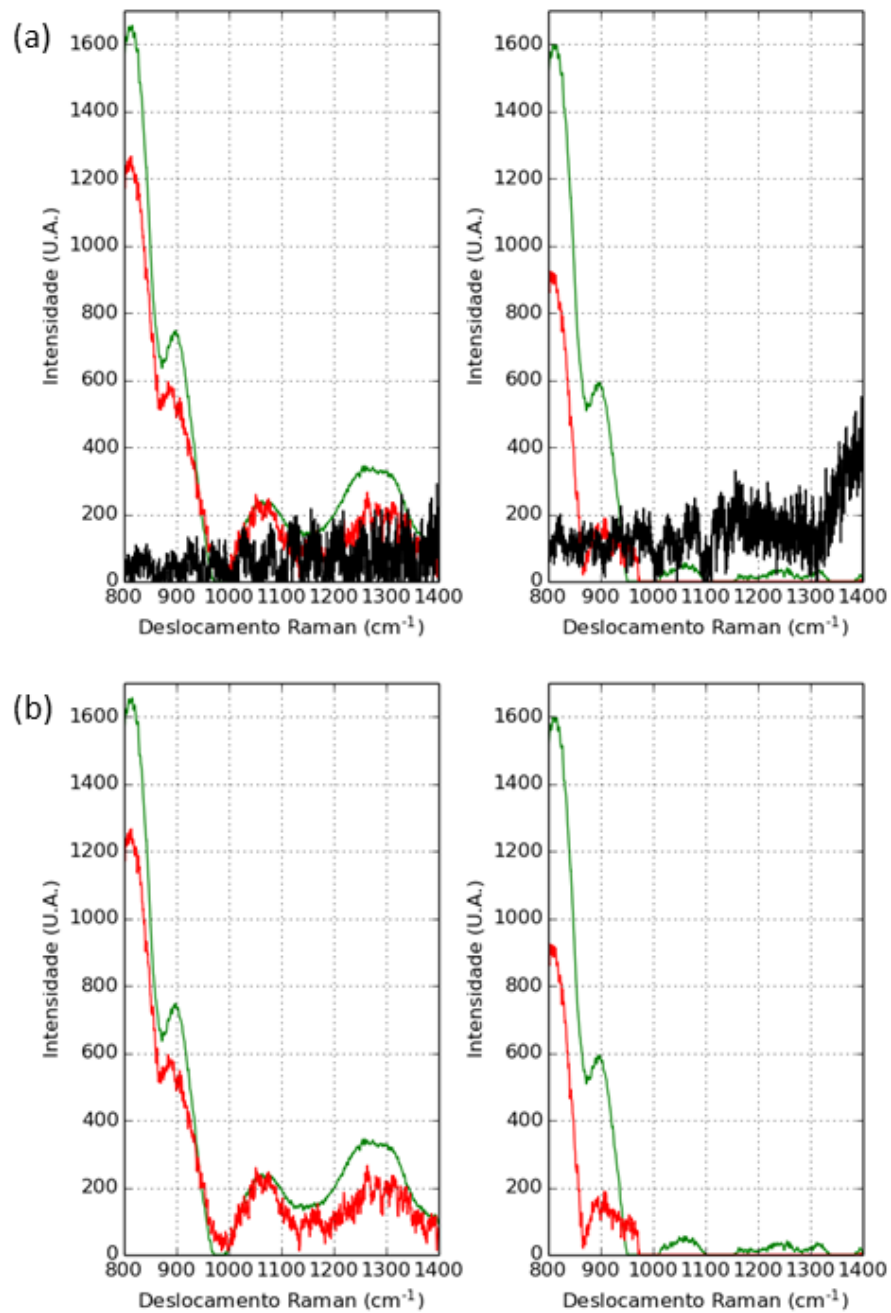


Figura 36 - Espectros Raman da espécie FLUMICRO\_AN05, na região de 800-1400 cm<sup>-1</sup>, após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP. (a) Utilizando os três lasers. (b) Sem a medida com o laser 785 nm.

Fonte: Elaborada pelo autor.

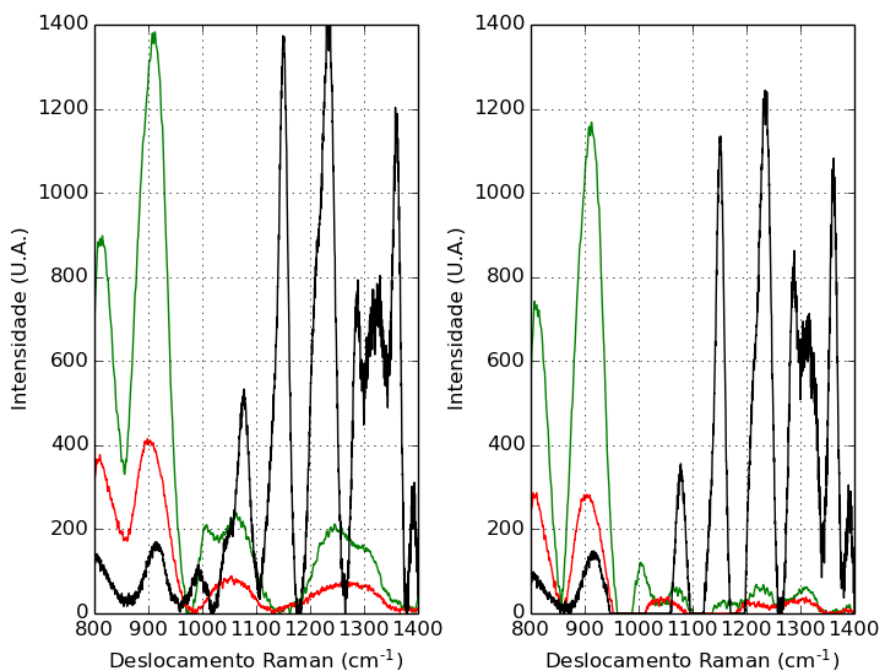


Figura 37 - Espectros Raman da espécie MICR\_AN06, na região de 800-1400  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

#### 4.2.5 DEFORMAÇÃO ANGULAR DE $\text{H}_2\text{O}$

Na região em torno de  $1600 \text{ cm}^{-1}$ , se procura a ocorrência de bandas de deformação angular de H-O-H. Nas amostras utilizadas neste trabalho apenas o espectro correspondente a MI\_AN4 apresentou bandas nesta região, como pode ser observado na figura 38, isto condiz com o resultado da composição química e com a literatura.<sup>81</sup>



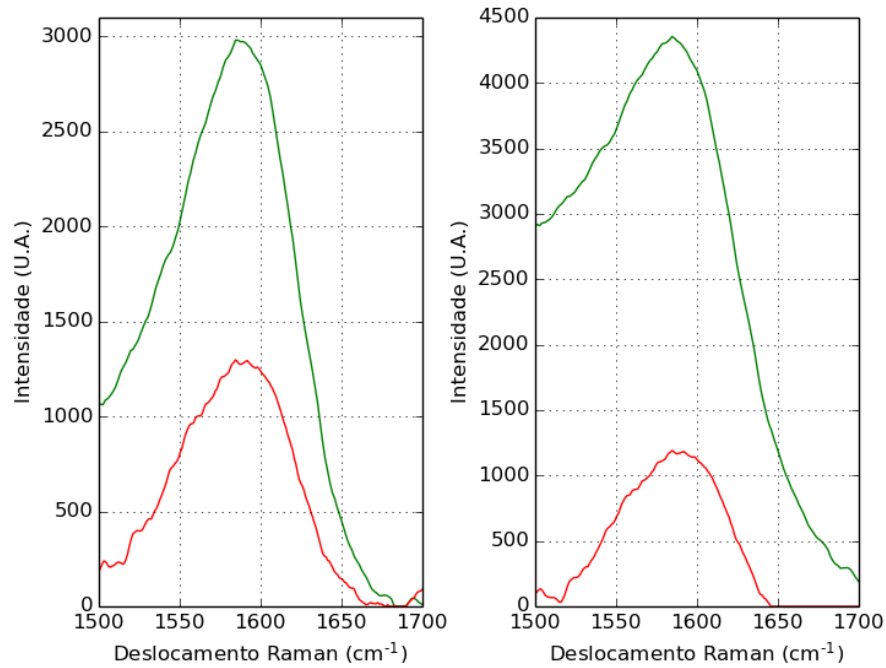


Figura 38 - Espectros Raman da espécie MI\_AN4, na região de 1400-1800  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

#### 4.2.6 ESTIRAMENTO DAS LIGAÇÕES O-H DA MOLÉCULA DE ÁGUA

As espécies do grupo do pirocloro apresentam bandas nesta região, como visto nas figuras 39, 40 e 41. A espécie MI\_AN4 apresenta uma banda em  $3600 \text{ cm}^{-1}$  (figura 42) que indica a presença de OH na estrutura. Para FLUMICRO\_AN05, não se observou bandas, como previsto na composição química.

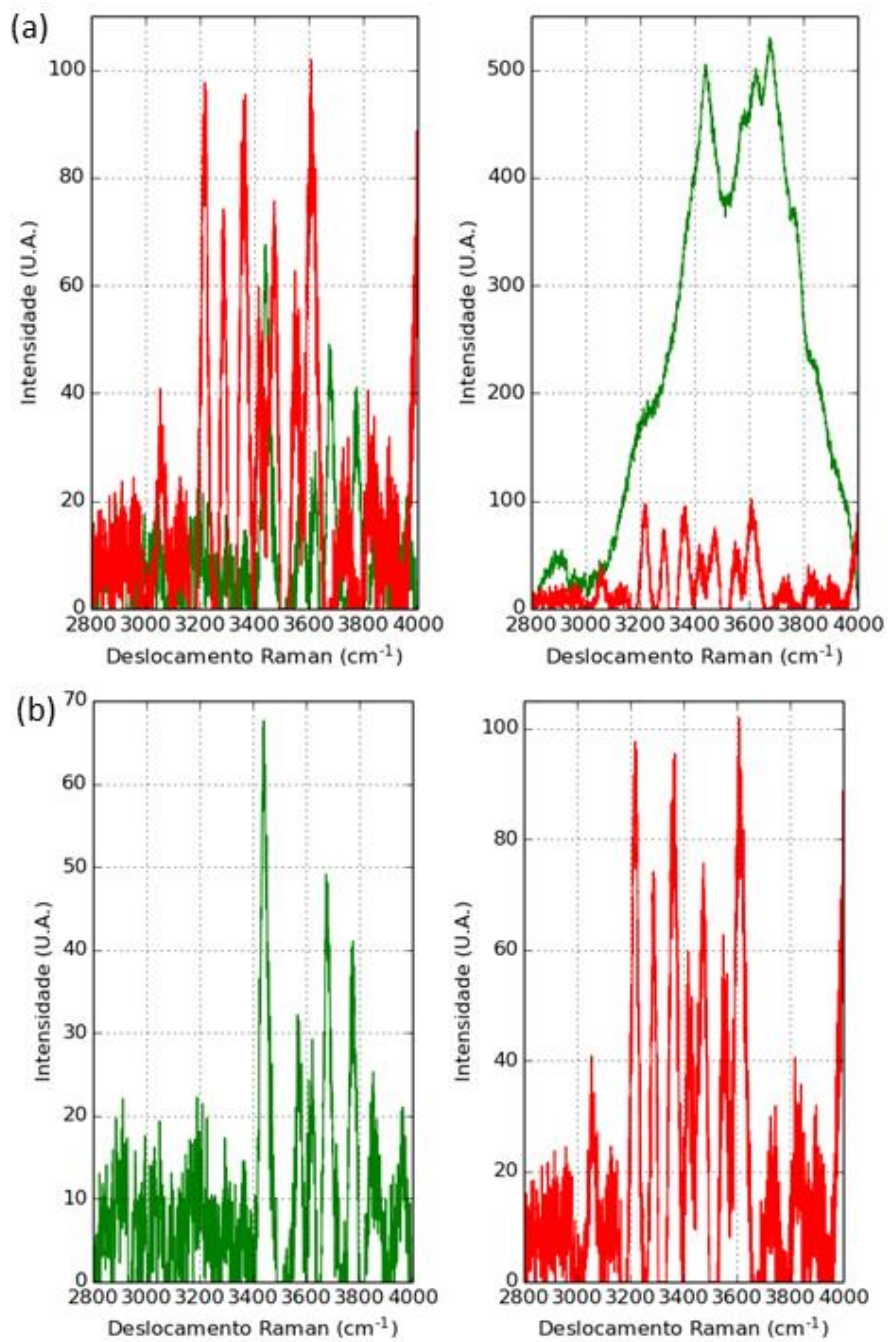


Figura 39 - Espectros Raman da espécie PY1\_AN1, na região de 2800-4000 cm<sup>-1</sup>, após a correção das linhas de base implementadas. Em (a) Esquerda por ALS e direita por VP. Em (b), espectros corrigidos por ALS utilizando os lasers 532 nm (em verde) e 633 nm (em vermelho).

Fonte: Elaborada pelo autor.

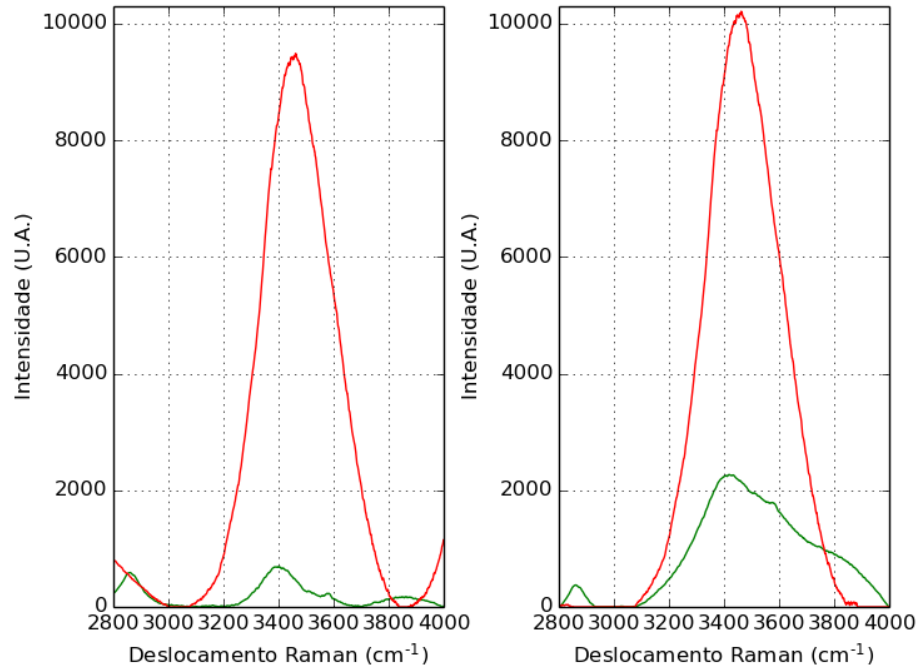


Figura 40 - Espectros Raman da espécie PY\_AN02, na região de 2800-4000  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

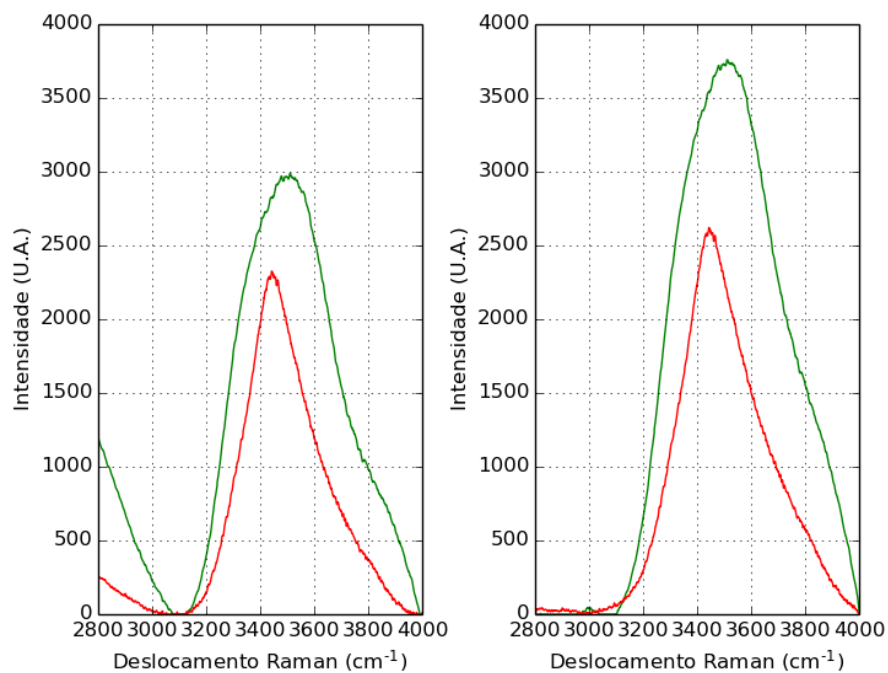


Figura 41 - Espectros Raman da espécie PY\_AN3, na região de 2800-4000  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

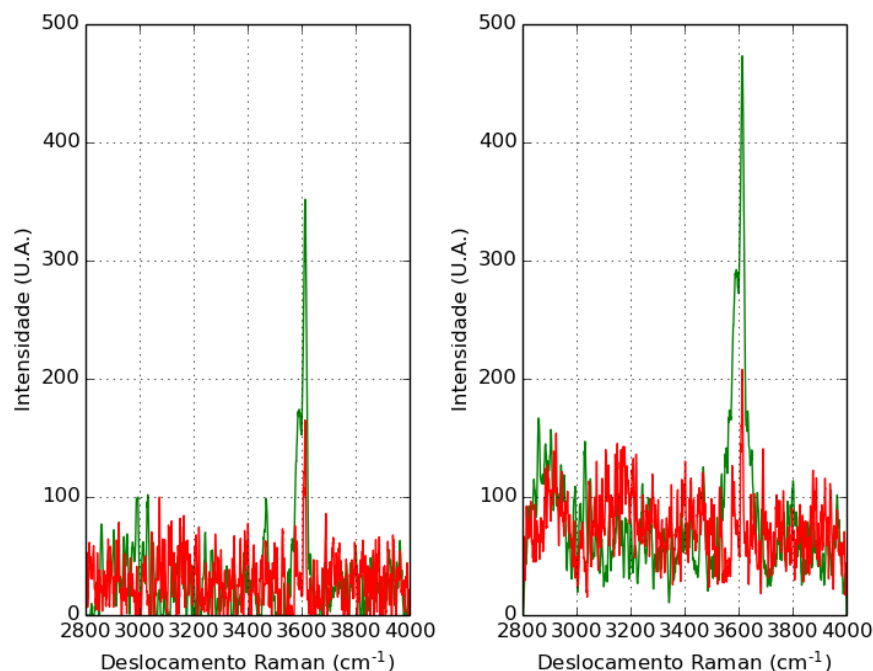


Figura 42 - Espectros Raman da espécie MI\_AN4, na região de 2800-4000  $\text{cm}^{-1}$ , após a correção das linhas de base implementadas. Esquerda por ALS e direita por VP.

Fonte: Elaborada pelo autor.

### 4.3 APRENDIZADO DA REDE

#### 4.3.1 TREINAMENTO E TESTES

A quantidade de espectros nos conjuntos de treinamento e de teste é representada pela Tabela 3. A região de 100-1400  $\text{cm}^{-1}$  possui mais espectros, em comparação com as outras regiões, devido ao fato de ser a região na qual se encontra as principais bandas características dos minerais (impressão digital) e, assim, estar mais presente na literatura. Como exemplo, até o presente momento a base de dados da Universidade de Parma<sup>79</sup> e do Laboratório de Geologia de Lyon<sup>80</sup> não foram disponibilizados espectros em outras regiões fora da região de “impressão digital”. O projeto RRUFF<sup>19</sup> possui os espectros Raman em alta resolução nesta faixa espectral e espectros de 200-4000  $\text{cm}^{-1}$ , mas em baixa resolução.

Para algumas espécies de minerais há poucas espectros Raman disponíveis para comparação. Em alguns casos como o do fluornatropirocloro, não há espectros presentes na base de dados do projeto RRUFF.

Tabela 3 - Número de espectros para cada classe em suas respectivas regiões.

Conjunto	Região (cm <sup>-1</sup> )	Grupos				
		pirocloro		microlita		outros
		Classes - espécies				
		0- fluornatro pirocloro	1- fluorcalcio pirocloro	2- fluorcalcio microlita	3- fluornatro microlita	4-outros
Treinamento	100- 1400	3	20	17	8	13
	1200- 1800	6	5	8	4	12
	2800- 4000	3	5	9	4	12
Teste	100- 1400	0	7	5	2	4
	1200- 1800	0	4	3	1	4
	2800- 4000	0	4	3	1	4

Fonte: Elaborada pelo autor

Durante a etapa de treinamento, em conjunto com a de testes, variou-se o número de camadas ocultas e de neurônios para encontrar a topologia com: maior acurácia, menor valor de função custo, para os dados do conjunto de testes, e menor complexidade (número de camadas ocultas e neurônios), como apresentado nas figuras 43-48.

Na procura da melhor topologia, variou-se primeiramente o número de camadas ocultas sem alterar a quantidade de neurônios em cada uma dessas camadas (mantida constante e igual a dez), como ilustrado nas figuras 43-45. Após essa etapa, alterou-se o número de neurônios pertencente a última camada oculta, mantendo constante o número de neurônios, dez, nas camadas ocultas anteriores, como ilustrado nas figuras 46-48 para o caso de duas camadas ocultas.

As redes foram treinadas com 1000 épocas, o valor da taxa de aprendizagem fixo em 1 e variando as quantidades de camadas ocultas. O tempo gasto para executar o treinamento de 10 vezes (para sair de casos de mínimos locais), com as três funções de ativação e salvamento das saídas em arquivos texto e imagem (os valores de acurácia, função custo e

as probabilidades de serem cada classe) foi entre 6 minutos, para as redes com menor número de camadas, até 22 minutos. Neste caso, foi observado que as estruturas com duas camadas ocultas tiveram maior acurácia, menor função custo e menor complexidade para as três regiões espectrais.

Como duas camadas tiveram o melhor desempenho, alterou-se o número de neurônios na última camada oculta. O tempo gasto, com iguais condições anteriores, foi de aproximadamente 6 minutos. Na maioria das regiões, figuras 46 e 48, a última camada oculta com 10 neurônios obteve melhores resultados. Apenas para a região de 1200-1800  $\text{cm}^{-1}$ , com 8 neurônios, foi melhor ao aplicar para ambas as remoções de ruído, presente na figura 47.

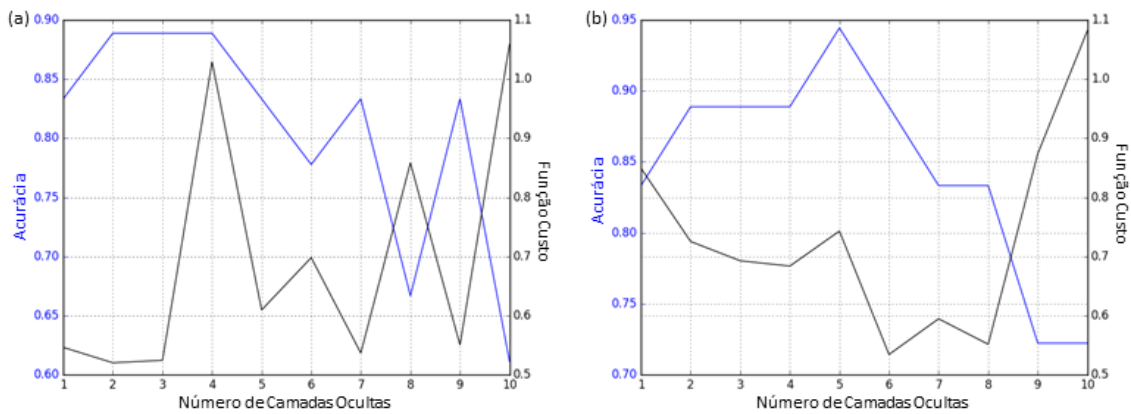


Figura 43 - Acurácia e função custo em função da variação do número de camadas ocultas para o conjunto da região de 100 – 1400  $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b).

Fonte: Elaborada pelo autor.

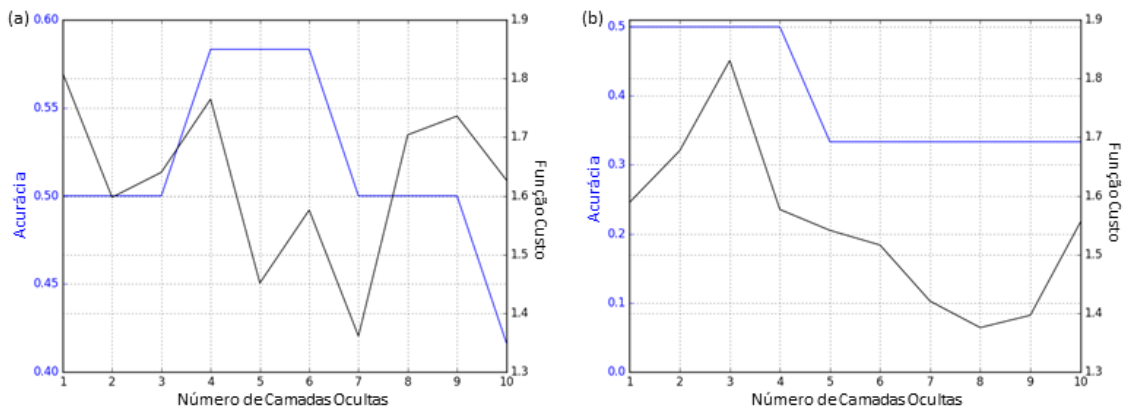


Figura 44- Acurácia e função custo em função da variação do número de camadas ocultas para o conjunto da região de 1200 – 1800  $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b).

Fonte: Elaborada pelo autor.

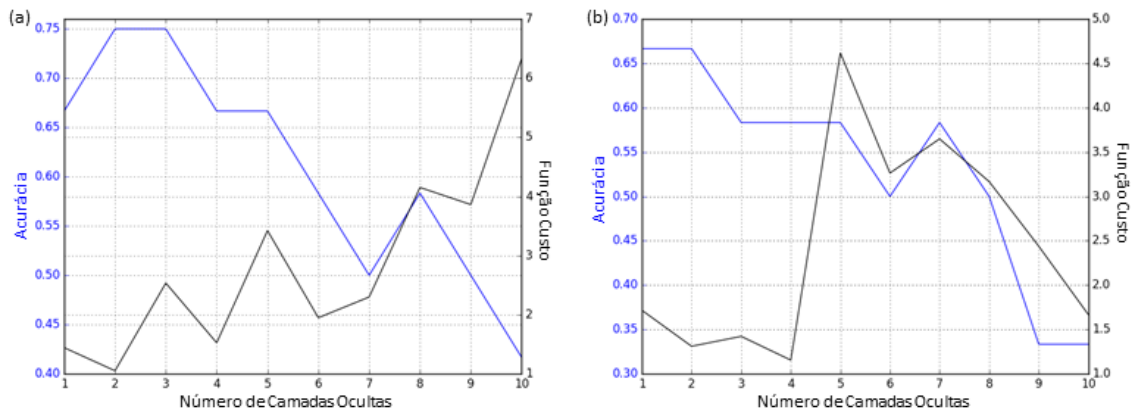


Figura 45 - Acurácia e função custo em função da variação do número de camadas ocultas para o conjunto da região de  $2800 - 4000 \text{ cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b).

Fonte: Elaborada pelo autor.

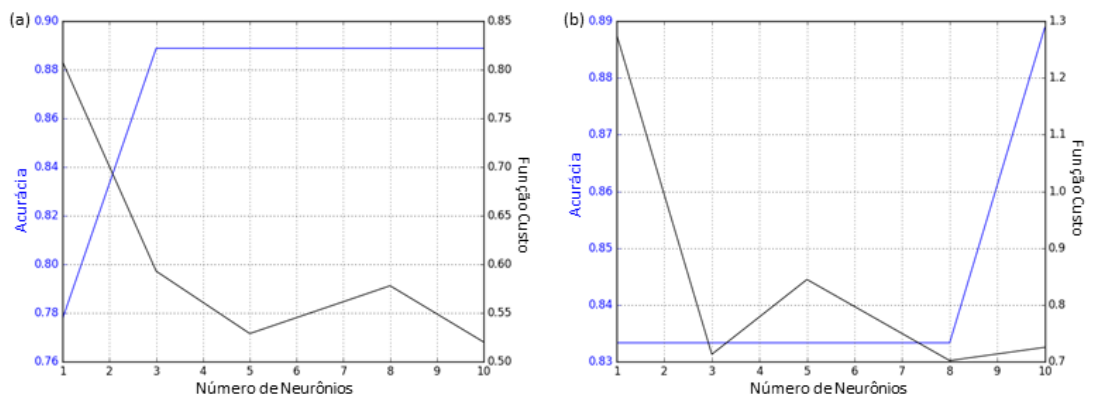


Figura 46 - Acurácia e função custo em função da variação do número de neurônios na última camada oculta para o conjunto da região de  $100 - 1400 \text{ cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b).

Fonte: Elaborada pelo autor.

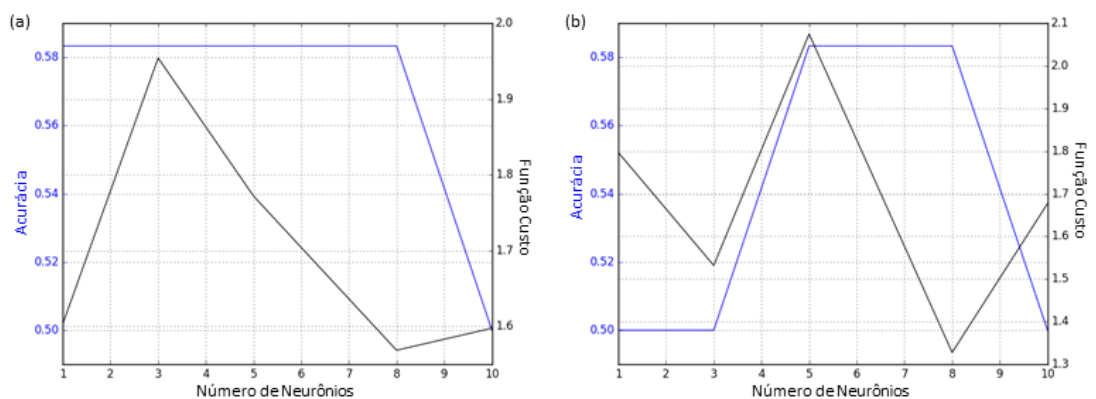


Figura 47 - Acurácia e função custo em função da variação do número de neurônios na última camada oculta para o conjunto da região de  $1200 - 1800 \text{ cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b).

Fonte: Elaborada pelo autor.

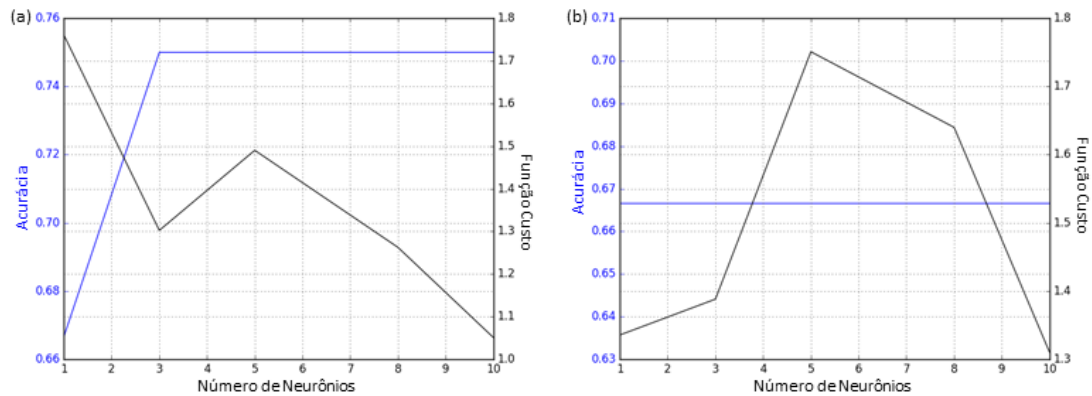


Figura 48 - Acurácia e função custo em função da variação do número de neurônios na última camada oculta para o conjunto da região de 2800 – 4000  $\text{cm}^{-1}$ . Conjunto de dados com remoção de ruído por: ALS (a) e VP (b).

Fonte: Elaborada pelo autor.

Com relação as funções de ativação, a ReLU apresentou a menor acurácia em comparação com as outras duas funções (logística e tangente hiperbólica – tanh). Utilizando como exemplo o melhor caso para a rede com duas camadas ocultas (10 neurônios em cada) na região de 100-1400  $\text{cm}^{-1}$  e com o conjunto de dados processados com ALS implementado. Nesta estrutura, a rede com a função ReLU obteve 0.777 de acurácia e 0.632 de valor da função custo, enquanto as outras funções obtiveram (0.8333, 0.737) para a função logística e (0.888, 0.520) para a tangente hiperbólica.

Por último, alterou-se a taxa de aprendizagem (TA) de 0.01 até 1 observando como a acurácia e a função custo são afetadas. A taxa de aprendizagem está relacionada diretamente com a velocidade de aprendizagem, quanto menor for esta variável maior será o número de épocas e o tempo de processamento. Mas se a taxa for grande, haverá oscilações na aprendizagem, como ilustrado na Figura 49. Para executar essa parte, foi necessário aumentar o número de épocas para 10000 e repetir 50 vezes. O tempo gasto foi de 300 minutos, o que corresponde ao o tempo gasto anteriormente para encontrar a melhor topologia (10 vezes mais épocas e repetido 5 vezes mais).



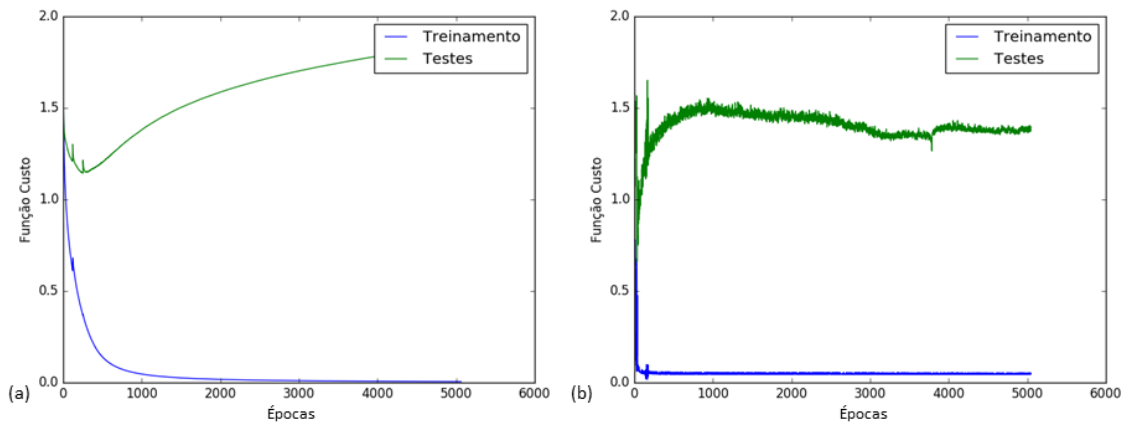


Figura 49 - Variação da função custo em relação à taxa de aprendizagem (TA): (a) TA = 0.01 e (b) TA = 1

Fonte: Elaborada pelo autor.

Na implementação das topologias, foi necessário à utilização da parada antecipada (“early stopping”) para diminuir os efeitos de sobreajuste (“Overfitting”), uma vez que os resultados de treinamento se distanciam enormemente dos resultados de testes, como apresentado nas figuras 49 e 50. A parada ocorre após se passar 500 ou 5000 épocas (dependendo da taxa de aprendizagem) depois de obter a época com maior acurácia para o conjunto de testes.

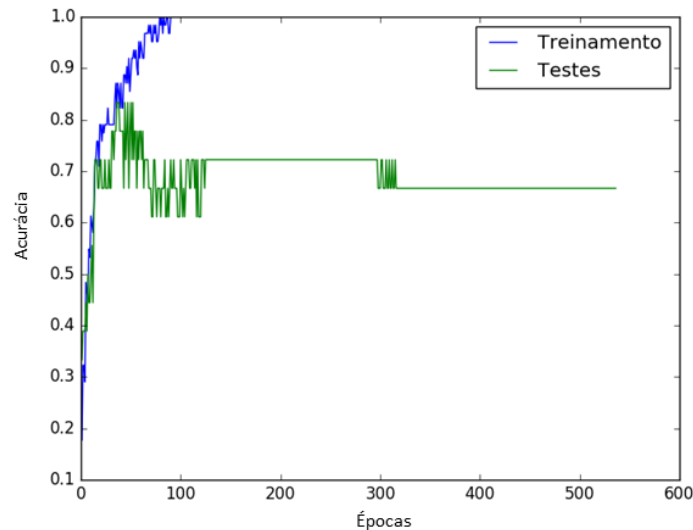


Figura 50 - Acurácia em relação ao número de épocas, indicação presença de sobreajuste

Fonte: Elaborada pelo autor.

A Tabela 4 mostra, de forma resumida, quais foram os melhores parâmetros para cada região. As figuras 51-56 apresentam exemplos das saídas de cada rede para as amostras do conjunto de testes com a respectiva classe alvo em vermelho. É importante notar que, como visto na seção 4.2, não foram observadas bandas Raman que identifiquem vibrações de sódio (Na) com as de cálcio (Ca), com isso não é possível diferenciar “fluornatro” de “fluorcalcio”. Assim, com a espectroscopia Raman é possível identificar

em qual grupo de minerais (neste trabalho: pirocloro ou microlita) uma determinada amostra pertence.

Tabela 4 - Parâmetros de rede para os melhores casos obtidos

Região (cm <sup>-1</sup> )	Remoção de ruído de fundo	Número de Camadas Ocultas	Número de Neurônios na última Camada Oculta	Taxa de Aprendizagem	Função de Ativação	Acurácia	Valor da função custo
100-1400	ALS	2	10	1	tanh	0.888	0.520
1200-1800			8	0.01		0.583	1.568
2800-4000			10	0.1		0.583	1.256
100-1400			10	1		0.888	0.725
1200-1800	VP	2	8	0.1	logística	0.583	1.327
2800-4000			10	0.01	tanh	0.666	1.310

Fonte: Elaborada pelo autor.

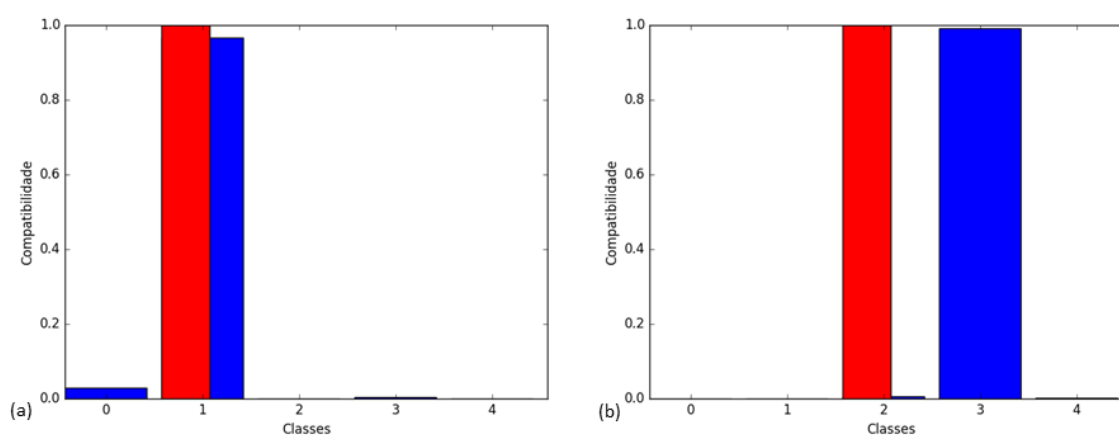


Figura 51 – Exemplos dos resultados da rede com conjunto de testes na região de 100-1400 cm<sup>-1</sup>, processado por ALS. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada e (b) caso de erro de identificação da classe, com acerto do grupo da microlita.

Fonte: Elaborada pelo autor.

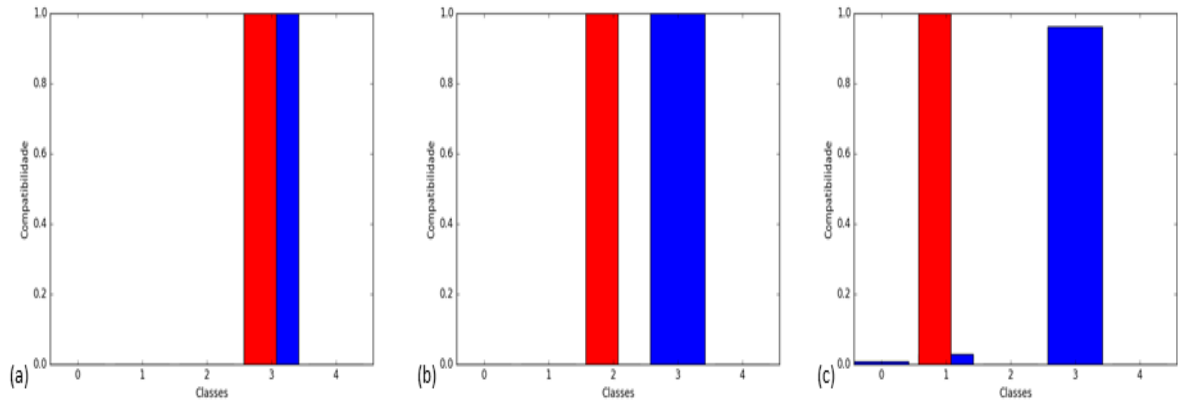


Figura 52 - Exemplos dos resultados da rede com conjunto de testes na região de 100-1400  $\text{cm}^{-1}$ , processado por VP. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo da microlita e (c) caso de erro na identificação de espécie e grupo.

Fonte: Elaborada pelo autor.

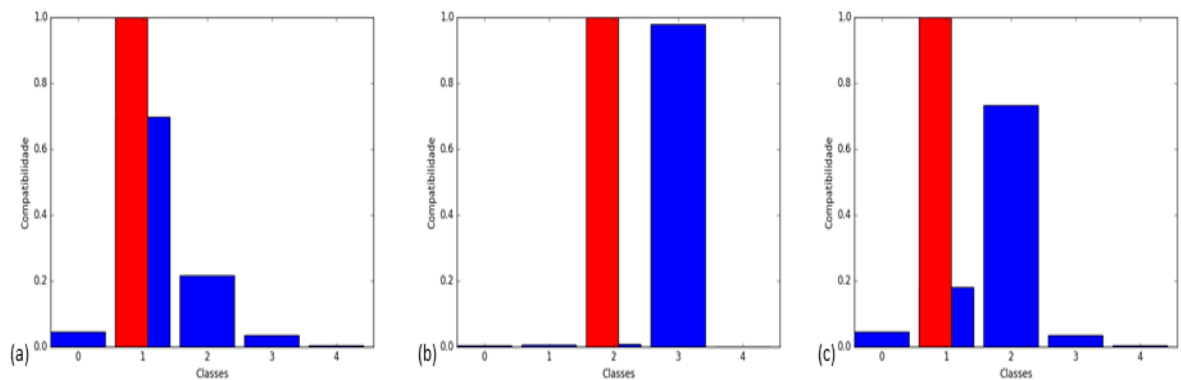


Figura 53 - Exemplos dos resultados da rede com conjunto de testes na região de 1200-1800  $\text{cm}^{-1}$ , processado por ALS. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo da microlita e (c) caso de erro na identificação de espécie e grupo.

Fonte: Elaborada pelo autor.

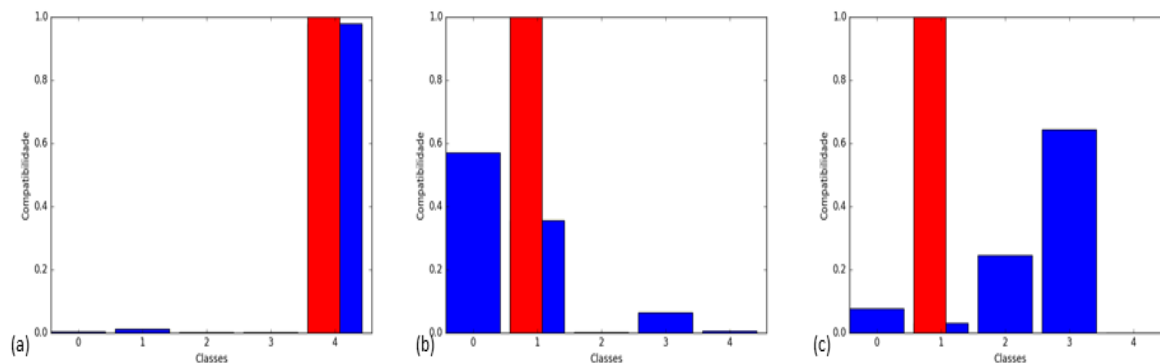


Figura 54 - Exemplos dos resultados da rede com conjunto de testes na região de 1200-1800  $\text{cm}^{-1}$ , processado por VP. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo do pirocloro e (c) caso de erro na identificação de espécie e grupo.

Fonte: Elaborada pelo autor.

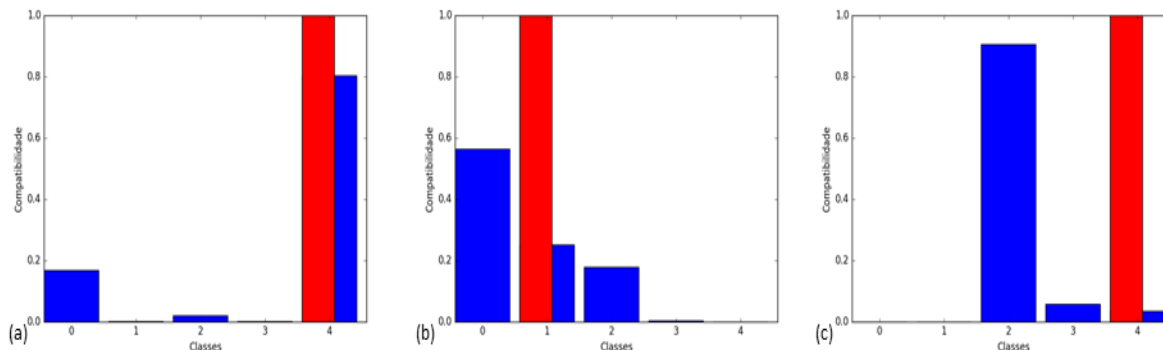


Figura 55 - Exemplos dos resultados da rede com conjunto de testes na região de 2800-4000  $\text{cm}^{-1}$ , processado por ALS. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo do pirocloro e (c) caso de erro na identificação de espécie e grupo.

Fonte: Elaborada pelo autor.

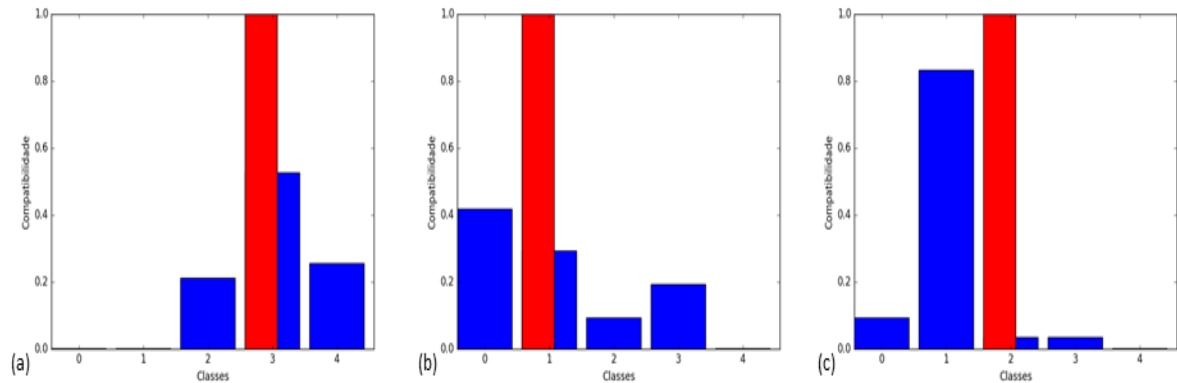


Figura 56 - Exemplos dos resultados da rede com conjunto de testes na região de 2800-4000  $\text{cm}^{-1}$ , processado por VP. Representada em azul a probabilidade da amostra ser de cada classe e em vermelho a resposta desejada. (a) caso de acerto de saída da rede com a resposta desejada, (b) caso de erro de identificação da classe, com acerto do grupo do pirocloro e (c) caso de erro na identificação de espécie e grupo.

Fonte: Elaborada pelo autor.

### 4.3.2 VALIDAÇÃO

O conjunto de validação contém as amostras nas quais se tem apenas os espectros Raman medidos e os possível grupos mineralógicos a que pertencem e também amostras de outras bases de dados (Universidade de Parma <sup>79</sup> e do Laboratório de Geologia de Lyon <sup>80</sup>). As saídas de cada uma das redes neurais artificiais para este conjunto estão presentes no apêndice A. O campo “Amostra Número” se refere ao número do espectro que será validado pela rede. Assim, para a região de 100-1400  $\text{cm}^{-1}$ : os números 1 a 15 são minerais com alta possibilidade de pertencerem ao grupo do pirocloro (classes 0 e 1), 16 a 38 de serem do grupo da microlita (classes 2 e 3) e 39 a 44 são minerais não pertencentes a estes dois grupos (classe 4). Para a região 1200-1800  $\text{cm}^{-1}$ : 1 a 8 podem ser do grupo do pirocloro e 9 a 20 podem ser da microlita. Na última região, 2800-4000  $\text{cm}^{-1}$ : 1 a 3 podem ser pirocloro e 4 a 19 podem ser microlita.

Os espectros deste conjunto foram comparados com a base de dados do projeto RRUFF <sup>19</sup>. Assim, as amostras de números 11 a 15 na região de 100-1400  $\text{cm}^{-1}$  pertencem a classe 4, por exemplo, como amostras de rutilo. Para ambos os métodos de remoção de ruído, a maior parte das classificações correspondem as informações prévias (os grupos mineralógicos), inclusive para os espectros de outras bases. O que muda de um modo de remoção para outro é o número de correspondência. Em virtude de ocorrer uma perda de informação maior em VP, tanto o número de acertos quanto a probabilidade de ser cada classe é menor em comparação com ALS.

Os dados apresentados na literatura utilizam a base de dados do projeto RRUFF para realizar a classificação de minerais.<sup>89-91</sup> Estes trabalhos focam em mais de quatro espécies de minerais, utilizam outras técnicas de pré-processamento de dados e outros métodos de aprendizado de máquina (além de *Perceptron* de múltiplas camadas, árvores de decisão, *Random Forest* e análise de componentes principais). Ao se comparar os resultados da literatura com os adquiridos neste trabalho na Tabela 4, é observado que se obteve acurácia próxima ou maior, isto sugere que as redes neurais aqui desenvolvidas apresentam bom desempenho nas classificações das espécies minerais estudadas neste trabalho.

## 5 CONCLUSÃO

### 5.1 CONCLUSÕES FINAIS

A espectroscopia Raman é uma ótima técnica para a caracterização de minerais. Em geral, não são necessários procedimentos muito elaborados na preparação das amostras e a técnica não é destrutiva.

Os espectros padrões escolhidos neste trabalho são formados pelos dados do conjunto de treinamento, para os quais são apresentados o resultados para espectroscopia Raman, difração de raios-X por monocristal e a análise de composição química. O intuito disso, é evitar qualquer ambiguidade na identificação das espécies minerais aqui investigadas.

Ao se comparar o tempo de processamento para a remoção do ruído de fundo dos espectros coletados, o método Vizinhos Próximos (VP) foi melhor que Mínimos Quadrados Assimétricos (ALS) implementados. No entanto, em VP houve regiões de informação que foram perdidas e isto provocou na maioria das vezes a redução da acurácia e o aumento da função custo das redes neurais artificiais, como foi observado na na seção 4.2 e Tabela 4.

A melhor topologia encontrada foi com duas camadas ocultas com 10 ou 8 neurônios na última camada oculta, utilizando como função de ativação, na maioria das redes, a tangente hiperbólica. A presença de sobreajuste detectada foi evitada pela parada antecipada ao se obter a maior acurácia para o conjunto de testes dentro de um determinado conjunto de épocas. Por exemplo, a análise da variação da taxa de aprendizagem foi utilizada uma espera de 5000 épocas para se verificar o caso de melhor acurácia.

Os resultados neste trabalho foram comparados com métodos de aprendizado de máquina apresentados na literatura e a acurácia (utilizando as espécies do conjunto de treinamento como referência) obtida pelas redes aqui trabalhadas foi próxima ou maior. Isto indica que as redes neurais artificiais desenvolvidas apresentam boa classificação de algumas espécies minerais do grupo do pirocloro (fluornatropirocloro e fluorcalciopirocloro, que possuem nióbio dominante numa das posições cristalográficas), da microlita (fluornatromicrolita e fluorcalciomicrolita, que possuem o elemento tântalo como dominante) e “outros” (como calcita, quartzo e rutilo).

## 5.2 PERSPECTIVAS FUTURAS

O processo de identificação automática de minerais utilizando a espectroscopia Raman e Redes Neurais artificiais apresentados nesta dissertação deu apenas os seus primeiros passos, com a demonstração de todas as suas etapas e da qualidade dos resultados obtidos.

Existem muitas melhorias e modificações a fazer, sobretudo na perspectiva de torná-lo num processo contínuo, com todas as etapas integradas numa só unidade. Para se poder considerar um processo integrado que inclui desde uma caracterização físico-química que gere um banco de dados mais robusto, o desenvolvimento de técnicas de replicação de banco de dados em software livre, além de correlacionar todas estas etapas no que respeita à velocidade de processamento, representa ainda muito esforço e trabalho futuro.

Face ao exposto, da confluência das visões oferecidas pelo estudo de aprendizado de máquina poderão ser propostos ajustes para a análise do comportamento das principais componentes espectrais de modelos teóricos de espectros Raman de minerais do supergrupo do pirocloro. Seguindo a linha de pesquisa desenvolvida nesta dissertação e tendo em vista os resultados obtidos, alguns tópicos tornam-se promissores como propostas de trabalhos futuros:

(1) Elaborar um banco de dados robusto através da coleta de dados da composição química, parâmetros cristalográficos e espalhamento Raman de minerais do supergrupo do pirocloro.

(2) Explorar as potencialidades de técnicas de aprendizado de máquina visando à obtenção de métodos rápidos e robustos que possam ser aplicados na caracterização de minerais.

(3) Desenvolver algoritmos com aplicabilidade voltada para a caracterização de minerais a partir de dados de espalhamento Raman.

(4) Testar a implementação em computação paralela dos algoritmos desenvolvidos para caracterização de minerais.

(5) Estudar as etapas e parâmetros envolvidos nas configurações dos algoritmos buscando introduzir melhorias no desempenho do mesmo quanto ao ajuste de espectros Raman teóricos e experimentais de minerais.



(6) Testar as eficiências de aprendizado de máquina desenvolvida quanto à obtenção de modelos mais robustos e representativos dos valores dos constituintes a serem estudados com o menor número de amostras possível.

(7) Fornecer elementos para ajustes dos modelos teóricos de espectros Raman.



## REFERÊNCIAS

- 1 ZHANG, B. On typical materials acting as the dividing standard of the developments stages of human substance civilization. **Interdisciplinary Description of Complex Systems**, v. 10, n. 2, p. 114-126, 2012.
- 2 KESLER, S. **Mineral supply and demand into the 21st Century**. p. 55-62, 2007. Disponível em: <<https://pubs.usgs.gov/circ/2007/1294/reports/paper9.pdf>>. Acesso em: 20 abr. 2019.
- 3 CARVALHO, F. P. Mining industry and sustainable development: time for change. **Food and Energy Security**, v. 6, n. 2, p. 61-77, 2017.
- 4 INTERNATIONAL COUNCIL ON MINING & METALS. **MINING'S contribution to sustainable development**. 2012. Disponível em: <<http://www.ibram.org.br/sites/1300/1382/00002639.pdf>>. Acesso em: 20 abr. 2019.
- 5 VILARDAGA, V. A corrida do nióbio. **IstoÉ**, n. 2560, 18 Janeiro 2019. Disponível em: <<https://istoe.com.br/a-corrída-do-niobio/>>. Acesso em: 12 maio 2019.
- 6 KLEIN, C.; DUTROW, B. **Manual of mineral science**. 23rd ed. New York: Wiley, 2007. 704 p.
- 7 ATENCIO, D. et al. The pyrochlore supergroup of minerals: nomenclature. **Canadian Mineralogist**, v. 48, p. 673-698, 2010. doi: 10.3749/canmin.48.3.
- 8 ATENCIO, D. et al. Ralstonite renamed hydrokenoralstonite, coulsellite renamed fluornatrocoulsellite, and their incorporation into the pyrochlore supergroup. **Canadian Mineralogist**, v.55, n. 1, p. 115-120, 2017.
- 9 ANDRADE, M. B. et al. The crystal structure of a microlite-group mineral with a formula near NaCaTa<sub>2</sub>O<sub>6</sub>F from the Morro Redondo mine, Coronel Murta, Minas Gerais, Brazil. **Canadian Mineralogist**, v. 49, n. 2, p. 615-621, 2011.
- 10 ALVES, A. R.; COUTINHO, D. R. The evolution of the niobium production in Brazil. **Materials Research**, v. 18, n. 1, p. 106-112, 2015.
- 11 GRILL, R.; GNADENBERGER, A. Niobium as mint metal: production: properties: processing. **International Journal of Refractory Metals & Hard Materials**, v. 24, n. 4, p. 275-282, 2006.
- 12 BAJWA, I. S.; CHOUDHARY, M. **Prediction of minerals by identifying various types of rocks using neural networks**. 2016. Disponível em: <<https://ieeexplore.ieee.org/stamp/stamp.jsp?tp=&arnumber=4106434>>. Acesso em: 30 abr. 2019.
- 13 THOMPSON, S.; FUETEN, F.; BOCKUS, D. Mineral identification using artificial neural networks and the rotating polarizer stage. **Computers & Geosciences**, v. 27, n. 9, p. 1081-1089, 2001.
- 14 IZADI, H.; SADRI, J.; MEHRAN, N.-A. Intelligent mineral identification using clustering and artificial neural networks techniques. In: PATTERN RECOGNITION AND IMAGE ANALYSIS (PRIA), 2013, Birjand. **Proceedings...** Birjand: IEEE, 2013

15 BAYKAN, N. A.; YILMAZ, N. Mineral identification using color space and artificial neural networks. **Computer & Geoscience**, v. 36, n. 1, p. 91-97, 2010.

16 BAYKAN, N. A.; YILMAZ, N. A Mineral classification system with multilayer artificial neural network using K-Fold cross validation. **Mathematical and Computational Applications**, v. 16, n. 1, p. 22-30, 2011.

17 FUETEN, F.; MASON, J. An artificial neural network assisted approach to editing edges in petrographic images collected with rotating polarizer stage. **Computers & Geosciences**, v. 33, n. 9, p. 1176-1188, 2007.

18 TSUJI, T. et al. Mineral classification from quantitative X-ray maps using neural network: Application to volcanic rocks. **Island Arc**, v. 19, n. 1, p. 105-119, 2010.

19 PROJETO RRUFF. Disponível em: <<http://rruff.info/>>. Acesso em: 20 abr. 2019.

20 RULL, F. The Raman effect and the vibrational dynamics of molecules and crystalline solids. In: DUBESSY, J.; CAUMON, M. C.; RULL, F. (Ed.). **Applications of Raman spectroscopy to earth sciences and cultural heritage**. Great Britain: European Mineralogical Union and Mineralogical Society of Great Britain & Ireland, 2012. v. 12, cap. 1, p. 1-60.

21 WOLVERSON, D. Raman spectroscopy. In: LAMBERTI, C.; AGOSTINI, G. (Ed.). **Characterization of semiconductor heterostructures and nanostructures**. 2nd ed. Amsterdã: Elsevier, 2013. cap. 17, p. 753-802.

22 MOMBERT, A. W. The 50th anniversary of the death of Adolf Gustav Smekal (1895–1959), a pioneer in materials physics. **Journal of Materials Science**. v. 45, n. 3, p. 750-758, 2010.

23 KNEIPP, K. Surface-enhanced Raman scattering. **Physics Today**, v. 60, n. 11, p. 40, 2007.

24 SALA, O. **Fundamentos da espectroscopia Raman e no infravermelho**. 2. ed. São Paulo: UNESP, 2008.

25 YU, P. Y.; CARDONA, M. **Fundamentals of semiconductors**. 4th ed. Berlin: Springer, 2010.

26 LOUDON, R. The Raman effect in crystals. **Advances in Physics**, v. 50, n. 7, p. 813-864, 2001.

27 HAYKIN, S. **Redes neurais: princípios e prática**. Tradução de Paulo Martins Engel. 2. ed. Porto Alegre: Bookman, 2001.

28 SILVA, I. N. D.; SPATTI, D. H.; FLAUZINO, R. A. **Redes neurais artificiais para engenharia e ciências aplicadas**. 2. ed. São Paulo: Artlibert, 2010.

29 MACUKOW, B. **Neural networks – state of art, brief history, basic models and architecture**. 2016. Disponível em: <[https://link.springer.com/content/pdf/10.1007%2F978-3-319-45378-1\\_1.pdf](https://link.springer.com/content/pdf/10.1007%2F978-3-319-45378-1_1.pdf)>. Acesso em: 23 jan. 2019.

30 DEEP Learning Book. 2018. Disponível em: <<http://www.deeplearningbook.com.br>>. Acesso em: 10 dez. 2018.

- 31 SNELL, R. S. **Clinical neuroanatomy**. 7th ed. Philadelphia: Ed. Lippincott Williams & Wilkins, 2010.
- 32 ROUGIER, N. P. **Scanning electron microscope image of a neuron**. Disponível em: <[https://www.upi.com/Science\\_News/2014/02/12/Scientists-locate-new-connection-between-fear-and-behavior-in-brain/4661392220731/](https://www.upi.com/Science_News/2014/02/12/Scientists-locate-new-connection-between-fear-and-behavior-in-brain/4661392220731/)>. Acesso em: 20 abr. 2019.
- 33 PANCHAL, G. et al. Behaviour analysis of multilayer perceptrons with multiple hidden neurons and hidden layers. **International Journal of Computer Theory and Engineering**, v. 3, n. 2, p. 332-337, 2011.
- 34 SCHMIDHUBER, J. Deep learning in neural networks: an overview. **Neural Networks**, v. 61, p. 85-117, 2015. doi:10.1016/j.neunet.2014.09.003
- 35 GALLO, C. Artificial neural networks: tutorial. In: KHOSROW-POUR, M. (Ed.). **Encyclopedia of information science and technology**. 3rd ed. 2015. p. 179-189. Disponível em: <[https://www.researchgate.net/publication/261392616\\_Artificial\\_Neural\\_Networks\\_tutorial](https://www.researchgate.net/publication/261392616_Artificial_Neural_Networks_tutorial)>. Acesso em: 20 abr. 2019.
- 36 LAROCHELLE, H. et al. Exploring strategies for training deep neural networks. **Journal of Machine Learning Research**, v. 1, p. 1-40, 2009. Disponível em: <<http://www.jmlr.org/papers/volume10/larochelle09a/larochelle09a.pdf>>. Acesso em: 30 abr. 2019.
- 37 ALAVAREZ, L. et al. (Ed.). **Progress in pattern recognition, image analysis, computer vision, and applications**. Berlin: Springer-Verlag, 2012.
- 38 ABDULJABBAR, R. et al. Applications of artificial intelligence in transport: an overview. **Sustainability**, v. 11, p. 1-24, 2019. doi: 10.3390/su11010189.
- 39 HONÓRIO, K. M. et al. Artificial neural networks and the study of the psychoactivity of cannabinoid compounds. **Chemical Biology & Drug Design**, v. 75, n. 6, p. 632-640, 2010.
- 40 HIMMELBLAU, D. M. Applications of artificial neural networks in chemical engineering. **Korean Journal of Chemical Engineering**, v. 17, n. 4, p. 373-392, 2000.
- 41 JIANG, J.; TRUNDLE, P.; REN, J. Medical image analysis with artificial neural networks. **Computerized Medical Imaging and Graphics**, v. 34, n. 8, p. 617-631, 2010.
- 42 NÆS, T. et al. Artificial neural networks in multivariate calibration. **Journal of Near Infrared Spectroscopy**, v. 1, n. 1, p. 1-11, 1993.
- 43 MUNK, M. E.; MADISON, M. S.; ROBB, E. W. Neural-network models for infrared spectrum interpretation. **Mikrochimica Acta**, v. 2, n. 1-6, p. 505-524, 1991.
- 44 MEYER, M.; WEIGELT, T. Interpretation of infrared spectra by artificial neural networks. **Analytica Chimica Acta**, v. 265, n. 2, p. 183-190, 1992.
- 45 SMITS, J. R. et al. Interpretation of infrared spectra with modular neural-network systems. **Chemometrics and Intelligent Laboratory Systems**, v. 18, n. 1, p. 27-39, 1993.
- 46 CIROVIC, D. A. Feed-forward artificial neural networks: applications to spectroscopy. **TrAC Trends in Analytical Chemistry**, v. 16, n. 3, p. 148-155, 1997.

- 47 LONG, J. R.; GREGORIOU, V. G.; GEMPERLINE, P. J. Spectroscopic calibration and quantitation using artificial neural networks. **Analytical Chemistry**, v. 62, p. 1791-1797, 1990. doi:10.1021/ac00216a013.
- 48 MCCANN, M. C. et al. Neural network analyses of infrared spectra for classifying cell wall architectures. **Plant Physiology**, v. 143, n. 3, p. 1314–1326, 2007.
- 49 TAKAHAMA, S. et al. Atmospheric particulate matter characterization by Fourier transform infrared spectroscopy: a review of statistical calibration strategies for carbonaceous aerosol quantification in US measurement networks. **Atmospheric Measurements Techniques**, v. 12, p. 525-567, 2019. doi: 10.5194/amt-12-525-2019.
- 50 AHMED, M. U. et al. Machine learning approach for biomass characterization. **Energy Procedia**, v. 158, p. 1279-1287, 2019. doi.org/10.1016/j.egypro.2019.01.316
- 51 LARKIN, P. **Infrared and Raman spectroscopy: principles and spectral interpretation**. San Diego: Elsevier, 2011. 230 p.
- 52 CHUKANOV, N. V.; CHERVONNYI, A. D. **Infrared spectroscopy of minerals and related compounds**. Moscow: Springer, 2016.
- 53 NASRAOUI, M.; BILAL, E.; GIBERT, R. Fresh and weathered pyrochlore studies by Fourier transform infrared spectroscopy coupled with thermal analysis. **Mineralogical Magazine**, v. 63, n. 4, p. 567-578, 1999.
- 54 ALI, S. M. et al. A comparison of Raman, FTIR and ATR-FTIR micro spectroscopy for imaging human skin tissue sections. **Analytical Methods**, v. 5, n. 9, p. 2281-2291, 2013.
- 55 GRIFFITH, W. P. Raman spectroscopy of minerals. **Nature**, v. 224, p. 264-266, 1969. doi: 10.1038/224264a0.
- 56 HASKIN, L. A. et al. Raman spectroscopy for mineral identification and quantification for in situ planetary surface analysis: A point count method. **Journal of Geophysical Research Atmospheres**, v. 102, n. 8, p. 293-306, 1997.
- 57 BERSANI, D.; LOTTICI, P. P. Raman spectroscopy of minerals and mineral pigments in archaeometry. **Journal of Raman Spectroscopy**, v. 47, n. 5, p. 499–530, 2016.
- 58 WHITE, W. B. Identification of cave minerals by Raman spectroscopy: new technology for non-destructive analysis. **International Journal of Speleology**, v. 35, n. 2, p. 103-107, 2006.
- 59 SRODEK, D.; DULSKI, M.; GALUSKINA, I. Raman imaging as a new approach to identification of the mayenite group minerals. **Scientific Reports**, v. 8, 2018. doi: 10.1038/s41598-018-31809-4.
- 60 WHITE, S. N. Laser Raman spectroscopy as a technique for identification of seafloor hydrothermal and cold seep minerals. **Chemical Geology**, v. 259, n. 3-4, p. 240-252, 2009.
- 61 WEBER, I. et al. Mineralogical and Raman spectroscopy studies of natural olivines exposed to different planetary environments. **Planetary and Space Science**, v. 104, Part B, p. 163–172, 2014.

62 ANDRADE, M. B. et al. Hydroxycalciumicrolite,  $\text{Ca}_{1.5}\text{Ta}_2\text{O}_6(\text{OH})$ , a new member of the microlite group from Volta Grande pegmatite, Nazareno, Minas Gerais, Brazil. **Mineralogical Magazine**, v. 81, n. 3, p. 555-564, 2017.

63 ANDRADE, M. B. et al. The crystal structure of a microlite-group mineral with a formula near  $\text{NaCaTa}_2\text{O}_6\text{F}$  from the Morro Redondo mine, Coronel Murta, Minas Gerais, Brazil. **Canadian Mineralogist**, v. 49, p. 615-621, 2011. doi: 10.3749/canmin.49.2.615.

64 AZO MATERIALS. **LabRAM HR UV-VIS-NIR Raman microscope from HORIBA**. Disponível em: <<https://www.azom.com/equipment-details.aspx?EquipID=1390>>. Acesso em: 25 abr. 2019.

65 TUSCHEL, D. Selecting an excitation wavelength for raman spectroscopy. **Spectroscopy**, v. 31, n. 3, p. 14-23, 2016.

66 WALRAFEN, G. E. Raman spectral studies of water structure. **Journal of Chemical Physics**, v. 40, n. 11, p. 3249-3256, 1964.

67 WALRAFEN, G. E. Raman spectral studies of the effects of temperature on water structure. **Journal of Chemical Physics**, v. 47, n. 1, p. 114-126, 1967.

68 WANG, Z. et al. Vibrational substructure in the OH stretching transition of water and HOD. **Journal of Physical Chemistry A**, v. 108, n. 42, p. 9054-9063, 2004.

69 ADAR, F. Introduction to Interpretation of Raman Spectra Using Database Searching and Functional Group Detection and Identification. **Spectroscopy**, v. 31, n. 7, p. 16-23, 2016.

70 SUN, Q. The Raman OH stretching bands of liquid water. **Vibrational Spectroscopy**, v. 51, n. 2, p. 213-217, 2009.

71 LAFUENTE, B. et al. The power of databases: the RRUFF project. In: ARMBRUSTER, T.; DANISI, R. M. **Highlights in mineralogical crystallography**. Berlin: Walter de Gruyter GmbH, 2015. cap. 1, p. 1-30.

72 EILERS, P. H. C.; BOELEN, H. F. M. **Baseline correction with asymmetric least squares smoothing**. 2005. Disponível em: <[https://zanran\\_storage.s3.amazonaws.com/www.science.uva.nl/ContentPages/443199618.pdf](https://zanran_storage.s3.amazonaws.com/www.science.uva.nl/ContentPages/443199618.pdf)>. Acesso em: 23 jan. 2019.

73 SUN, K. et al. Baseline correction for Raman spectra based on piecewise linear fitting. **Spectroscopy Online**, v. 29, n. 2, 2014. Disponível em: <<http://www.spectroscopyonline.com/baseline-correction-raman-spectra-based-piecewise-linear-fitting>>. Acesso em: 22 maio 2019

74 SCIPY.ORG. **Scipy.signal.savgol\_filter**. Disponível em: <[https://docs.scipy.org/doc/scipy-0.15.1/reference/generated/scipy.signal.savgol\\_filter.html](https://docs.scipy.org/doc/scipy-0.15.1/reference/generated/scipy.signal.savgol_filter.html)>. Acesso em: 20 abr. 2019.

75 GÉRON, A. **Hands-on machine learning with scikit-learn and tensorflow**. New York: O'Reilly Media, 2017.

76 ABADI, M. et al. **Tensorflow: large-scale machine learning on heterogeneous distributed systems**. 2015. Disponível em: <<http://download.tensorflow.org/paper/whitepaper2015.pdf>>. Acesso em: 23 abr. 2019.

77 CHOLLET, F. **Keras**: the Python deep learning library. Disponível em: <<https://keras.io/>>. Acesso em: 23 abr. 2019.

78 PONTI, M. A.; COSTA, G. B. P. Como funciona o Deep Learning. In: VIERA, V.; RAZENTE, H. L.; BARIONI, M. C. N. **Tópicos em gerenciamento de dados e informações 2017**. Uberlandia: SBC, 2017. cap. 3, p. 63-93.

79 LABORATORY of Photoinduced Effects Vibrational and X-ray Spectroscopies. Disponível em: <<https://www.fis.unipr.it/phevix/ramandb.php>>. Acesso em: 15 maio 2019.

80 HANDBOOK of Raman Spectra for geology. Disponível em: <<https://www.fis.unipr.it/phevix/ramandb.php>>. Acesso em: 15 maio 2019.

81 ANDRADE, M. B. et al. Fluorcalciomicrolite,  $(Ca,Na,&)_2Ta_2O_6F$ , a new microlite-group mineral from Volta Grande pegmatite, Nazareno, Minas Gerais, Brazil. **Mineralogical Magazine**, v. 77, n. 7, p. 2989-2996, 2013.

82 ATENCIO, D.; CIRIOTTI, M. E.; ANDRADE, M. B. Fluorcalcioroméite,  $(Ca,Na)_2Sb_5+2(O,OH)_6F$ , a new roméite-group mineral from Starlera mine, Ferrera, Grischun, Switzerland: description and crystal structure. **Mineralogical Magazine**, v. 77, n. 4, p. 467-473, 2013.

83 BELAKOVSKIY, D. I.; CÁMARA, F.; GAGNÉ, O. C. New mineral names. **American Mineralogist**, v. 103, n.12, p. 2036-2048, 2018.

84 OLIVEIRA, E. A. et al. Crystal structure and vibrational spectrum of the  $NaCaMg_2F_7$  pyrochlore. **Journal of Solid State Chemistry**, v. 177, n. 8, p. 2943-2950, 2004.

85 GARBOUT, A. et al. Raman scattering and X-ray diffraction on  $YBiTi_2O_7$  prepared at low temperature. **Journal of Raman Spectroscopy**, v. 39, n. 10, p. 1469-1474, 2008.

86 FUENTES, A. F. et al. Synthesis of disordered pyrochlores,  $A_2Ti_2O_7$  (A=Y, Gd and Dy), by mechanical milling of constituent oxides. **Solid State Sciences**, v. 7, n. 4, p. 343-353, 2005.

87 KONG, L. et al. Aqueous chemical synthesis of  $Ln_2Sn_2O_7$  pyrochlore-structured ceramics. **Journal of the American Ceramic Society**, v. 96, n. 9, p. 2994-3000, 2013.

88 ARENAS, D. J. et al. Raman study of phonon modes in bismuth pyrochlores. **Physical Review B**, v. 82, n. 21, p. 214302, 2010.

89 CAREY, C. et al. Machine learning tools for mineral recognition and classification from Raman spectroscopy. **Journal of Raman Spectroscopy**, v. 46, n. 10, p. 894-903, 2015.

90 ISHIKAWA, S. T.; GULICKAC, V. C. An automated mineral classifier using Raman spectra. **Computers & Geosciences**, v. 54, p. 259-268, 2013. doi:10.1016/j.cageo.2013.01.011

91 SEVETLIDIS, V.; PAVLIDIS, G. Effective Raman spectra identification with tree-based methods. **Journal of Cultural Heritage**, v. 37, p. 121-128, 2018. doi: 10.1016/j.culher.2018.10.016.



## APÊNDICE A – SAÍDA DAS REDES NEURAS ARTIFICIAIS PARA O CONJUNTO DE VALIDAÇÃO

O campo “Amostra Número” se refere qual o número do espectro que será validado pela rede, “Classificado pela Rede” é a classificação da rede neural artificial e “Probabilidade de ser Cada Classe” é qual a probabilidade de uma amostra ser de cada uma das classes em um intervalo de 0 a 1.

### REGIÃO DE 100-1400 $\text{cm}^{-1}$

Por ALS:

Amostra Número: 10

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.02808402 0.6516894 0.08563016 0.21488999 0.01970641]

Amostra Número: 11

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [1.0261138e-03 2.7025873e-03 1.6828447e-04 1.1012064e-03 9.9500185e-01]

Amostra Número: 12

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [9.7702525e-04 3.9253505e-03 1.5672199e-04 1.5007901e-03 9.9344009e-01]

Amostra Número: 13

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [2.0393319e-03 4.1236319e-03 3.1332741e-04 1.2997852e-03 9.9222386e-01]

Amostra Número: 14

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [1.9774747e-03 3.0689377e-03 2.6049287e-04 9.0758206e-04 9.9378550e-01]

Amostra Número: 15

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [1.5441503e-03 9.6615022e-03 2.0221647e-04 2.8937499e-03 9.8569840e-01]

Amostra Número: 16

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00492456 0.00186547 0.86321145 0.12700263 0.00299584]

Amostra Número: 17

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00509102 0.00177593 0.8687099 0.12160431 0.00281875]

Amostra Número: 18

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [5.8030978e-02 9.6647397e-02 4.8676702e-01 3.5812965e-01 4.2501639e-04]

Amostra Número: 19

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00970808 0.00606664 0.55264145 0.23893432 0.19264951]

Amostra Número: 1

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03497982 0.93930554 0.00820796 0.01625133 0.00125549]

Amostra Número: 20

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00485306 0.00118728 0.85267836 0.1319074 0.00937376]

Amostra Número: 21

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [2.1363553e-02 9.6379858e-01 2.9782832e-03 1.1515543e-02 3.4398166e-04]

Amostra Número: 22

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00642249 0.00651649 0.85455924 0.12395701 0.00854479]

Amostra Número: 23

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00621731 0.00623229 0.8535437 0.12590133 0.00810546]

Amostra Número: 24

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00466169 0.00156857 0.8011861 0.18567596 0.00690776]

Amostra Número: 25

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01849685 0.02146497 0.6527081 0.15293752 0.15439251]

Amostra Número: 26

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00629517 0.00546477 0.85464793 0.12449572 0.00909644]

Amostra Número: 27

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00590558 0.0108072 0.801357 0.1685433 0.0133869 ]

Amostra Número: 28

Classificado pela Rede: 3

Probabilidade de ser Cada Classe: [0.00135191 0.06346968 0.04932879 0.87874776 0.00710185]

Amostra Número: 29

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.03916493 0.26970848 0.44638675 0.19507341 0.04966643]

Amostra Número: 2

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03752299 0.94117624 0.00700508 0.01318078 0.00111474]

Amostra Número: 30

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [2.61392482e-02 9.60220456e-01 2.74824398e-03 1.06088435e-02 2.83135130e-04]

Amostra Número: 31

Classificado pela Rede: 3

Probabilidade de ser Cada Classe: [0.01514003 0.08711544 0.12081232 0.6367532 0.14017911]

Amostra Número: 32

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.03973059 0.06783165 0.73147416 0.08801947 0.07294413]

Amostra Número: 33

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [2.5410732e-02 9.6336716e-01 2.1726443e-03 8.7437211e-03 3.0582037e-04]

Amostra Número: 34

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00566387 0.00666541 0.8325 0.14785662 0.00731416]

Amostra Número: 35

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00532104 0.0043739 0.85437167 0.13051386 0.00541955]

Amostra Número: 36

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01898196 0.05783706 0.6595748 0.21446638 0.04913977]

Amostra Número: 37

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00424778 0.00480625 0.78591686 0.19709897 0.00793008]

Amostra Número: 38

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [2.5282312e-02 9.1620630e-01 1.2647891e-02 4.5378398e-02 4.8501202e-04]

Amostra Número: 39

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.00348904 0.00398665 0.00186027 0.00596984 0.9846942 ]

Amostra Número: 3

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.08324939 0.66537327 0.12445834 0.10325793 0.02366112]

Amostra Número: 40

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.00776137 0.0128743 0.00697422 0.0283895 0.94400054]

Amostra Número: 41

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [3.4753438e-03 1.3719125e-02 9.0617186e-04 7.1165501e-03 9.7478288e-01]

Amostra Número: 42

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [2.2380683e-03 8.7849246e-03 4.8291634e-04 4.1181594e-03 9.8437595e-01]

Amostra Número: 43

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [5.5734771e-03 4.1224030e-03 4.3873984e-04 6.5418129e-04 9.8921120e-01]

Amostra Número: 44

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [3.7781107e-03 3.7495710e-03 2.4551552e-04 6.0797844e-04 9.9161875e-01]

Amostra Número: 4

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03544061 0.79322875 0.0683687 0.08248805 0.02047386]

Amostra Número: 5

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.02102609 0.9358397 0.00939114 0.03242101 0.00132199]

Amostra Número: 6

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.04556119 0.13252857 0.35015574 0.2152254 0.25652912]

Amostra Número: 7

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.06364485 0.85575444 0.03582486 0.03487948 0.00989649]

Amostra Número: 8

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03570814 0.9121975 0.01397971 0.03529258 0.00282212]

Amostra Número: 9

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.02724077 0.35393053 0.07055712 0.17851152 0.36976007]

Por VP:

Amostra Número: 10

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.01700125 0.8299562 0.02391546 0.03781102 0.09131601]

Amostra Número: 11

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.01169756 0.01051008 0.00220096 0.0069954 0.968596 ]

Amostra Número: 12

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.01169756 0.01051008 0.00220096 0.0069954 0.968596 ]

Amostra Número: 13

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.01399851 0.02352944 0.00705673 0.01892521 0.93649006]

Amostra Número: 14

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.00963513 0.01384421 0.00466284 0.01181958 0.96003824]

Amostra Número: 15

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.02911536 0.39883003 0.01046849 0.03171848 0.52986765]

Amostra Número: 16

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00874617 0.00964437 0.7996247 0.17577411 0.00621071]

Amostra Número: 17

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00625269 0.0075642 0.7840453 0.19532715 0.00681068]

Amostra Número: 18

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00878904 0.01147397 0.7731079 0.2011832 0.00544594]

Amostra Número: 19

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01583497 0.00888845 0.6225769 0.30750772 0.04519197]

Amostra Número: 1

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03558213 0.95085984 0.00514106 0.00442111 0.00399584]

Amostra Número: 20

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00852464 0.01047516 0.7733266 0.20044295 0.00723075]

Amostra Número: 21

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01783428 0.01054266 0.80108774 0.16019692 0.01033835]

Amostra Número: 22

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.02661587 0.01962172 0.7860897 0.15999012 0.00768256]

Amostra Número: 23

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.02257128 0.01780291 0.78639734 0.16607705 0.00715136]

Amostra Número: 24

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00987867 0.01728674 0.72474384 0.24206774 0.00602302]

Amostra Número: 25

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.04667755 0.03298282 0.5639644 0.3141617 0.04221348]

Amostra Número: 26

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01283426 0.01554819 0.75681514 0.20751068 0.00729167]

Amostra Número: 27

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.0069129 0.00907126 0.7062169 0.26780087 0.00999804]

Amostra Número: 28

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00709327 0.010355 0.78724146 0.18860109 0.00670921]

Amostra Número: 29

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.0383296 0.91564566 0.02172345 0.01811977 0.0061815 ]

Amostra Número: 2

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03558213 0.95085984 0.00514106 0.00442111 0.00399584]

Amostra Número: 30

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01493296 0.00800428 0.8264982 0.13490391 0.01566063]

Amostra Número: 31

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.0107213 0.01064058 0.67914283 0.271534 0.02796125]

Amostra Número: 32

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.04990408 0.59562206 0.20891318 0.13196313 0.01359753]

Amostra Número: 33

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01753058 0.00734567 0.8042802 0.15254375 0.01829983]

Amostra Número: 34

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00798355 0.01204899 0.7479328 0.22674581 0.00528882]

Amostra Número: 35

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00823482 0.01235404 0.74739784 0.22670901 0.00530433]

Amostra Número: 36

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.06347686 0.15971541 0.4934123 0.25790548 0.02548989]

Amostra Número: 37

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.01426633 0.02712543 0.72457576 0.22730711 0.00672522]

Amostra Número: 38

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.00787002 0.01196678 0.7725669 0.20191269 0.00568355]

Amostra Número: 39

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.02030837 0.01759619 0.01501426 0.02811818 0.918963 ]

Amostra Número: 3

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03757666 0.9239871 0.01834857 0.01381894 0.00626871]

Amostra Número: 40

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.04027188 0.06449526 0.05524142 0.08888656 0.7511049 ]

Amostra Número: 41

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.02914998 0.04162477 0.02611824 0.05169903 0.851408 ]

Amostra Número: 42

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.01848076 0.02798746 0.01326825 0.02680704 0.9134565 ]

Amostra Número: 43

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.02946617 0.0211513 0.00213779 0.00778834 0.9394564 ]

Amostra Número: 44

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.03174484 0.01705086 0.00194676 0.00646831 0.94278926]

Amostra Número: 4

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.01528151 0.9288216 0.00832239 0.01165462 0.0359199 ]

Amostra Número: 5

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.01787351 0.92524135 0.01379139 0.01611017 0.02698372]

Amostra Número: 6

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.01778618 0.7012174 0.03072315 0.06441639 0.18585674]

Amostra Número: 7

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.03467955 0.92657423 0.01804251 0.01333711 0.00736666]

Amostra Número: 8

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.04110869 0.9166794 0.01934188 0.01288605 0.0099841 ]

Amostra Número: 9

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.02801937 0.6888748 0.05941318 0.09437823 0.12931447]



**REGIÃO DE 1200-1800 cm<sup>-1</sup>**

Por ALS:

Amostra Número: 10

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.30219948 0.28688192 0.22904517 0.1103579 0.07151554]

Amostra Número: 11

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.07194213 0.20822491 0.5138053 0.12473278 0.08129492]

Amostra Número: 12

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.07816683 0.15375565 0.46920756 0.11706378 0.18180616]

Amostra Número: 13

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.18163486 0.21778867 0.21243116 0.07458138 0.3135639 ]

Amostra Número: 14

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.06041911 0.05996352 0.13577357 0.03234954 0.7114942 ]

Amostra Número: 15

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.13123581 0.23821467 0.4222455 0.1194764 0.0888276 ]

Amostra Número: 16

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.03255138 0.13309968 0.6288932 0.16657533 0.03888042]

Amostra Número: 17

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.22779846 0.26959342 0.2181524 0.07733764 0.20711803]

Amostra Número: 18

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.06714997 0.09575986 0.2927663 0.07064535 0.4736784 ]

Amostra Número: 19

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.02743415 0.08056223 0.40286624 0.09172299 0.3974144 ]

Amostra Número: 1

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.37630296 0.21340856 0.19099438 0.09670734 0.12258674]

Amostra Número: 20

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.07598394 0.11871728 0.35646942 0.09063011 0.35819918]

Amostra Número: 2

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.24880752 0.20102148 0.24041428 0.1276526 0.18210416]

Amostra Número: 3

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.07812119 0.11971794 0.31544653 0.10477506 0.38193932]

Amostra Número: 4

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.4384139 0.20220569 0.12195607 0.08622574 0.1511986 ]

Amostra Número: 5

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.0269016 0.05138648 0.18747957 0.04815599 0.6860764 ]

Amostra Número: 6

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.6749198 0.19196142 0.03661932 0.03281088 0.06368858]

Amostra Número: 7

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.11693057 0.1620968 0.27744654 0.10882994 0.33469617]

Amostra Número: 8

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.5942406 0.18616727 0.04924708 0.04997237 0.12037266]

Amostra Número: 9

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.07632176 0.19886158 0.4886531 0.19059563 0.04556793]

Por VP:

Amostra Número: 10

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.24631612 0.26519957 0.1657458 0.14332017 0.17941837]

Amostra Número: 11

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.25104973 0.29304937 0.14254738 0.13456976 0.17878377]

Amostra Número: 12

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.2453037 0.24842998 0.16823733 0.14897136 0.18905757]

Amostra Número: 13

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.20040321 0.2558174 0.18797708 0.1261521 0.22965023]

Amostra Número: 14

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.2548231 0.27653027 0.14762613 0.14163615 0.17938441]

Amostra Número: 15

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.24670124 0.30960527 0.14156307 0.12769382 0.17443657]

Amostra Número: 16

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.16871421 0.16874829 0.27718407 0.14158 0.24377343]

Amostra Número: 17

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.22421217 0.30060264 0.14716542 0.12189668 0.20612307]

Amostra Número: 18

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.22841953 0.24154827 0.18839607 0.14484282 0.19679332]

Amostra Número: 19

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.08814709 0.08409329 0.34790412 0.11211817 0.36773738]

Amostra Número: 1

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.26747063 0.27626836 0.15397537 0.1469023 0.1553833 ]

Amostra Número: 20

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.21866086 0.23083128 0.2025395 0.14466129 0.20330706]

Amostra Número: 2

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.24324328 0.2860095 0.15584746 0.13509403 0.17980582]

Amostra Número: 3

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.1499863 0.16489026 0.2908149 0.13012443 0.26418418]

Amostra Número: 4

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.24044377 0.27199015 0.15813822 0.13868554 0.19074233]

Amostra Número: 5

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.10607101 0.11275306 0.3425534 0.11647997 0.3221426 ]

Amostra Número: 6

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.22320925 0.23002397 0.1893372 0.14661767 0.21081199]

Amostra Número: 7

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.11123298 0.12280113 0.32413328 0.11625709 0.32557547]

Amostra Número: 8

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.24369624 0.30321002 0.13748772 0.12815502 0.18745096]

Amostra Número: 9

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.19480638 0.19713137 0.24293715 0.14586927 0.21925585]

**REGIÃO DE 2800-4000 cm<sup>-1</sup>**

Por ALS:

Amostra Número: 10

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.03678498 0.07557427 0.6097251 0.22771665 0.05019893]

Amostra Número: 11

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.04493729 0.06997742 0.6809576 0.15683661 0.04729107]

Amostra Número: 12

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.04429495 0.07218421 0.67405474 0.1581364 0.0513297 ]

Amostra Número: 13

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.16501287 0.7320955 0.05181221 0.02822767 0.02285169]

Amostra Número: 14

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.20616114 0.5699312 0.05506152 0.04192917 0.12691696]

Amostra Número: 15

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.23245728 0.62830794 0.04579242 0.02949763 0.06394474]

Amostra Número: 16

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.1599782 0.66762763 0.09905647 0.04578843 0.02754928]

Amostra Número: 17

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.27435288 0.57590103 0.04673247 0.02831044 0.07470308]

Amostra Número: 18

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.273902 0.5762922 0.04677113 0.02832422 0.07471044]

Amostra Número: 19

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.03798253 0.02362959 0.01684492 0.01090816 0.91063476]

Amostra Número: 1

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.57213616 0.20947778 0.12190532 0.02442804 0.07205269]

Amostra Número: 2

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.0680922 0.13350731 0.1702961 0.06232048 0.5657839 ]

Amostra Número: 3

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.05349474 0.0931576 0.08599494 0.03927518 0.72807753]

Amostra Número: 4

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.11682624 0.78324926 0.05525998 0.02932469 0.01533989]

Amostra Número: 5

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.19431295 0.5502688 0.17139661 0.05688608 0.02713557]

Amostra Número: 6

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.04025617 0.02361272 0.01853992 0.01111478 0.90647644]

Amostra Número: 7

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.03692041 0.06839869 0.7591966 0.11156109 0.02392316]

Amostra Número: 8

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.06608089 0.10061049 0.53422004 0.20742501 0.09166362]

Amostra Número: 9

Classificado pela Rede: 2

Probabilidade de ser Cada Classe: [0.06356847 0.09176102 0.72870094 0.08436468 0.03160489]

Por VP:

Amostra Número: 10

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.18990164 0.2103351 0.19546726 0.19135903 0.21293707]

Amostra Número: 11

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.23736016 0.20621416 0.18502694 0.10794488 0.26345387]

Amostra Número: 12

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.23802207 0.20941919 0.18701434 0.10690236 0.2586421 ]

Amostra Número: 13

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.37007445 0.2358614 0.09739637 0.03756864 0.25909913]

Amostra Número: 14

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.3983427 0.20062529 0.08694938 0.03306559 0.281017 ]

Amostra Número: 15

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.39866403 0.20122467 0.08672792 0.03351676 0.27986664]

Amostra Número: 16

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.31286535 0.24613844 0.12256902 0.0369018 0.2815254 ]

Amostra Número: 17

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.38897163 0.1809359 0.09204775 0.03685582 0.3011889 ]

Amostra Número: 18

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.38916954 0.18104501 0.09204603 0.0368033 0.3009361 ]

Amostra Número: 19

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.11122616 0.2067063 0.16681837 0.01998086 0.49526832]

Amostra Número: 1

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.3967402 0.18567961 0.11610301 0.03373573 0.2677415 ]

Amostra Número: 2

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.19435266 0.22638234 0.20109345 0.05740034 0.3207712 ]

Amostra Número: 3

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.19435266 0.22638234 0.20109345 0.05740034 0.3207712 ]

Amostra Número: 4

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.39237303 0.2029418 0.0914616 0.03383555 0.27938804]

Amostra Número: 5

Classificado pela Rede: 0

Probabilidade de ser Cada Classe: [0.31165624 0.25187233 0.12303907 0.03858811 0.27484423]

Amostra Número: 6

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.24114664 0.18906091 0.13523403 0.01630963 0.41824877]

Amostra Número: 7

Classificado pela Rede: 1

Probabilidade de ser Cada Classe: [0.20635088 0.2258235 0.21538132 0.13827153 0.2141727 ]

Amostra Número: 8

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.19458652 0.22479895 0.22375572 0.11744824 0.2394106 ]

Amostra Número: 9

Classificado pela Rede: 4

Probabilidade de ser Cada Classe: [0.2154159 0.24934083 0.20731038 0.06630885 0.2616241 ]