

UNIVERSIDADE ESTADUAL DE PONTA GROSSA
SETOR DE CIÊNCIAS EXATAS E NATURAIS
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS

GIOVANNY LUTIERRI VIANA

**PREDIÇÃO DOS PARÂMETROS GAUSSIANOS DOS ESPECTROS RAMAN DE
VIDROS TELURETOS UTILIZANDO REDES NEURAIS PERCEPTRON
MULTICAMADAS**

PONTA GROSSA

2025

GIOVANNY LUTIERRI VIANA

**PREDIÇÃO DOS PARÂMETROS GAUSSIANOS DOS ESPECTROS RAMAN DE
VIDROS TELURETOS UTILIZANDO REDES NEURAIS PERCEPTRON
MULTICAMADAS**

Dissertação apresentada junto ao Programa de Pós-Graduação em Ciências da Universidade Estadual de Ponta Grossa como requisito parcial para obtenção do título de mestre em Ciências/Física.

Orientador: Prof. Dr. Aloisi Somer.

PONTA GROSSA

2025

V614 Viana, Giovanni Luttieri
Predição dos parâmetros gaussianos dos espectros raman de vidros teluretos utilizando redes neurais perceptron multicamadas / Giovanni Luttieri Viana. Ponta Grossa, 2026.
138 f.

Dissertação (Mestrado em Ciências - Área de Concentração: Física), Universidade Estadual de Ponta Grossa.

Orientador: Prof. Dr. Aloisi Somer.
Coorientadora: Profa. Dra. Andressa Novatski.

1. Espectroscopia Raman. 2. Aprendizado de máquina. 3. Ajuste gaussiano. 4. Análise espectral. 5. Materiais vítreos. I. Somer, Aloisi. II. Novatski, Andressa. III. Universidade Estadual de Ponta Grossa. Física. IV.T.

CDD: 535

TERMO DE APROVAÇÃO

Giovanny Luttieri Viana

PREDIÇÃO DOS PARÂMETROS GAUSSIANOS DOS ESPECTROS RAMAN DE VIDROS TELURETOS UTILIZANDO REDES NEURAIS PERCEPTRON MULTICAMADAS

Dissertação aprovada como requisito parcial para obtenção do grau de Mestre em Ciências no Programa de Pós-Graduação em Ciências, Área de Concentração: Física dos Materiais, da Universidade Estadual de Ponta Grossa, pela seguinte banca examinadora:

Orientador -



Documento assinado digitalmente
ALOISI SOMER
Data: 09/02/2026 10:59:06-0300
Verifique em <https://validar.iti.gov.br>

Aloisi Somer – UEPG – Presidente



Documento assinado digitalmente
ANTONIO MARCOS BATISTA
Data: 09/02/2026 11:36:22-0300
Verifique em <https://validar.iti.gov.br>

Antonio Marcos Batista – UEPG – Titular



Documento assinado digitalmente
DANIELE TONIOLO DIAS FERREIRA ROSA
Data: 09/02/2026 11:11:37-0300
Verifique em <https://validar.iti.gov.br>

Daniele Toniolo Dias Ferreira Rosa – UTFPR-PG – Titular

Ponta Grossa, 09 de fevereiro de 2026

AGRADECIMENTOS

Agradeço ao Prof. Dr. Aloisi Somer, meu orientador, pela amizade, pelas contribuições fundamentais ao longo do Mestrado, pelos ensinamentos científicos e pelas oportunidades que foram essenciais para o desenvolvimento deste trabalho, bem como pelas correções realizadas na etapa de qualificação. Muito obrigada por tudo!

Agradeço também à Profa. Dra. Andressa Novatski pela amizade e pelas valiosas sugestões oferecidas durante o Mestrado. Meu sincero agradecimento!

Expresso minha gratidão aos demais professores e funcionários do Departamento de Física e do PPG Ciências/Física, que, direta ou indiretamente, contribuíram para este processo.

Sou profundamente grato aos meus pais, Selma e Samuel, por toda a educação, pelos valores transmitidos e por serem tão importantes em minha vida. Aos meus irmãos, Hayssa e Henrique, agradeço pela compreensão em todos os momentos desta e de outras jornadas.

Agradeço igualmente aos meus colegas do grupo de pesquisa GPEViM — Guilherme Teixeira Santos, Marcos Vinícius Gauglitz e Jaqueline Valeski Gunha — pelos inúmeros momentos agradáveis e pela colaboração ao longo desses anos.

Meu agradecimento se estende a todos os amigos da graduação, da pós-graduação e de todos os meus amigos da Universidade Estadual de Ponta Grossa que compartilharam os inúmeros desafios enfrentados, sempre com espírito colaborativo

À Universidade Estadual de Ponta Grossa, agradeço pela oportunidade e pela infraestrutura disponibilizada para a realização dos estudos.

Às agências de fomento FINEP, CAPES, CNPq e Fundação Araucária, agradeço pelo imprescindível suporte financeiro.

Aos funcionários do C-LABMU/UEPG, meu muito obrigado pelo suporte técnico oferecido, essencial ao desenvolvimento desta pesquisa.

Agradeço ainda a todos aqueles que, mesmo não citados nominalmente, fizeram parte desta trajetória. Muito obrigada!

Por fim, mas não menos importante, agradeço ao Criador do Universo, pois sem Ele nada seria possível.

"Deixem que o futuro diga a verdade e avalie cada um de acordo com o seu trabalho e realizações. O presente pertence a eles, mas o futuro pelo qual eu sempre trabalhei pertence a mim."
(Nikola Tesla)

RESUMO

Esta dissertação apresenta a aplicação de redes neurais artificiais (RNAs) do tipo perceptrons multicamadas (MLP) para a análise e extração automática de parâmetros físicos de espectros Raman de vidros teluretos. Esses materiais, à base de TeO_2 , possuem propriedades ópticas e eletrônicas relevantes, porém apresentam espectros Raman complexos e sobrepostos em função de sua estrutura amorfa, o que torna os métodos tradicionais de deconvolução gaussiana trabalhosos e limitados, especialmente em condições de ruído. Para contornar essas limitações, foi construído um banco de dados teórico composto por curvas Raman simuladas com diferentes níveis de ruído (1%, 2% e 5%), utilizado no treinamento e na validação das redes neurais. Após o treinamento, as MLPs foram aplicadas a espectros Raman experimentais de vidros teluretos, permitindo a extração automática dos parâmetros de cada banda espectral, como posição central, largura e área, tradicionalmente obtidos por procedimentos iterativos de ajuste. Os resultados demonstraram que as redes neurais foram capazes de ajustar adequadamente os espectros Raman, apresentando alta correlação entre os valores teóricos e experimentais, mesmo em condições de elevado ruído, além de reduzir significativamente o tempo de análise. Dessa forma, o método proposto configura-se como uma alternativa robusta e eficiente aos métodos convencionais de deconvolução gaussiana, mostrando-se uma ferramenta promissora para o ajuste automático e a interpretação de espectros Raman, contribuindo para o avanço da aplicação de técnicas de inteligência artificial na caracterização de materiais vítreos.

Palavras-chave: Espectroscopia Raman; Aprendizado de Máquina; Ajuste Gaussiano; Análise Espectral; Materiais Vítreos.

ABSTRACT

This dissertation presents the application of artificial neural networks (ANNs), specifically Multilayer Perceptron (MLP) architectures, for the analysis and automatic extraction of physical parameters from Raman spectra of tellurite glasses. These TeO_2 -based materials exhibit relevant optical and electronic properties; however, due to their amorphous structure, they produce complex and overlapping Raman spectra, making traditional Gaussian deconvolution methods labor-intensive and limited, particularly under noisy conditions. To overcome these limitations, a theoretical database composed of simulated Raman curves with different noise levels (1%, 2%, and 5%) was constructed and used for training and validation of the neural networks. After training, the MLP models were applied to experimental Raman spectra of tellurite glasses, enabling the automatic extraction of spectral band parameters such as central position, width, and area, which are traditionally obtained through iterative fitting procedures. The results demonstrate that the neural networks are capable of accurately fitting Raman spectra, showing high correlation between theoretical and experimental values even at high noise levels, while significantly reducing the analysis time. Therefore, the proposed method represents a robust and efficient alternative to conventional Gaussian deconvolution techniques, demonstrating strong potential for the automatic fitting and interpretation of Raman spectra and contributing to the advancement of artificial intelligence applications in the characterization of glassy materials.

Keywords: Raman Spectroscopy; Machine Learning; Gaussian Fitting; Spectral Analysis; Glassy Materials.

LISTA DE FIGURAS

Figura 1 – a) Sólido cristalino apresentando estruturas periódicas a longo alcance; b) Vidro silicato sem estruturas periódicas de longo alcance.	24
Figura 2 – (a) Bipirâmide trigonal TeO_4 ; (b) Poliedro TeO_{3+1} ; (c) Pirâmide trigonal TeO_3	26
Figura 3 – Representação esquemática do espalhamento elástico e inelástico.	28
Figura 4 – Representação das intensidades do espalhamento elástico e inelástico Stokes e anti-Stokes.	30
Figura 5 – Exemplo de Espectro Raman de um vidro TLZ e unidades estruturais.	32
Figura 6 – Modelo de um neurônio, o perceptron.	36
Figura 7 – Cálculos do neurônio artificial.	39
Figura 8 – Demonstração da Função de Ativação Sigmoide.	41
Figura 9 – Diagrama de Comparação entre as Funções Sigmoide e Tanh.	43
Figura 10 – Função de ativação Ramp ou ReLu.	44
Figura 11 – Representação da rede Multiperceptron.	45
Figura 12 – Obtenção do espectro Raman das vibrações modelulares pela remoção da linha de base.	51
Figura 13 – Exemplo de ajuste por deconvolução gaussiana no Origin.	52
Figura 14 – Resultados dos ajustes pelo método clássico de deconvolução gaussiana dos espectros de intensidade normalizada dos Deslocamentos Raman pelo <i>Origin</i> para as oito amostras de TLZ dopadas com Cr_2O_3	54
Figura 15 – Exemplo de curvas de deslocamento Raman teóricas sem ruído geradas para base de dados para validação de redes neurais artificiais MLP.	58
Figura 16 – Representação gráfica dos resultados teóricos 1 e 2.	63
Figura 17 – Espectroscopia Raman obtidos.	64
Figura 18 – Resultados de validação e treinamento das redes neurais 2C-1 e 2C-5 na base 5.	66
Figura 19 – Resultados experimentais da rede na base 5 da rede MLP 2C-1.	68
Figura 20 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 de 10 mil curvas.	71
Figura 21 – Resultados do espectro predito pela rede 1C-11Ag treinada na base 5 para todos os resultados experimentais.	72
Figura 22 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com 1% de ruído.	76
Figura 23 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com 1% de ruído.	76
Figura 24 – Resultados da predição 1C-11Ag na base 5.	78
Figura 25 – Resultados da predição 2C-3Ag na base 5.	79

Figura 26 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com 2% de ruído.	80
Figura 27 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com 2% de ruído.	81
Figura 28 – Resultados da predição 1C-11Ag na base 5.	82
Figura 29 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com 5% de ruído.	84
Figura 30 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com 5% de ruído.	85
Figura 31 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com baseline na base de dados.	87
Figura 32 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com baseline na base de dados.	88
Figura 33 – Resultados experimentais da rede 1C-1Ag com baseline na base de dados na base 5.	89
Figura 34 – Resultados experimentais da rede 1C-2Ag com baseline na base de dados na base 5.	90
Figura 35 – Resultados experimentais da rede 1C-3Ag com baseline na base de dados na base 5.	92
Figura 36 – Resultados experimentais da rede 1C-11Ag com baseline na base de dados na base 5.	93
Figura 37 – Áreas das bandas obtidas pela: (a) Deconvolução gaussiana no <i>Origin</i> e (b) Predição usando a rede neural 1C-11Ag.	95
Figura 38 – Centros das bandas obtidas pela predição usando a rede neural 1C-11Ag na base de dados normalizada com 1% de ruído e linha de base.	96
Figura 39 – Áreas das bandas obtidas pela predição usando as cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.	98
Figura 40 – Centros das bandas obtidas pela predição usando as cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.	100
Figura 41 – Largura das bandas obtidas pela predição usando as cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.	101
Figura 42 – Resultados da predição 2C-2Ag na base 5.	115
Figura 43 – Resultados da predição 2C-3Ag na base 5.	116
Figura 44 – Resultados da predição 2C-4Ag na base 5.	117
Figura 45 – Resultados da predição 2C-5Ag na base 5.	118
Figura 46 – Resultados da predição 1C-1Ag na base 5.	119
Figura 47 – Resultados da predição 1C-2Ag na base 5.	121
Figura 48 – Resultados da predição 1C-3Ag na base 5.	122
Figura 49 – Resultados da predição 1C-4Ag na base 5.	123

Figura 50 – Resultados da predição 2C-3Ag na base 5.	124
Figura 51 – Resultados da predição 1C-1Ag na base 5.	126
Figura 52 – Resultados da predição 1C-2Ag na base 5.	128
Figura 53 – Resultados da predição 1C-3Ag na base 5.	129
Figura 54 – Resultados da predição 1C-4Ag na base 5.	130
Figura 55 – Resultados da predição 1C-1Ag na base 5.	132
Figura 56 – Resultados da predição 1C-2Ag na base 5.	134
Figura 57 – Resultados da predição 1C-3Ag na base 5.	135
Figura 58 – Resultados da predição 1C-4Ag na base 5.	136
Figura 59 – Resultados da predição 1C-11Ag na base 5.	138

LISTA DE TABELAS

Tabela 1 – História de pesquisa em Rede Neurais.	35
Tabela 2 – Nomenclatura e composição das amostras preparadas.	49
Tabela 3 – Intervalos da espectroscopia Raman obtidos das deconvoluções pelo <i>Origin</i>	53
Tabela 4 – Intervalos dos parâmetros utilizados para geração da base de dados para os convolução dos espectros de Deslocamento Raman sem normalização	56
Tabela 5 – Intervalos dos parâmetros utilizados para geração da base de dados para os convolução dos espectros de Deslocamento Raman normalizados	57
Tabela 6 – Parâmetros das funções gaussianas utilizadas na geração dos sinais teóricos.	63
Tabela 7 – Resultados do treinamento, validação e R^2 das redes sem normalização com duas camadas. Todas as redes com função de ativação Ramp (ReLU).	66
Tabela 8 – Erros percentuais dos resultados teóricos 1 das redes 2C-1 e 2C-5.	67
Tabela 9 – Erros percentuais dos parâmetros das gaussianas ajustadas para as redes 2C-1Ag, das amostras TLZCr-0.10 e TLZCr-0,75.	69
Tabela 10 – Resultados do treinamento, validação e R^2 das redes normalizadas com duas camadas para base de 10 mil curvas, todas com função de ativação Ramp.	69
Tabela 11 – Resultados do treinamento, validação e R^2 das redes normalizadas com duas camadas para base de 20 mil curvas, todas com função de ativação Ramp.	70
Tabela 12 – Erros percentuais dos resultados teóricos 1 normalizados das redes 1C-11Ag (melhor) e 2C-3Ag (pior).	71
Tabela 13 – Erros percentuais dos parâmetros das gaussianas ajustadas para as redes 1C- 11Ag, das amostras TLZCr-0.35 e TLZCr-075 incluindo centro, largura e área de cada banda.	73
Tabela 14 – Resultados do treinamento, validação e R^2 das redes normalizadas e 1% de ruído com duas camadas para base de 10 mil curvas, todas com função de ativação Ramp.	74
Tabela 15 – Resultados do treinamento, validação e R^2 das redes normalizadas e 1% de ruído com duas camadas para base de 20 mil curvas, todas com função de ativação Ramp.	75
Tabela 16 – Erros percentuais dos resultados teóricos 1 normalizados com 1% de ruído das redes 1C-11Ag e 2C-3Ag.	77
Tabela 17 – Parâmetros das gaussianas ajustadas para as redes 1C-11Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda.	77
Tabela 18 – Erros percentuais dos resultados teóricos 1 normalizados com 2% de ruído das redes 1C-11Ag e 2C-3Ag.	81
Tabela 19 – Parâmetros das gaussianas ajustadas para a rede 1C-11Ag nas amostras TLZ- Base e TLZCr-2, incluindo centro, largura e área de cada banda.	83

Tabela 20 – Erros percentuais dos resultados teóricos 1 normalizados com 5% de ruído das redes 1C-11Ag e 2C-3Ag.	86
Tabela 21 – Erros percentuais dos resultados teóricos 1 normalizados com baseline e 1% de ruído das redes 1C-11Ag e 2C-3Ag.	88
Tabela 22 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0.35 e TLZCr-050 incluindo centro, largura e área de cada banda. . .	89
Tabela 23 – Parâmetros das gaussianas ajustadas para as redes 1C-2Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda. . .	91
Tabela 24 – Parâmetros das gaussianas ajustadas para as redes 1C-2Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda. . .	92
Tabela 25 – Parâmetros das gaussianas ajustadas para as redes 1C-11Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda. . .	94
Tabela 26 – Valores médios e respectivos erros para as áreas das gaussianas das amostras analisadas obtidos pelas cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.	97
Tabela 27 – Valores médios e respectivos erros para os centros das gaussianas das amostras analisadas obtidos pelas cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.	99
Tabela 28 – Valores médios e respectivos erros para as larguras das gaussianas das amostras analisadas obtidos pelas cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.	100
Tabela 29 – Parâmetros das gaussianas ajustadas para as redes 2C-2Ag, das amostras TLZCr-01 e TLZCr-0,50 incluindo centro, largura e área de cada banda. . .	115
Tabela 30 – Parâmetros das gaussianas ajustadas para as redes 2C-3Ag, das amostras TLZCr-0,25 e TLZCr-1 incluindo centro, largura e área de cada banda. . . .	116
Tabela 31 – Parâmetros das gaussianas ajustadas para as redes 2C-4Ag, das amostras TLZCr-0,35 e TLZCr-2 incluindo centro, largura e área de cada banda. . . .	117
Tabela 32 – Parâmetros das gaussianas ajustadas para as redes 2C-4Ag, das amostras TLZCr-0,1 e TLZCr-75 incluindo centro, largura e área de cada banda. . . .	118
Tabela 33 – Resultados da rede 1C-1Ag na base 5.	119
Tabela 34 – Resultados da rede 1C-2Ag na base 5.	121
Tabela 35 – Parâmetros das gaussianas ajustadas para as redes 1C-3Ag, das amostras TLZBase e TLZCr-050 incluindo centro, largura e área de cada banda. . . .	122
Tabela 36 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0.35 e TLZCr-075 incluindo centro, largura e área de cada banda. . .	123
Tabela 37 – Parâmetros das gaussianas ajustadas para as redes 2C-3Ag, das amostras TLZCr-0.35 e TLZCr-075 incluindo centro, largura e área de cada banda. . .	125
Tabela 38 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0.1 e TLZCr-035 incluindo centro, largura e área de cada banda. . .	126

Tabela 39 – Parâmetros das gaussianas ajustadas para as redes 1C-2Ag, das amostras TLZCr-0.50 e TLZCr-2 incluindo centro, largura e área de cada banda. . . .	127
Tabela 40 – Parâmetros das gaussianas ajustadas para as redes 1C-3Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda. . . .	129
Tabela 41 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda. . . .	130
Tabela 42 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0,75 e TLZCr-2 incluindo centro, largura e área de cada banda. . . .	132
Tabela 43 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0,50 e TLZCr-2 incluindo centro, largura e área de cada banda. . . .	134
Tabela 44 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0,50 e TLZCr-0,75 incluindo centro, largura e área de cada banda. . . .	135
Tabela 45 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0,1 e TLZCr-2 incluindo centro, largura e área de cada banda. . . .	137
Tabela 46 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0,1 e TLZCr-2 incluindo centro, largura e área de cada banda. . . .	137
Tabela 47 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZBase e TLZCr-2 incluindo centro, largura e área de cada banda. . . .	138

LISTA DE ABREVIações

BO Oxigênios ligantes. 33

CLS Método dos mínimos quadrados clássicos. 17

CNNs Convolutional neural network. 18

FTIR Espectroscopia de Infravermelho por Transformada de Fourier. 16

Li₂O Óxido de Lítio. 32

LSR líquidos super-resfriados. 22

MLP Multilayer Perceptron. 20

MSE Erro quadratico médio. 57

NBO Oxigênios não ligantes. 31

TeO₂ Dióxido de Telúrio. 18

TeO₃ Unidade estrutural piramidal trigonal. 18

TeO₄ Unidade estrutural de bipirâmide trigonal. 18

TeO₃₊₁ Unidade estrutural poliédrica de transição (bipirâmide trigonal distorcida). 18

TLZ Te-Li-Zn. 18, 20

UV-Vis Ultravioleta-Visível. 16

ZnO Óxido de Zinco. 18

SUMÁRIO

1	INTRODUÇÃO	16
2	FUNDAMENTAÇÃO TEÓRICA	21
2.1	Vidros Teluretos	21
2.1.1	Definição do termo vidro	21
2.1.2	Formação dos vidros teluretos	23
2.1.3	Estrutura dos vidros teluretos	25
2.1.4	Vibrações moleculares espectroscopia Raman	27
2.1.5	Espectros Raman dos vidros teluretos	31
2.2	Redes Neurais	35
2.2.1	Breve histórico	35
2.2.2	Modelo de um neurônio artificial - Perceptron	38
2.3	Neurônio artificial	39
2.3.1	Funções de ativação	40
2.4	Redes Neurais Multilayer Perceptron	44
2.4.1	Definição da camada densa	45
2.4.2	Empilhamento (entrada–ocultas–saída)	46
2.4.3	Fluxo feed-forward e por que isso resolve mapeamentos não lineares de espectros	46
2.5	Definição de <i>Overfitting</i>	47
3	MATERIAIS E MÉTODOS	49
3.1	Preparação das amostras	49
3.2	Espectroscopia Raman	50
3.3	Pré-processamento - remoção da baseline e normalização pela área	50
3.3.1	Deconvolução gaussiana	51
3.4	Redes Neurais Artificiais	55
3.4.1	Base de dados sem normalização	56
3.4.2	Base de dados com normalização	57
3.4.3	Bases de dados normalizadas com adição de ruído	58
3.4.4	Arquiteturas das redes neurais	59
3.4.5	Nomenclatura das redes neurais	61
3.5	Curvas de teste teóricas	62
4	RESULTADOS E DISCUSSÕES	64
4.1	Espectros Raman	64
4.2	Predição de parâmetros e espectros Raman sem normalização	65
4.2.1	Resultados das curvas experimentais da rede 2C-1	67
4.3	Predição de parâmetros e espectros Raman com normalização	69

4.3.1 Resultados das curvas experimentais da rede 1C-11Ag base 5	72
4.4 Redes neurais aplicadas a espectros normalizados com adição de 1% de ruído	74
4.4.1 Resultados das curvas experimentais da rede 1C-11Ag com 1% de ruído	77
4.4.2 Resultados das curvas experimentais da rede 2C-3Ag com 1% de ruído	79
4.5 Redes neurais aplicadas a espectros normalizados com adição de 2% de ruído	79
4.5.1 Resultados das curvas experimentais da rede 1C-11Ag com 2% de ruído	82
4.6 Redes neurais aplicadas a espectros normalizados com adição de 5% de ruído	84
4.7 Aplicação de redes neurais com 1% em espectros normalizados com linha de base	86
4.7.1 1C-1Ag base 5 – 1C-1Ag base 6 com baseline %	88
4.7.2 1C-2Ag base 5 – 1C-2Ag base 6 com baseline	90
4.7.3 1C-3Ag base 5 – 1C-3Ag base 6 com baseline %	91
4.7.4 1C-11Ag base 5 – 1C-11Ag base 6 com baseline %	93
4.8 Comportamento das áreas	94
4.8.1 Comportamento dos centros pela 1C-11Ag	95
4.8.2 Comportamento das áreas para todas as redes	96
4.8.3 Comportamento dos centros para todas as redes	98
4.8.4 Comportamento das larguras para todas as redes	99
5 CONCLUSÃO	102
6 PERSPECTIVAS FUTURAS	105
REFERÊNCIAS	106
APÊNDICE A – RESULTADOS DA PREDIÇÃO DOS RESULTADOS EXPERIMENTAIS PELAS REDES SEM NORMALIZAÇÃO	115
APÊNDICE B – RESULTADOS DA PREDIÇÃO EXPERIMENTAL PELAS REDES COM NORMALIZAÇÃO	119
APÊNDICE C – RESULTADOS DAS DA PREDIÇÃO DOS RESULTADOS EXPERIMENTAIS PELAS REDES COM NORMALIZAÇÃO E 1% DE RUÍDO	126
APÊNDICE D – RESULTADOS DA PREDIÇÃO EXPERIMENTAL PELAS REDES COM NORMALIZAÇÃO E 2% DE RUÍDO	132

1 INTRODUÇÃO

As Redes Neurais Artificiais (RNAs) são modelos computacionais inspirados no funcionamento do cérebro humano, compostos por unidades interconectadas denominadas neurônios artificiais, capazes de aprender padrões e realizar tarefas complexas de reconhecimento, classificação e predição a partir de dados experimentais (1, 2). Nos últimos anos, o uso dessas redes tem se expandido significativamente no campo das técnicas espectroscópicas, impulsionando avanços na análise qualitativa e quantitativa de materiais (3). As RNAs têm sido aplicadas em técnicas espectroscópicas como Raman, FTIR e UV-Vis, que permitem identificar grupos funcionais, avaliar ligações químicas, determinar composições moleculares, analisar modificações estruturais e as alterações dos níveis eletrônicos dos materiais (4–8). Porém, a interpretação dos espectros obtidos pode ser complexa, sobretudo pela sobreposição de bandas e pela presença de ruídos e interações vibracionais que dificultam a análise quantitativa e qualitativa precisa (9).

Apesar de amplamente utilizada, a extração de informações físico-químicas a partir de espectros Raman com bandas sobrepostas ainda enfrenta um gargalo metodológico importante: a deconvolução por múltiplas Gaussianas é um problema inverso sensível às condições iniciais e à correlação entre parâmetros. Em vidros teluretos, nos quais é comum a presença de cinco ou mais bandas na mesma faixa espectral (10–12), o ajuste por múltiplos picos em *softwares* (por exemplo, *Fityk* e *Origin*) pode apresentar duas limitações centrais. Quando se deixam livres os três parâmetros principais de cada Gaussiana, centro, largura e área, é frequente obter ajustes com coeficiente de determinação elevado (R^2 próximo de 1), porém com soluções não únicas e, em alguns casos, fisicamente inconsistentes, como larguras irreais, deslocamentos incompatíveis e até áreas negativas, que não têm significado físico nesse contexto. Para contornar esses efeitos, a prática usual exige um ajuste manual iterativo: fixam-se centros e larguras, ajusta-se parcialmente, libera-se um parâmetro por vez e repete-se o procedimento até alcançar uma solução estável e interpretável. Esse processo, além de demorado, depende fortemente da perícia do operador, exige treinamento e reduz a reprodutibilidade dos resultados entre diferentes analistas; adicionalmente, a fixação de parâmetros pode introduzir viés e mascarar fenômenos reais de interesse — como alargamentos de banda ou deslocamentos de centro associados a mudanças estruturais no vidro.

Diante dessa lacuna, esta dissertação propõe investigar o uso de Redes Neurais Perceptron Multicamadas (MLP) como uma alternativa para automatizar a deconvolução de espectros Raman de vidros teluretos, predizendo simultaneamente os parâmetros Gaussianos de múltiplas bandas (centro, largura e área), totalizando, por exemplo, 15 parâmetros para cinco bandas. A hipótese é que, uma vez treinada de forma adequada, a MLP pode fornecer estimativas consistentes e reprodutíveis, reduzindo a interferência humana no ajuste e preservando variações fisicamente relevantes do espectro, contribuindo para análises mais rápidas, padronizadas e interpretáveis em conjuntos extensos de dados experimentais.

Nesse contexto, as redes neurais artificiais surgem como uma ferramenta poderosa para reconhecimento de padrões e extração automática de características em espectros complexos. Modelos como as redes perceptrons multicamadas (MLPs) têm demonstrado excelente desempenho em tarefas de identificação de componentes (13–15). Essas redes são capazes de aprender relações não lineares entre as variáveis espectrais, superando as restrições impostas pelos métodos lineares tradicionais.

As redes perceptron, especialmente em suas versões multicamadas (MLPs), têm sido amplamente utilizadas para resolver problemas básicos e intermediários na análise de espectros Raman (16). Embora seu uso mais comum envolva tarefas de classificação, nas quais cada variável de entrada representa as intensidades dos picos espectrais e a saída corresponde à classe ou estado do material analisado (17), sua aplicação vai além dessa função. Devido à complexidade e sobreposição dos sinais espectrais, as MLPs mostram-se particularmente adequadas para lidar com relações não lineares entre as variáveis (18). Assim, essas redes permitem modelar e distinguir padrões em dados Raman, possibilitando avanços na identificação de componentes químicos, na diferenciação de fases de materiais e em diagnósticos biomédicos de caráter inicial.

Estudos envolvendo RNAs aplicadas à espectroscopia Raman de materiais vítreos são relativamente escassos, mas vêm crescendo nos últimos anos. Historicamente, as primeiras investigações da técnica unindo Raman e redes neurais ocorreram nos anos 1990, ainda que não focadas diretamente em vidros, período em que a capacidade de processamento computacional começou a permitir o tratamento de espectros complexos. Os primeiros estudos buscavam demonstrar que as RNAs poderiam superar limitações dos métodos quimiométricos clássicos, como o método dos mínimos quadrados clássicos (CLS) e a regressão linear, especialmente em sistemas com ruído, sobreposição de bandas e não linearidades. Em um dos primeiros trabalhos, soluções aquosas de neurotransmissores de pequenas moléculas, mostrou que MLPs treinadas por retropropagação eram capazes de identificar com alta acurácia os perfis Raman e, mais que isso, evidenciou um aspecto metodológico crucial: funções de transferência senoidais nas camadas ocultas aceleravam o treinamento e melhoravam a discriminação entre espectros de compostos estruturalmente próximos (19). No ano seguinte os mesmos autores aplicaram redes neurais multicamadas (MLP) à análise de misturas de neurotransmissores a partir de espectros Raman (20). O objetivo era identificar e quantificar componentes como dopamina, serotonina e acetilcolina em soluções aquosas. O estudo comparou o desempenho das RNAs com o método CLS e mostrou que as redes eram mais tolerantes a ruído, deslocamentos espectrais e distorções, evidenciando seu potencial para decomposição de espectros complexos e análises quantitativas em sistemas biológicos.

Outro trabalho pioneiro estendeu o uso de MLPs ao problema de classificação de materiais lignocelulósicos, acoplando espectroscopia FT-Raman no infravermelho próximo à computação neural para a extração de características e o reconhecimento de diferentes tipos de madeira

(21). Ao tratar um ambiente espectral mais complexo do que soluções químicas relativamente simples, esse estudo demonstrou a viabilidade do uso de MLPs na análise de dados Raman para tarefas de classificação, evidenciando a capacidade dessas redes de identificar padrões vibracionais relevantes associados à diferenciação entre classes naturais.

Um exemplo recente é o trabalho de Wu et al. (2025) (22), que propôs um algoritmo de espectroscopia Raman baseado em CNNs e MLPs (RS-MLP), capaz de realizar análises qualitativas e quantitativas de misturas químicas com alta precisão ($R^2 > 0.99$), indicando que o modelo demonstra boas condições.

Mais recentemente, observa-se um crescimento expressivo na integração entre espectroscopia Raman e métodos de aprendizado de máquina. Revisões da literatura mostram que arquiteturas modernas, incluindo redes profundas, métodos híbridos (físico-informados) e algoritmos de pré-processamento, têm melhorado significativamente a extração de características e a interpretação de espectros (23). Além disso, modelos especializados para sinais Raman, como redes do tipo RamanNet, têm sido desenvolvidos para lidar com ruído, sobreposição de bandas e variações (24).

A espectroscopia Raman é uma das principais ferramentas experimentais utilizadas para investigar a estrutura e as interações internas em materiais vítreos (10). Por meio do espalhamento inelástico da luz, a técnica fornece informações sobre as vibrações fundamentais da rede, permitindo identificar unidades estruturais (25, 26). A sensibilidade da técnica a pequenas variações estruturais torna o Raman essencial para estudar os efeitos de dopantes e modificadores na rede vítrea, correlacionando estrutura, ligações químicas e propriedades ópticas. Na literatura, encontramos cinco bandas características para o TeO₂ puro, sendo elas: 450, 615, 668, 732 e 782 cm⁻¹ (27, 28), as quais se referem aos modos vibracionais do dióxido de telúrio. No estudo de Novatski et al. (2019) (10) sobre vidros do sistema TLZ TeO₂-Li₂O-ZnO TLZ, a espectroscopia Raman foi empregada para compreender como a substituição gradual de TeO₂ por ZnO influencia a estrutura da rede. Foram observadas cinco bandas características, centradas em aproximadamente 450, 611, 660, 730 e 780 cm⁻¹, associadas às vibrações das unidades estruturais TeO₄, TeO₃₊₁, TeO₃.

No contexto específico de vidros, o uso de técnicas de aprendizado de máquina tem se mostrado cada vez mais promissor para prever propriedades a partir de sua composição e estrutura. Um dos primeiros trabalhos nessa direção foi o estudo de Dreyfus (2003) (29), que empregou uma rede neural MLP para estimar a temperatura líquida de misturas de óxidos formadoras de vidro, demonstrando a capacidade dessas redes de modelar relações não lineares complexas. Posteriormente, Brauer et al. (2007) (30) aplicaram redes neurais feed-forward à previsão da durabilidade química de vidros fosfatos, relacionando composição e solubilidade em água. O modelo mostrou boa capacidade preditiva, evidenciando o potencial das RNAs na modelagem de propriedades difíceis de descrever por métodos tradicionais.

Krishnan e colaboradores (2018) (31) utilizaram redes neurais artificiais para prever pa-

rômetros de durabilidade de vidros silicatos a partir de espectros Raman combinados com outras técnicas estruturais. Mesmo com um conjunto de dados relativamente reduzido, o modelo apresentou bom desempenho preditivo, indicando a capacidade das RNAs de explorar correlações entre características extraídas dos espectros e propriedades macroscópicas dos materiais.

O uso de técnicas de aprendizado de máquina tem avançado rapidamente na ciência dos vidros, permitindo prever propriedades termodinâmicas e estruturais diretamente a partir da composição química ou de assinaturas espectroscópicas. Trabalhos pioneiros demonstraram que redes neurais artificiais podem prever propriedades como a temperatura de transição vítrea (T_g) com alta precisão (32). A ideia central dessas técnicas consiste em treinar modelos computacionais capazes de aprender relações não lineares entre dados de entrada e propriedades físicas conhecidas, ajustando iterativamente seus parâmetros internos de modo a minimizar o erro de previsão.

Esforços recentes incluem o uso de redes profundas multitarefa, capazes de prever simultaneamente múltiplas propriedades físico-químicas a partir de extensas bases de dados composicionais (33), bem como estudos que empregam técnicas de aprendizado de máquina para investigar correlações entre composição e propriedades em famílias de materiais, como vidros calcogenetos (34). Esses avanços evidenciam duas tendências principais: a expansão para modelos com maior capacidade de generalização, aptos a explorar amplos espaços composicionais, e a crescente incorporação da espectroscopia Raman como fonte de informação estrutural para modelos baseados em redes neurais.

Vale ressaltar que redes do tipo MLP apresentam melhor desempenho quando os dados espectrais são previamente normalizados e submetidos a etapas de pré-processamento destinadas à redução de ruído e interferências, como a fluorescência (35). Após esse processamento, as redes são treinadas com conjuntos de dados de referência, ajustando iterativamente seus parâmetros internos de forma a minimizar o erro entre as saídas previstas e os valores experimentais (36).

Entretanto, observa-se uma lacuna na literatura no que se refere à aplicação de redes neurais artificiais para o ajuste e a predição de parâmetros espectrais em vidros a partir de espectros Raman. Na maioria das investigações existentes, essas redes são empregadas predominantemente em tarefas de classificação. É importante destacar que vidros apresentam estrutura amorfa e desordenada, o que resulta em espectros Raman complexos e amplamente sobrepostos. No caso específico dos vidros teluretos, os espectros Raman situam-se, na maior parte das vezes, no intervalo de 300 a 1000 cm^{-1} (10).

Redes neurais artificiais geralmente requerem grandes volumes de dados para um treinamento eficaz (37–39). Diante desse cenário, o uso de redes do tipo Perceptron Multicamadas (perceptrons multicamadas, MLP) treinadas com conjuntos de dados gerados teoricamente, com base em características observadas experimentalmente, constitui uma estratégia para viabilizar a predição de parâmetros espectrais em situações nas quais a disponibilidade de dados experimentais é limitada.

Esta dissertação tem como objetivo aplicar Redes Neurais perceptrons multicamadas (MLP) para regressão e predição dos parâmetros de ajuste dos espectros de Deslocamento Raman dos vidros teluretos TLZ dopados com Cr_2O_3 .

Como objetivos específicos, pretende-se:

- Aplicar redes neurais perceptron multicamadas (MLP) em diferentes arquiteturas (número de camadas, neurônios, funções de ativação e tamanho da base de dados) para verificação da que melhor se aplica ao problema de regressão dos espectros Raman de vidros teluretos;
- Verificar a influência da normalização dos dados de treino e ajuste na predição dos parâmetros gaussianos de curva teórica e resultados experimentais;
- Analisar o efeito da adição de diferentes níveis de ruído e linha de base às curvas de treino sobre a predição do espectro e dos parâmetros de ajuste;

Dessa forma, esta dissertação está organizada, além desta Introdução, em quatro Capítulos e quatro Apêndices, a saber:

No capítulo 2 (Fundamentação Teórica) são apresentadas a definição do que é um vidro, a formação e estrutura dos vidros teluretos, uma breve explicação sobre o efeito Raman e sua aplicação na caracterização da estrutura dos vidros teluretos, um breve histórico sobre RNAs e uma base sobre as redes perceptron multicamadas (MLP).

No capítulo 3 são apresentados os Materiais e Métodos, iniciando com a preparação das amostras de vidros TLZ dopados com Cr_2O_3 . Após isso são apresentados a caracterização pela espectroscopia Raman e ajuste convencional do espectro por deconvolução gaussiana utilizando o software Origin. Esse passo é importante para conhecer os intervalos esperados para posterior criação das bases de dados (de 500 a 20000 curvas) que serão utilizadas para treinamento das redes neurais utilizando o Wolfram Mathematica, cuja metodologia é apresentada ao final do capítulo.

No Capítulo 4 são apresentados os principais resultados obtidos. Inicialmente, discutem-se os testes de predição de parâmetros teóricos e experimentais com redes MLP, variando-se o número de camadas e de neurônios, bem como o tamanho da base de dados, sem considerar ruído e nem normalização de todos os parâmetros. Na sequência, analisa-se de forma sistemática a influência da normalização das variáveis de entrada e de saída. Por fim, investiga-se o efeito da adição de diferentes níveis de ruído ao espectro Raman e de uma linha de base sobre o desempenho das redes.

Por fim, no capítulo 5 são apresentadas as Conclusões com base nos resultados obtidos e a indicação de trabalhos futuros. Os apêndices trazem resultados das predições dos resultados experimentais das demais redes neurais não incluídas no capítulo Resultados.

2 FUNDAMENTAÇÃO TEÓRICA

Neste capítulo, são apresentados os fundamentos teóricos que dão suporte ao desenvolvimento desta dissertação. Inicialmente, discutem-se os conceitos essenciais relacionados aos vidros teluretos, abordando a definição do termo vidro, os principais critérios de formação vítrea, a estrutura local das unidades formadoras à base de TeO_2 e as contribuições da espectroscopia Raman para a caracterização desses materiais. Em seguida, são introduzidos os conceitos de redes neurais artificiais, com ênfase nas redes do tipo perceptrons multicamadas (MLP), contemplando desde o modelo de neurônio artificial e as funções de ativação até a organização em camadas e os aspectos relacionados a sobreajuste (overfitting). Essa fundamentação teórica é importante para compreender as etapas metodológicas adotadas, bem como a interpretação dos resultados obtidos ao se aplicar redes neurais à predição dos parâmetros gaussianos dos espectros Raman de vidros teluretos.

2.1 Vidros Teluretos

Os vidros teluretos constituem uma importante classe de materiais vítreos baseados em TeO_2 , amplamente investigados devido às suas propriedades ópticas e eletrônicas distintivas, como altos índices de refração, boa transmitância no infravermelho, baixa temperatura de fusão e elevada estabilidade térmica. Essas características tornam esses vidros particularmente atrativos para aplicações em fotônica, guias de onda, dispositivos optoeletrônicos e sensores (10, 27).

Nas subseções a seguir discutem-se primeiramente, a definição do termo vidro e os critérios clássicos de formação vítrea, com ênfase nos sistemas óxidos. Em seguida, aborda-se a formação específica dos vidros teluretos e sua estrutura em termos das diferentes unidades de coordenação do telúrio na rede (TeO_4 , TeO_{3+1} e TeO_3). Por fim, apresentam-se os princípios da espectroscopia Raman aplicada a esses materiais, destacando-se as vibrações moleculares características e as principais bandas observadas nos espectros Raman de vidros teluretos, que serão fundamentais para a análise desenvolvida nos capítulos seguintes.

2.1.1 Definição do termo vidro

Cientificamente, para se caracterizar um material como vidro é necessário considerar seu estado sólido, sua composição, sua estrutura e a sua forma de obtenção. Os primeiros estudos nesse sentido foram realizados por Michael Faraday, no ano de 1830. Uma definição proposta por ele é de que “vidros são materiais que mais aparentam uma solução de diferentes

substâncias do que um composto em si” (40). Em 1921, Lebedev propôs a Hipótese do Cristalito para tentar explicar a estrutura dos vidros e, segundo essa hipótese, “um vidro era como um fundido comum constituído de cristais altamente dispersos” (41). Com o passar do tempo tanto as técnicas de obtenção de vidros quanto as caracterização foram aprimoradas, assim, diferentes definições foram propostas para esses materiais. Em geral, essas definições concordem que o vidro é um sólido não cristalino, caracterizado pela ausência de simetria e de periodicidade de longo alcance, e que apresenta transição vítrea.

Uma das definições mais completas foi proposta por Shelby em 1997 (42), na qual o vidro é definido como um sólido amorfo com uma ausência completa de ordem e periodicidade, exibindo uma região de transição vítrea. Qualquer material inorgânico, orgânico ou metal formado por qualquer técnica que exibe um fenômeno de transição vítrea. O primeiro problema dessa definição seria que ainda é preciso elucidar qual é o fenômeno da transição vítrea. Outra questão é o fato de que as definições anteriores rotulam o vidro como sólido, enquanto sua estrutura pode ser melhor comparada a um líquido super-resfriado. Os vidros se assemelham aos sólidos em uma "escala de tempo humana", possuem rigidez e elasticidade mecânicas e podem fraturar como um sólido. No entanto, eles fluem a qualquer temperatura sob a ação da gravidade, relaxando em direção a um estado líquido super-resfriado (43). Além disso, sugere-se que, após um longo tratamento térmico ou no limite de tempo infinitamente longo, o vidro se cristalize, solidificando.

A definição mais recente aceita é a dada por Zanotto (43)

Vidro é um estado da matéria condensada fora do equilíbrio termodinâmico, não cristalino, que exibe uma transição vítrea. As estruturas dos vidros são semelhantes às dos seus líquidos super-resfriados *LSR* e relaxam espontaneamente em direção ao estado de *LSR*. Seu destino final, para tempos infinitamente longos, é cristalizar.

Portanto, podemos assumir que os vidros são sólidos amorfos, que possuem ausência de simetria e periodicidade de longo alcance. Entretanto, é necessário lembrar que ainda não existe um modelo realista capaz de descrever a natureza das interações de um sistema vítreo.

O primeiro trabalho sobre a estrutura dos vidros foi desenvolvido no artigo *The Atomic Arrangement in Glass* (O Arranjo Atômico em Vidros), publicado por Zacharysen (44), o qual afirmou que a base conceitual para a formação de vidro é por fusão/resfriamento. Para explicar como ocorre a formação de um material considerado vidro, duas teorias foram desenvolvidas: a estrutural e a cinética (45). A estrutural busca explicar a capacidade de formação dos vidros a partir de aspectos químicos e estruturais; já a cinética, explica a formação vítrea como a capacidade de qualquer material fundido em não cristalizar-se. Os vidros são obtidos a partir do método de fusão/resfriamento (46).

2.1.2 Formação dos vidros teluretos

Embora existam muitos critérios para a formação estrutural, é difícil encontrar o que pode ser aplicado em todas as substâncias formadoras de vidro. Nesse sentido, abordaremos os vidros óxidos.

- ✓ **Critério de Goldschmidt (1926)** (47): na busca por critérios que explicassem a vitrificação da fórmula estequiométrica R_nO_m , supunha que a formação era a razão entre o raio iônico do cátion, R , e o do íon do oxigênio que se encontravam entre 0,2 e 0,4. Seguindo as condições clássicas para a estrutura iônica, a razão R está correlacionada com o número da coordenação do cátion central. A configuração proposta implica em coordenação tetraédrica (42).
- ✓ **Regra de Zachariasen (1932)** (44): uma análise completa mostrou que o critério de Goldschmidt era limitado, pois o BeO satisfaz a razão R e não pode vitrificar. Zachariasen reconsiderou o problema e, por meio de um discurso prático, determinou algumas regras de uma gama importante sobre investigação vítrea. Sua análise tem a seguinte afirmação: A força de ligação interatômica e as propriedades mecânicas devem ser semelhantes tanto no vidro quanto nos cristais. Embora os cristais apresentem uma rede tridimensional bem definida, o vidro exibe uma característica difusa no espectro de raios X, que descreve uma rede periódica, porém não simétrica.

Considerando que a formação dos vidros óxidos depende de diferentes coordenações, Zachariasen estabeleceu algumas regras para os arranjos estruturais, as quais estão descritas abaixo (44):

De acordo com as regras de Zachariasen, o oxigênio liga-se preferencialmente, a no máximo, dois cátions formadores de rede. Além disso, esses cátions devem apresentar baixa coordenação com átomos de oxigênio, formando poliedros que compartilham apenas vértices, evitando o compartilhamento de arestas ou faces. Para garantir a formação de uma rede vítrea tridimensional contínua, cada poliedro deve compartilhar pelo menos três de seus vértices com poliedros vizinhos.(48);

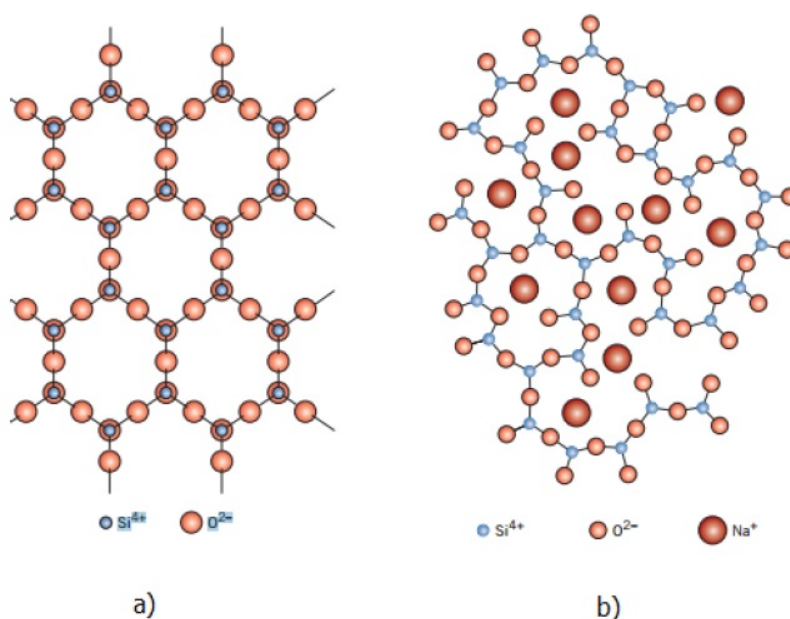
Zachariasen (44), generalizou suas regras para vidros óxidos em geral, acrescentando que um vidro óxido (complexo) pode formado se (48):

A estrutura vítrea é favorecida quando a amostra contém uma alta porcentagem de cátions circundados por átomos de oxigênio na forma de tetraedros ou triângulos. Esses tetraedros ou triângulos devem compartilhar apenas vértices entre si, evitando o compartilhamento de faces ou de arestas. Além disso, parte dos átomos de oxigênio deve estar ligada a apenas dois cátions da rede, sem formar ligações adicionais com outros cátions.

Isso significa que os vidros óxidos devem conter uma quantidade mínima de cátions capazes de formar óxidos que vitrificam, responsáveis pela formação da rede vítrea. Os demais cátions podem atuar como modificadores ou intermediários da rede, sem impedir a vitrificação. Esse mesmo autor denominou os cátions B^{3+} , Si^{4+} , Ge^{4+} , P^{5+} , As^{5+} , P^{3+} , Sb^{3+} , Nb^{5+} e Ta^{5+} , como formadores de vidro, os quais, quando associados aos átomos de oxigênio, constituem a rede vítrea (48).

A Figura 53a) ilustra a representação esquemática de Zachariasen para um cristal, apresentando uma representação bidimensional da estrutura que possui configuração ordenada com cadeias simétricas e periodicidade de longo alcance. Enquanto a Figura 53b) traz a representação bidimensional de um vidro que tem uma estrutura desordenada; em outras palavras, há ausência de simetria ou nenhuma periodicidade na rede (49).

Figura 1 – a) Sólido cristalino apresentando estruturas periódicas a longo alcance; b) Vidro silicato sem estruturas periódicas de longo alcance.



Fonte: JR, W. D. C.; RETHWISCH, D. G. Fundamentals of materials science and engineering: an integrated approach. [S.l.]: John Wiley & Sons, 2012

Devido a essas condições, definem-se como óxidos formadores de rede aqueles que participam diretamente da estrutura vítrea, estabelecendo a conectividade da rede. Por outro lado, os óxidos modificadores não integram a rede covalente contínua, atuando na sua despolimerização. Nesse contexto, o termo modificador justifica-se pelo efeito observado na estrutura da rede quando um óxido não formador é adicionado, como ocorre com a incorporação de (por exemplo, o Na_2O) na sílica (SiO_2). Os oxigênios introduzidos participarão da rede, provocando a quebra de um dado número de ligações. Para cada molécula de Na_2O introduzida na matriz, uma ligação de $Si-O-Si$ é quebrada e o oxigênio suplementar formará um par $Si-O^-$. As cargas

negativas de oxigênios serão compensadas pela presença de um par de cátions vizinhos de Na^+ , possibilitando a neutralidade do conjunto. Isso ocorre porque, durante a fusão, a rede primitiva SiO_2 fica quebrada – despolimerizada (48).

Zachariasen (1932) (50), observa que as condições apresentadas anteriormente não são suficientes para que um vidro se forme. É necessário que o material fundido seja resfriado adequadamente, ou seja, ele já antecipava as teorias de formação vítrea baseadas na cinética do processo.

- ✓ **Necessidade de ligações mistas** (51): Smekal (1951) propôs a presença de ligações de caráter parcialmente iônico e parcialmente covalente. No entanto, uma ligação puramente covalente, como a $\text{C}=\text{C}$ do diamante, provoca ligações fixas muito rígidas, o que seria tão inapropriado quanto as ligações iônicas. Dessa forma, para a formação de um vidro necessitamos de óxidos como SiO_2 e de B_2O_3 para os quais as ligações são parcialmente iônicas e parcialmente covalentes.
- ✓ **Critério de eletronegatividade** (52): Stanworth (1952) sugeriu uma correlação entre o caráter covalente da ligação e a tendência de um óxido à formação vítrea. A diferença de eletronegatividade entre os dois elementos possibilita determinar a porcentagem de ionicidade da ligação. Quanto menor for a diferença, mais covalente seria a ligação. Como a eletronegatividade do oxigênio é 3,5, os cátions podem ser agrupados como:
 - * eletronegatividade entre 1,8 e 2,1, são formadores de rede; ($\text{Si}=1,8$);
 - * eletronegatividade entre 1,5 e 1,8, são intermediários; ($\text{Al}=1,5$);
 - * eletronegatividade menor que 1,5, são modificadores de rede; ($\text{Mg}=1,2$ e $\text{Ca}=1,0$).

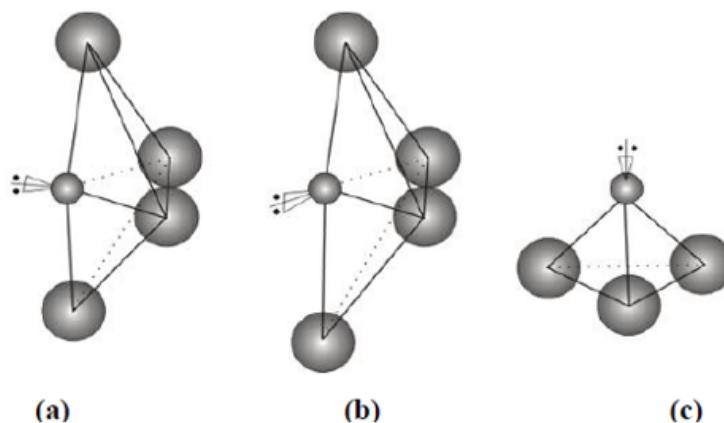
2.1.3 Estrutura dos vidros teluretos

Os vidros à base de óxido de telúrio possuem aplicações importantes e propriedades de grande interesse, como baixos pontos de fusão, boa transmitância no infravermelho, altas constantes dielétricas, durabilidade química elevada, além de altos índices de refração e boa estabilidade térmica.

Conforme comentado anteriormente, vidros que possuem como óxido formador de rede o TeO_2 são classificados como teluretos. Na forma cristalina, o TeO_2 se apresenta basicamente em duas maneiras estáveis (53), com duas modificações, $\alpha\text{-TeO}_2$ (paratelureto) $\beta\text{-TeO}_2$ (telureto) (54). Normalmente, um vidro telureto é formado por diferentes polimorfos de TeO_2 , sendo estes, uma distorção do $\alpha\text{-TeO}_2$ (paratelureto).

No entanto, o telúrio se apresenta de três formas na estrutura TeO_n , como ilustrado na Figura 2, onde n pode assumir os valores de 4, 3 ou um intermediário entre 3 e 4 (descrito com $3 + 1$).

Figura 2 – (a) Bipirâmide trigonal TeO_4 ; (b) Poliedro TeO_{3+1} ; (c) Pirâmide trigonal TeO_3 .



Fonte: Gunha, Jaqueline Valeski. "Caracterização estrutural e cinética de cristalização do sistema vítreo $\text{TeO}_2\text{-Na}_2\text{O-TiO}_2$."(2020) (55)

Na coordenação de número quatro, o átomo de telúrio está no centro da bipirâmide trigonal. Podemos observar essa configuração na Figura 2-a), onde há 2 átomos de oxigênio axiais e 2 átomos de oxigênio equatoriais, sendo que uma terceira posição equatorial é ocupada por um par de elétrons isolado (56). A distância da ligação Te-O é maior nas ligações axiais, sendo esta da ordem de $2,08 \text{ \AA}$, enquanto as ligações equatoriais são da ordem de $1,90 \text{ \AA}$. O par de elétrons livres está disposto no plano formado pelos oxigênios equatoriais (11).

Na Figura 2-b), o átomo de telúrio com coordenação 3+1 é um poliedro assimétrico, onde uma ligação axial Te-O é menor e a outra é mais alongada, que é representado do $\beta\text{-TeO}_2$, estrutura bipirâmide trigonal. Quando a quarta ligação axial, Te-O_{ax} , se alonga significativamente, sua interação torna-se fraca e pode ser considerada desprezível do ponto de vista estrutural. Como o átomo de telúrio está ligado a três átomos de oxigênio, e a um quarto átomo a uma distância superior a $2,20 \text{ \AA}$, diz-se que essa unidade é do tipo TeO_{3+1} . Essas unidades surgem devido à incorporação de óxidos modificadores na composição vítrea, fazendo com que as unidades de TeO_4 passem à unidades de TeO_{3+1} (11,57).

O agrupamento desses átomos está representado na Figura 2-c), que descreve uma pirâmide trigonal TeO_3 , onde, no topo, está localizado o átomo de telúrio, e a base é formada pelos átomos de oxigênio. A estrutura é formada quando a distância da ligação Te-O_{ax} na coordenação de TeO_{3+1} é maior que $2,58 \text{ \AA}$. As outras três ligações curtas Te-O apresentam uma distância em torno de $1,88 \text{ \AA}$ (58). Essas três ligações são curtas Te-O ; ou seja, podemos afirmar que a coordenação resulta em uma evolução progressiva: $\text{TeO}_4 \rightarrow \text{TeO}_{3+1} \rightarrow \text{TeO}_3$ (59). Do ponto de vista da elaboração do vidro, essa evolução estrutural está associada à incorporação de óxidos modificadores na composição vítrea. A adição desses óxidos promove a quebra de

ligações Te–O–Te, gerando oxigênios não ponte e reduzindo a conectividade da rede. Como consequência, ocorre a conversão progressiva de unidades.

2.1.4 Vibrações moleculares espectroscopia Raman

A técnica de Espectroscopia Raman deve-se ao espalhamento inelástico de uma radiação monocromática que incide sobre uma molécula. A atividade que resulta nesse comportamento está ligada à variação do momento de dipolo induzido na molécula pelo campo elétrico da radiação incidente (60).

Ao incidir um feixe monocromático na amostra, a luz interage com a molécula e, na maioria dos casos, essa radiação é espalhada elasticamente, não mudando sua frequência inicial; esse processo é chamado de espalhamento Rayleigh. Contudo, um pequeno número de fótons é espalhado inelasticamente, mudando sua frequência inicial, e é conhecido como espalhamento Raman (61), na figura 3 temos a representação dos espalhamentos Raman. O deslocamento para frequências menores que a radiação incidente é o espalhamento Stokes, e o deslocamento para frequências maiores é o espalhamento anti-Stokes, (62). Essa mudança no comprimento de onda é causada pela interação da luz com os modos vibracionais do material (61).

A energia do fóton espalhado será $h\nu_0 + E_v$ e, conseqüentemente, a frequência do fóton espalhado pode ser expressa como

$$\nu_r = \nu_0 + \frac{E_v}{h}, \quad (2.1)$$

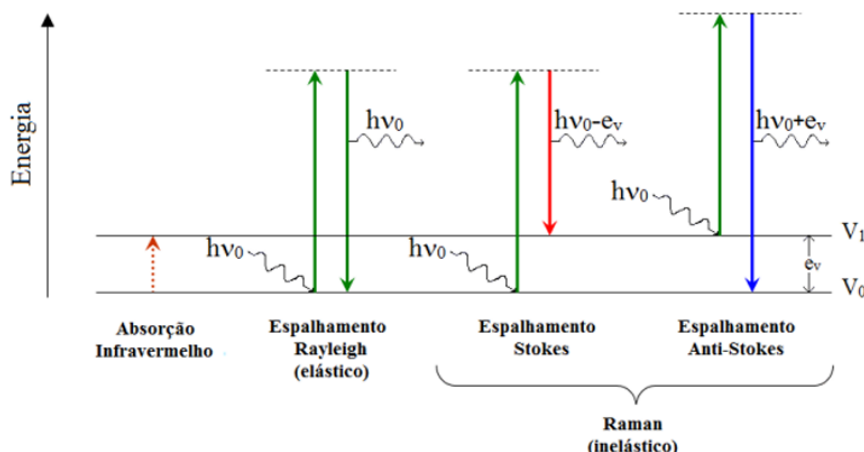
sendo, portanto, maior que a frequência do fóton incidente (61).

O espalhamento Rayleigh acontece quando um fóton de energia $h\nu$, onde h é a constante de Planck e ν é a frequência do fóton, interage com uma molécula no estado fundamental V_0 e a excita para um nível de energia mais alto V_1 . Após isso, a molécula decai para o mesmo nível de energia inicial. O estado inicial com uma frequência igual à do fóton incidente, pois não há mudança no estado vibracional da molécula.

No espalhamento Stokes, quando o fóton incidente interage com a molécula, ele transfere parte de sua energia para a rede, resultando na criação de um fônon — uma excitação vibracional no material (63). Ele excita a molécula para um estado vibracional mais alto V_1 , o que significa que parte da energia do fóton é transferida para a molécula. A energia do fóton ao ser espalhado será $h\nu_0 - e_v$, neste caso, a frequência do fóton espalhado será menor que a frequência do fóton incidente.

Embora o espalhamento Raman envolva uma mudança na frequência do fóton espalhado, o processo é descrito, do ponto de vista físico, como uma interação óptica linear. Isso ocorre porque o espalhamento Raman resulta da modulação da polarizabilidade da molécula pelo

Figura 3 – Representação esquemática do espalhamento elástico e inelástico.



Fonte: ISMAIL, K. Fabrication and characterisation of SERS substrates through photo-deposition of gold nanoparticles. Tese (Doutorado), 2015.

campo elétrico da radiação incidente, sendo permitido apenas para modos vibracionais que induzem uma variação da polarizabilidade molecular.

O espalhamento Raman anti-Stokes ocorre quando uma molécula ou rede cristalina já está em um estado vibracional excitado V_1 , antes da interação com um fóton incidente. Nesse processo, o fóton incidente interage com a molécula ou rede que possui um fônon e ele pode aniquilar esse fônon, absorvendo sua energia (64). Este processo eleva temporariamente a molécula a um nível de energia mais alto V_1 . Quando a molécula decai de volta ao seu estado fundamental V_0 , o fóton espalhado possui uma frequência maior do que a do fóton incidente, porque carrega não apenas a energia original, mas também a energia extra proveniente do fônon. A energia do fóton ao ser espalhado será $h\nu_0 + e_v$, e a frequência do fóton espalhado será maior que a frequência do fóton incidente (61).

O efeito Raman é observado quando a radiação incidente interage com o dipolo elétrico de uma molécula, o campo elétrico da radiação pode induzir ou alterar o momento de dipolo da molécula. De acordo com a teoria clássica, a magnitude do momento de dipolo está diretamente relacionada à intensidade do campo elétrico incidente segundo a equação:

$$P = \alpha \cdot E \quad (2.2)$$

onde P é o momento de dipolo elétrico induzido, E é o campo elétrico da radiação incidente e α é a polarizabilidade da molécula. A polarizabilidade pode ser expressa como Uma expansão em séries de Taylor com relação à coordenada de vibração q , onde q descreve o deslocamento das moléculas em relação à sua posição de equilíbrio durante a vibração:

$$\alpha = \alpha_0 + \frac{d\alpha}{dq_0} \cdot q + \dots \quad (2.3)$$

E escrevendo:

$$E = E_0 \cdot \cos(\nu_0 \cdot t) \quad (2.4)$$

$$q = q_0 \cdot \cos(\nu_v \cdot t) \quad (2.5)$$

onde ν_v é a frequência do modo de vibração e ν_0 a frequência da radiação incidente. É possível calcular o dipolo induzido:

$$P = \alpha \cdot E = \alpha_0 \cdot E_0 \cdot \cos(\nu_0 \cdot t) + \frac{d\alpha}{dq_0} \cdot q_0 \cdot E_0 \cdot \cos(2\pi \cdot \nu_0 \cdot t) \cdot \cos(2\pi \cdot \nu_v \cdot t) \quad (2.6)$$

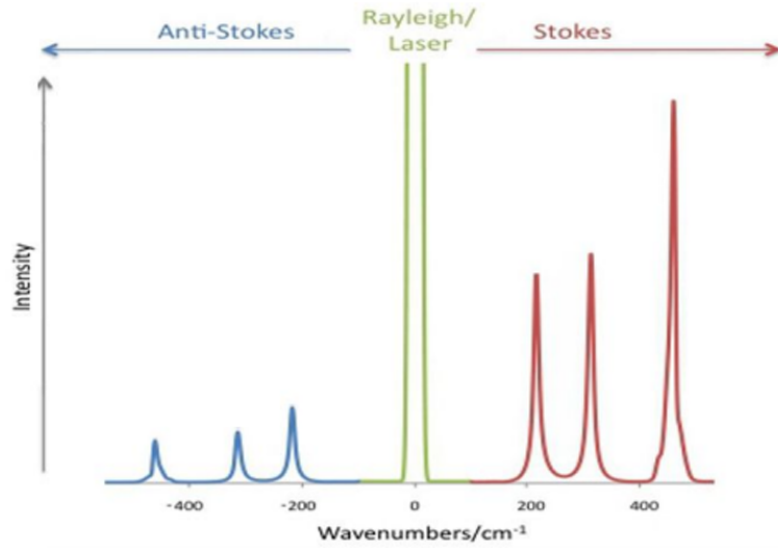
Usando a identidade trigonométrica $A \cos B = \frac{1}{2} [\cos(A + B) + \cos(A - B)]$, obtém-se:

$$P = \alpha_0 \cdot E_0 \cdot \cos(\nu_0 \cdot t) + \frac{1}{2} \frac{d\alpha}{dq_0} \cdot q_0 \cdot E_0 \cdot [\cos 2\pi[(\nu_0 + \nu_v)t] + \cos 2\pi[(\nu_0 - \nu_v)t]] \quad (2.7)$$

O momento de dipolo induzido em uma molécula pode ser decomposto em três termos distintos. O primeiro termo, que é função de α_0 , está relacionado ao espalhamento Rayleigh, caracterizando o espalhamento elástico da radiação incidente (63). Os dois termos subsequentes envolvem derivadas de α em relação à coordenada normal de vibração e apresentam oscilações associadas a $(\nu_0 - \nu_v)$ e $(\nu_0 + \nu_v)$. Esses últimos termos representam, respectivamente, os fenômenos de espalhamento Raman Stokes e anti-Stokes, revelando as mudanças na frequência vibracional da molécula devido à interação com a radiação. Para que o segundo termo contribua, a derivada de α deve ser diferente de zero, ou seja, é necessário que a polarizabilidade da molécula varie com a vibração (65–67).

O modelo clássico demonstra uma eficácia na predição das frequências no espectro Raman. No entanto, sua precisão é limitada quando se trata da relação de intensidades entre os espalhamentos Raman Stokes e anti-Stokes. De acordo com a teoria clássica, esses dois tipos de espalhamento deveriam apresentar a mesma intensidade; contudo, observa-se experimentalmente que as linhas correspondentes ao espalhamento Stokes exibem uma intensidade superior às associadas ao espalhamento anti-Stokes (63). Os espalhamentos Stokes e anti-Stokes envolvem a interação de fótons com modos vibracionais de uma molécula pela criação e aniquilação de fônons, e a existência de fônons necessários para o processo anti-Stokes depende da temperatura e é descrita pela distribuição de Bose-Einstein (67). Em condições normais de temperatura, a maioria das moléculas ou fônons em um material estão no estado fundamental de energia vibracional. É esperado que a maioria dos espectros Raman possuam apenas linhas Stokes, como demonstrado na Figura 4, onde as linhas são muito mais intensas.

Figura 4 – Representação das intensidades do espalhamento elástico e inelástico Stokes e anti-Stokes.



Fonte: ISMAIL, K. Fabrication and characterisation of SERS substrates through photo-deposition of gold nanoparticles. Tese (Doutorado), 2015.

A relação entre as intensidades anti-Stokes e Stokes pode ser obtida como: (62, 68):

$$\frac{I_{Anti-Stoke}}{I_{Stoke}} = \left(\frac{\nu_0 + \nu_v}{\nu_0 - \nu_v} \right)^4 \cdot \exp \left(-\frac{e_v}{kT} \right) \quad (2.8)$$

Sendo a intensidade da banda Raman dada por:

$$I_r \propto \nu^4 \cdot I_0 \cdot N \cdot \left(\frac{\partial \alpha}{\partial Q} \right)^2 \quad (2.9)$$

na qual,

- I_0 é a intensidade do laser incidente;
- N é o número de moléculas espalhadas em um determinado estado;
- ν é a frequência do laser excitante;
- α é a polarizabilidade das moléculas;
- Q é a amplitude vibracional;
- $\left(\frac{\partial \alpha}{\partial Q} \right)$ = Seção transversal Raman.

A intensidade do sinal Raman depende fortemente da frequência do laser, pois a aceleração de uma carga induzida pela luz é proporcional ao quadrado da frequência da luz ($a \propto \nu^2$). Como a intensidade da radiação emitida é proporcional ao quadrado da aceleração ($I_r \propto a^2$), a

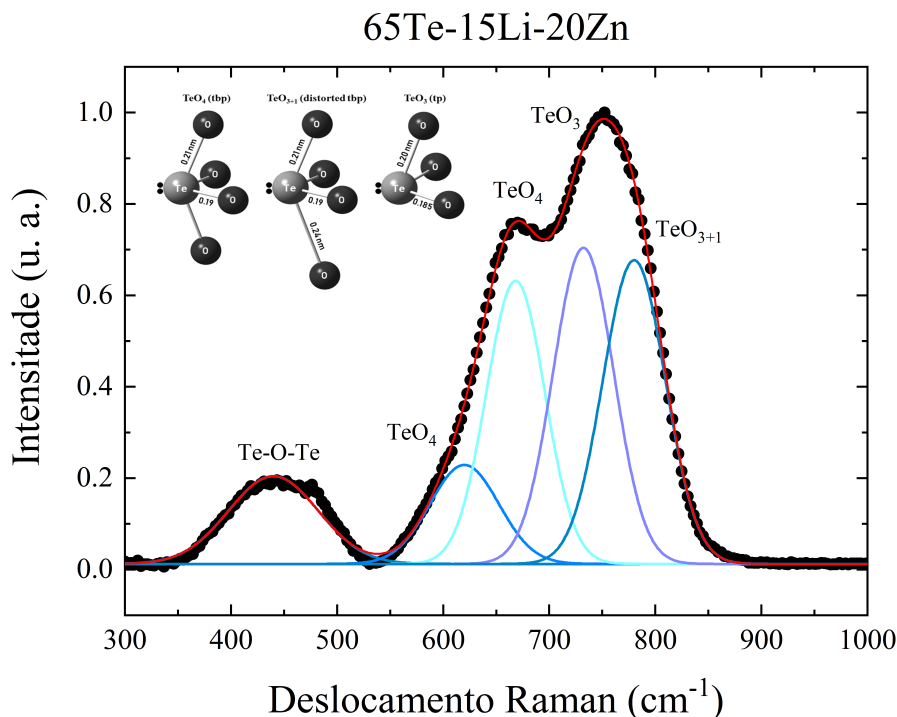
intensidade do espalhamento Raman acaba sendo proporcional à quarta potência da frequência do laser (63). Assim, se a seção transversal Raman e a intensidade do laser incidente permanecerem constantes, a intensidade da banda Raman (I_r) é diretamente proporcional à concentração da amostra, desde que mantidas constantes as condições experimentais e na ausência de efeitos significativos de reabsorção ou fluorescência. Ou seja, o sinal depende da concentração e também da potência do fluxo do laser incidente, e apenas as vibrações moleculares que causam uma mudança na polarizabilidade são Raman ativas (63, 69). Apenas uma pequena porção de fótons, aproximadamente um em 10^5 a 10^7 fótons, sofre dispersão de maneira inelástica, isto constitui uma pequena fração da radiação eletromagnética que se propaga com energia distinta da dispersão Rayleigh. A seção transversal Raman situa-se em torno de $\approx 10^{-32}$ a 10^{-29} cm², tornando o sinal Raman inerentemente fraco. Em contraste, a fluorescência (que pode competir como um processo alternativo) possui uma seção transversal de $\approx 10^{-17}$ cm², e, portanto, quando presente, pode sobrepor o sinal Raman (63, 70).

2.1.5 Espectros Raman dos vidros teluretos

Nos vidros do sistema TeO₂–Li₂O–ZnO (TLZ), a região mais importante do espectro Raman para análise estrutural está entre aproximadamente 300 e 1000 cm⁻¹. Nesse intervalo, os espectros podem ser descritos por cinco bandas principais, centradas em torno de 450, 611, 660, 730 e 780 cm⁻¹, como apresentado por Em Novatski et al. (2019) em (10). Um exemplo de espectro Raman para uma amostra de 65Te-15Li-20Zn é apresentado na Figura 5.

Estudos clássicos de TeO₂ puro e de vidros mostram que essas bandas estão ligadas às diferentes unidades de coordenação do telúrio na rede vítrea, principalmente TeO₄ em geometria de bipirâmide trigonal, TeO₃₊₁ e TeO₃, com oxigênios não ligantes NBO. (71–74)

Figura 5 – Exemplo de Espectro Raman de um vidro TLZ e unidades estruturais.



Fonte: (10), Novatski et al. (2019).

A primeira banda encontra-se usualmente centrada em torno de $440\text{--}460\text{ cm}^{-1}$, como relatado em estudos de vidros TeO_2 e sistemas MO--TeO_2 , bem como em trabalhos mais recentes sobre os vidros TLZ e $\text{TeO}_2\text{--Na}_2\text{O--TiO}_2$ (10, 71–75). Esta banda de baixa frequência corresponde, de forma geral, aos modos de flexão e estiramento das ligações Te--O--Te em unidades TeO_4 em geometria de bipirâmide trigonal, formadas por poliedros de telúrio que compartilham vértices. Em Novatski et al. (2019) (10), a banda análoga é modelada em torno de 450 cm^{-1} e está associada à conectividade da rede telureta: à medida que óxidos modificadores como Li_2O e, sobretudo, ZnO , são adicionados, a área desta banda diminui, indicando a quebra progressiva de ligações Te--O--Te e a conversão de unidades mais conectadas em unidades com oxigênios não ligantes (NBO). Assim, pequenas variações na posição podem ocorrer, mas o efeito mais marcante é a redução da intensidade relativa, sugerindo uma diminuição na conectividade da rede vítrea.

A segunda banda, na região de $600\text{--}620\text{ cm}^{-1}$, é associada aos modos de estiramento antissimétrico Te--O--Te de uma rede contínua composta por unidades TeO_4 em geometria de bipirâmide trigonal. Em vidros teluretos binários MO--TeO_2 , Sekiya et al. observam essa banda em torno de 610 cm^{-1} e a atribuem explicitamente à vibração da rede contínua formada por TeO_4 tbp (73). Para o TLZ, a banda é centrada em 611 cm^{-1} em Novatski et al. (2019) (10) e é atribuída principalmente ao estiramento antissimétrico de uma rede contínua composta por unidades TeO_4 com todos os quatro oxigênios ligados. Ela funciona, portanto, como um segundo

marcador de uma rede mais polimerizada, na qual o telúrio permanece majoritariamente em coordenação quatro, com predominância de oxigênios ligados (BO). Os resultados de Novatski et al. (2019) (10) mostram que, com o aumento da fração de ZnO no sistema TLZ, a área desta banda também diminui, reforçando a interpretação de que a introdução de modificadores converte gradualmente TeO_4 em unidades menos coordenadas (TeO_{3+1} e TeO_3) e prepara o cenário para o surgimento de bandas de mais alta frequência associadas a NBO.

A terceira banda, localizada na região de $650\text{--}670\text{ cm}^{-1}$, é amplamente reportada em vidros teluretos. Em sistemas binários MO--TeO_2 , Sekiya et al. descrevem um pico intenso em torno de 660 cm^{-1} e o atribuem a vibrações de estiramento assimétrico Te--O--Te envolvendo ligações entre um Te--O curto e um Te--O longo em unidades TeO_4 em bipirâmide trigonal, nas quais o oxigênio atua como oxigênio ponte (BO) entre dois átomos de telúrio (73). Para o sistema TLZ, entretanto, o modo em 660 cm^{-1} é interpretado por Novatski et al. (10) como uma vibração de estiramento assimétrico $\text{Te}_{\text{ax}}\text{--O}_{\text{eq}}\text{--Te}$ associada a unidades TeO_4 completamente ligadas (unidades com quatro BO), de modo que essa banda funciona como um indicador direto da fração de TeO_4 na rede. Consequentemente, a diminuição da área dessa banda com o aumento da concentração de ZnO no TLZ reflete a conversão gradual dessas unidades altamente conectadas em TeO_{3+1} e TeO_3 , acompanhada do aumento relativo das bandas de mais alta frequência associadas a NBO.

A quarta banda, na região de $720\text{--}740\text{ cm}^{-1}$, é reportada em vidros teluretos associada a modos de estiramento Te--O^- envolvendo oxigênios não ligantes (NBO). Em vidros MO--TeO_2 , Sekiya et al. observam um pico intenso em torno de 730 cm^{-1} e o atribuem explicitamente ao estiramento entre o telúrio e átomos de oxigênio não ligantes em unidades parcialmente despolimerizadas, do tipo TeO_{3+1} (73). Nessas estruturas, o poliedro TeO_{3+1} corresponde a um telúrio tetracoordenado com três ligações Te--O mais curtas e um quarto oxigênio ligado por uma ligação Te--O mais longa, identificado como NBO (73); assim, a banda em torno de 730 cm^{-1} é dominada pelo estiramento da ligação Te--O^- terminal (NBO), enquanto os oxigênios ligantes BO contribuem principalmente para as bandas de menor frequência. No sistema TLZ, Novatski et al. (10) ajustam essa banda com centro em 730 cm^{-1} e a associam às unidades $Q_{T,1}^3$ (TeO_{3+1}), nas quais a presença de NBO indica uma rede parcialmente despolimerizada. Do ponto de vista estrutural, essa banda está associada a mudanças na conectividade da rede vítrea, indicando a transição gradual entre uma rede predominantemente formada por unidades TeO_4 sem oxigênios não ligantes e uma rede na qual parte dos poliedros de telúrio passa a apresentar oxigênios não ligantes (NBOs). Os resultados de Novatski et al. mostram que a área dessa banda aumenta com a elevação da concentração de ZnO no TLZ, enquanto as bandas ligadas a TeO_4 diminuem, o que é consistente com a interpretação de que o ZnO (e, de forma análoga, Li_2O) favorece a conversão de TeO_4 em TeO_{3+1} e TeO_3 e o aumento do número de NBO na rede vítrea.

A quinta banda, na região de $770\text{--}790\text{ cm}^{-1}$, é descrita na literatura de vidros teluretos como associada a modos de estiramento Te--O^- envolvendo oxigênios não ligantes (NBO) em

unidades fortemente despolimerizadas. Em vidros TeO_2 modificados, Sekiya et al. observam um pico em torno de 780 cm^{-1} e o relacionam tanto à vibração da rede contínua de TeO_4 quanto ao estiramento Te-O em poliedros TeO_{3+1} e TeO_3 (73). Trabalhos posteriores de Raman em vidros teluretos dopados com terras-raras refinam essa interpretação e atribuem a banda em $\sim 770\text{ cm}^{-1}$ especificamente ao modo de estiramento $(\text{Te-O})_{\text{NBO}}$ em pirâmides TeO_{3+1} e/ou TeO_3 , isto é, ao movimento do oxigênio terminal não ligante ligado a um único átomo de telúrio (12, 72, 76). Nessas unidades TeO_3 trigonais piramidais, o telúrio encontra-se em coordenação três, com pelo menos um NBO, de modo que a presença e o aumento de intensidade dessa banda são marcadores diretos de uma rede vítrea rica em NBO e com baixa conectividade. No sistema TLZ, Novatski et al. ajustam essa quinta banda com centro em 780 cm^{-1} e a atribuem aos modos de estiramento Te-O^- de unidades TeO_3 com oxigênios não ligantes (10). Os resultados mostram que a área dessa banda aumenta com a elevação da concentração de ZnO , enquanto as bandas associadas a TeO_4 diminuem, indicando a conversão progressiva de TeO_4 em unidades TeO_{3+1} e TeO_3 ricas em NBO.

De forma simplificada, as três bandas de menor frequência (450 , 611 e 660 cm^{-1}) estão associadas aos modos de flexão e estiramento Te-O-Te em unidades TeO_4 mais conectadas, típicas de uma rede rica em oxigênios ligados. Já as duas bandas de maior frequência (em torno de 730 e 780 cm^{-1}) estão ligadas a estiramentos Te-O^- em unidades TeO_{3+1} e TeO_3 , nas quais aparecem NBO e a rede se torna menos polimerizada. A literatura mostra que, conforme óxidos modificadores são adicionados à matriz telureta (como Li_2O , Na_2O , ZnO ou outros óxidos de metais de transição), ocorre uma conversão gradual de TeO_4 em TeO_{3+1} e TeO_3 , observada no aumento relativo das bandas de alta frequência e na diminuição das bandas ligadas às unidades mais conectadas (72, 77, 78).

Em composições TLZ dopadas (por exemplo, com Cr_2O_3 ou outros óxidos de metais de transição), espera-se que esses cátions influenciem determinadas unidades TeO_3 e de TeO_{3+1} ou promovam a geração adicional de NBO, reforçando as bandas de alta frequência (730 e 780 cm^{-1}) e, assim, tendo pequenas alterações entre $\text{TeO}_3/\text{TeO}_4$ e a conectividade global da rede vítrea.

Essa evolução estrutural não é apenas qualitativa. A análise das áreas das bandas deconvoluídas permite relacionar diretamente o conteúdo de NBO e a razão $\text{TeO}_3/\text{TeO}_4$ com propriedades macroscópicas como temperatura de transição vítrea, densidade, módulo elástico e índice de refração, como discutido em vários trabalhos (10, 77–79). Em geral, o aumento de NBO (refletido pelo crescimento das bandas em torno de 730 e 780 cm^{-1}) enfraquece a rigidez da rede, diminui a conectividade entre os poliedros de telúrio e modifica o comportamento térmico e óptico dos vidros. (76, 77)

2.2 Redes Neurais

Redes neurais artificiais podem ser definidas como estruturas computacionais projetadas com o objetivo de explorar os princípios organizacionais existentes nos sistemas neurais biológicos. Por sua natureza, o estudo de redes neurais é multidisciplinar, envolvendo pesquisadores de diversas áreas, como neurofisiologia, psicologia, física, computação e engenharia (80).

Nesta seção, são apresentados os conceitos fundamentais das redes neurais artificiais, abordando inicialmente seu desenvolvimento histórico e a motivação para sua criação. Em seguida, descreve-se o modelo matemático do neurônio artificial no contexto do perceptron, seguido das principais funções de ativação e da estrutura das redes MLP. Por fim, discutem-se os aspectos relacionados ao processo de treinamento, como a retropropagação do erro (*back-propagation*), a normalização de dados e o sobreajuste (*overfitting*).

2.2.1 Breve histórico

As redes neurais artificiais passaram por um interessante processo de evolução, marcado por um período de grande atividade, seguido por anos de estagnação nas pesquisas e pelo ressurgimento do interesse científico como consequência do desenvolvimento de novas tecnologias e fundamentos teóricos. A seguir, é apresentado um breve histórico da pesquisa em redes neurais, sendo enfatizados alguns resultados e conceitos considerados relevantes no desenvolvimento deste trabalho (80).

Alguns dos mais destacados pesquisadores envolvidos no estudo e na aplicação de redes neurais, nas últimas três décadas, estão relacionados na Tabela 1. Esta tabela é dividida de forma a ressaltar o período cronológico mais significativo na atividade científica de cada pesquisador na área de redes neurais e é considerada uma base no processo de sequenciamento histórico.

Tabela 1 – História de pesquisa em Rede Neurais.

1943	McCulloch e Pitts
1948	Wiener
1949	Hebb
1957	Rosenblatt
1958	Widrow e Hoff
1969	Minsky e Papert
1960-1980	Kohonen, Grossberg, Widrow, Anderson, Caianiello, Fukushima,
1974	Werbos
1982	Hopfield
1986	Rumelhart e McClelland

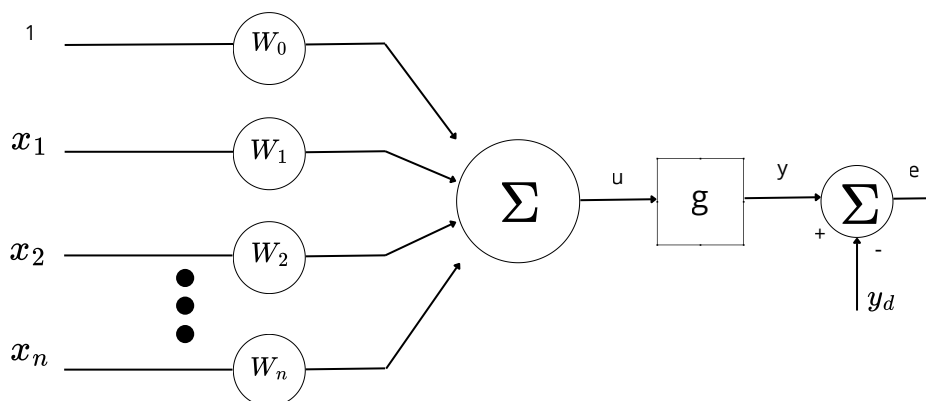
Em 1943, McCulloch e Pitts (81), dois pesquisadores da área da biologia, propuseram um modelo de neurônio como uma unidade de processamento binária e provaram que essas unidades são capazes de executar muitas operações que podem ser descritas como termos lógicos. Este modelo, apesar de muito simples, trouxe uma grande contribuição para as discussões sobre a construção dos primeiros computadores digitais, permitindo a criação dos primeiros modelos matemáticos de dispositivos artificiais que buscavam analogias biológicas.

Em 1948, N. Wiener (82) criou a palavra cibernética para descrever, de forma unificada, o controle e a comunicação nos organismos vivos e nas máquinas.

Em 1949, D. O. Hebb (83) apresentou uma hipótese a respeito de como a força das sinapses no cérebro se altera em resposta à experiência. Em particular, ele sugeriu que as conexões entre células que são ativadas ao mesmo tempo tendem a se fortalecer, enquanto as outras conexões tendem a se enfraquecer. Esta hipótese passou a influir decisivamente na evolução da teoria de aprendizagem em redes neurais artificiais.

Em 1957, F. Rosenblatt (84) introduziu uma nova abordagem para o problema de reconhecimento de padrões com o desenvolvimento do perceptron, cuja representação diagramática é apresentada na Figura 6. Rosenblatt também propôs um algoritmo para problemas linearmente separáveis. Por volta do mesmo período, B. Widrow (85) e seus colaboradores desenvolveram o adaline ("*Adaptive Linear Element*") para uso em processamento de sinais, o qual apresenta uma estrutura semelhante à do perceptron.

Figura 6 – Modelo de um neurônio, o perceptron.



Fonte: O autor.

Apesar do sucesso do perceptron e do adaline, a pesquisa em redes neurais passou gradualmente a se deparar com dois problemas fundamentais. Devido ao fato de a maior parte das pesquisas desenvolvidas até então ser de natureza heurística, o primeiro problema estava vinculado à carência na área, o que ocasionou a redução na produção de novas ideias. O segundo

problema, e talvez o de maior significado histórico, foi a expectativa exagerada criada pelos próprios pesquisadores dessa área, não acompanhada de resultados à altura, o que acelerou a queda dos financiamentos para a pesquisa. Mas foi após a publicação, em 1969, do livro "Perceptron" de autoria de M.L. Minsky e S.A. Papert (86), que as pesquisas na área de redes neurais sofreram uma retração significativa. Neste livro, são apresentados conceitos para analisar as capacidades adaptativas e computacionais de neurônios do tipo apresentado na Figura 6. Como resultado, os autores demonstraram que o perceptron, apesar de ser capaz de executar as operações booleanas AND e OR, não é capaz de executar outras operações elementares, como XOR (OU-exclusivo). Além disso, esses autores não acreditam que uma arquitetura adequada, juntamente com um algoritmo de ajuste de pesos, possa ser desenvolvida de forma a superar essa limitação. Após a publicação desses resultados, a maior parte dos pesquisadores da área de redes neurais passou a buscar alternativas dentro do campo da engenharia e, principalmente, da lógica matemática, que, na época, encontrava-se em franca expansão devido às grandes conquistas realizadas na área de computação.

Apesar deste êxodo generalizado, um número de pesquisadores continuou a trabalhar com redes neurais nos anos 70. Os nomes de T. Kohonem (Finlândia), S. Grossberg, B. Widrow e J. Anderson (Estados Unidos), E. Caianiello (Itália) e K. Fukushima (Japão) estão associados a este período.

Nos anos 80, muitos fatores contribuíam para o ressurgimento definitivo das pesquisas em redes neurais: neurofisiologistas foram adquirindo um maior conhecimento acerca do processamento de informações em organismos vivos; avanços tecnológicos tornaram disponível um maior potencial computacional a baixo custo, viabilizando ou facilitando simulações e testes com modelos de redes neurais; novas teorias para a implementação de algoritmos adaptativos foram desenvolvidas, permitindo a aplicação em sistemas reais.

Em 2024, a relevância das redes neurais artificiais alcançou seu mais alto reconhecimento científico. Nesse ano, John Hopfield e Geoffrey Hinton foram laureados com o Prêmio Nobel de Física por contribuições pioneiras que não apenas transformaram a compreensão teórica das redes neurais (87), mas também estabeleceram os fundamentos essenciais sobre os quais o machine learning moderno foi construído. Seu trabalho, iniciado décadas antes da atual explosão da inteligência artificial, mostrou como princípios da física podem orientar métodos de aprendizagem capazes de reconhecer padrões, reconstruir informações e extrair estruturas ocultas dos dados, abrindo caminho para muitas das tecnologias que hoje consideramos indispensáveis (87).

A consolidação de redes neurais como um novo ramo de pesquisa é evidenciada pela organização de numerosas conferências e pela fundação de revistas científicas devotadas exclusivamente ao tratamento deste tema, que recebeu a denominação de teoria conexionista.

2.2.2 Modelo de um neurônio artificial - Perceptron

O Perceptron foi introduzido por Frank Rosenblatt em 1958 (88) como um modelo matemático inspirado no funcionamento de neurônios biológicos, capaz de aprender a partir de exemplos e reconhecer padrões. A proposta de Rosenblatt (88) consistia em criar um sistema computacional que apresentasse princípios semelhantes aos mecanismos de percepção do cérebro humano, utilizando, para tanto, uma rede de unidades interconectadas que sofriam ajustes ao longo do processo de aprendizagem. Seu modelo tornou-se a primeira formulação formal de uma rede neural artificial treinável.

Rosenblatt descreveu o Perceptron como composto por três conjuntos principais de unidades: a camada sensorial (S-units), a camada associativa (A-units) e a camada de resposta (R-units). As unidades sensoriais são responsáveis por receber diretamente os estímulos externos; podem representar, por exemplo, pixels de uma imagem ou sinais provenientes de sensores. Essas unidades transmitem sinais para a camada associativa por meio de conexões com pesos ajustáveis, que desempenham um papel equivalente às sinapses nos neurônios biológicos. A camada associativa realiza o processamento interno, combinando e ponderando as entradas recebidas. Por fim, a camada de resposta atua como a unidade decisora, produzindo uma saída que representa a classificação do padrão apresentado. A arquitetura geral do perceptron é, portanto, a de um classificador linear baseado na soma ponderada das entradas (88).

Matematicamente, Rosenblatt explicou o funcionamento do Perceptron a partir do cálculo da soma ponderada dos sinais de entrada. Cada entrada é multiplicada por um peso sináptico, e o conjunto desses produtos é somado e comparado a um valor limiar. Caso essa soma ultrapasse o limiar, a unidade de resposta é ativada; caso contrário, permanece inativa. Essa operação pode ser descrita como uma função de ativação do tipo degrau, que transforma a saída da soma ponderada em uma decisão binária. A interpretação oferecida por Rosenblatt enfatiza o caráter probabilístico do modelo, que buscava aproximar o comportamento dos sistemas nervosos sem exigir uma representação exata de suas estruturas biológicas (88).

Um dos aspectos mais importantes da formulação de Rosenblatt é o mecanismo de aprendizagem, conhecido como Perceptron Learning Rule. Segundo o autor, o perceptron deve ser capaz de ajustar seus pesos sinápticos de modo a produzir as respostas desejadas a partir da apresentação de exemplos. O aprendizado ocorre sempre que o perceptron comete um erro. Se a rede deveria ativar-se, mas não ativou, os pesos associados às entradas responsáveis são aumentados. Se ativou quando não deveria, os pesos são reduzidos. Esse processo leva gradualmente a uma configuração de pesos capaz de separar linearmente os padrões apresentados. Rosenblatt (88) demonstrou matematicamente que, caso exista uma solução linearmente separável, o perceptron sempre será capaz de encontrá-la após um número finito de iterações.

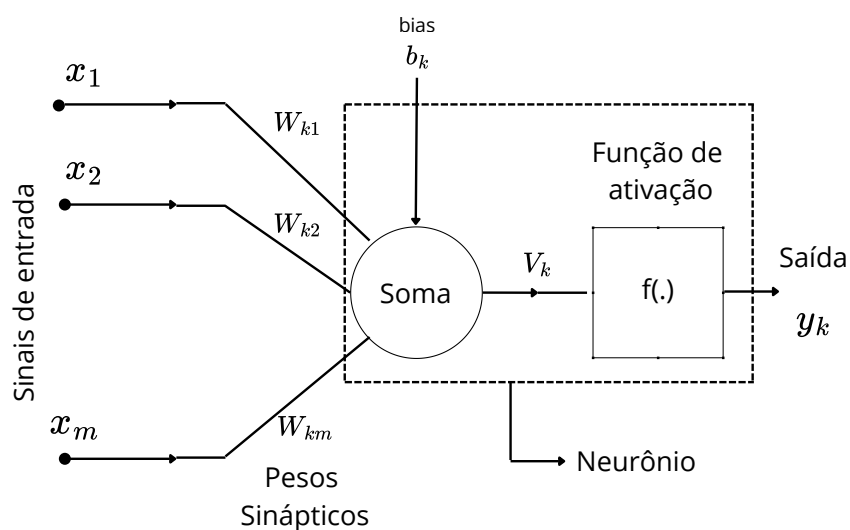
Rosenblatt (88) também discutiu as capacidades e limitações do modelo. Ele afirmou

que o perceptron é eficaz para problemas que podem ser separados por um hiperplano e apresenta diversos exemplos práticos, como reconhecimento de letras, discriminação de figuras e classificação geométrica simples. Por outro lado, o autor reconheceu implicitamente que padrões mais complexos exigiriam arquiteturas mais elaboradas. Esses limites foram formalizados anos depois por Minsky e Papert, que demonstraram que um perceptron de camada única não é capaz de resolver problemas que exigem separações não lineares, como o famoso problema XOR.

Apesar dessas limitações, o trabalho de Rosenblatt (88) estabeleceu as bases da área de redes neurais artificiais ao introduzir um dos primeiros modelos de aprendizado supervisionado, no qual os parâmetros sinápticos eram ajustados por meio de uma regra de aprendizado baseada em exemplos, fornecendo uma estrutura teórica que influenciou gerações posteriores de modelos neurais.

2.3 Neurônio artificial

Figura 7 – Cálculos do neurônio artificial.



Fonte: O autor.

O neurônio artificial é a unidade básica de processamento de dados de uma rede neural, cujo projeto baseia-se no modelo representado pelo diagrama de blocos da Figura 7 (89). Nela, podem ser identificados elementos fundamentais desse modelo, os quais são apresentados e descritos a seguir (89):

Conjunto de sinapses ou elos de conexão: são caracterizados por uma força própria, denominada *peso sináptico* (w_{kj}), o qual pode estar compreendido em um intervalo que abrange

valores positivos e negativos. Possuem a finalidade de armazenar informação, sendo modificados durante o processo de aprendizagem para alcançar os objetivos do projeto desejado;

Combinador linear ou somador: soma os sinais de entrada ponderados pelas respectivas sinapses do neurônio;

Função de ativação ($\varphi(\cdot)$): aplicada à saída do combinador linear para restringir a amplitude do sinal de saída do neurônio a um valor finito. Os três tipos básicos de funções de ativação mais comumente utilizados são: limiar, linear por partes e sigmoide. Em alguns casos, é desejável que as funções limiar e sigmoide sejam estendidas a valores negativos, passando, então, a serem chamadas de *função sinal* e *função tangente hiperbólica*;

Bias (b_k): parâmetro externo que pode ser comparado ao limiar do neurônio biológico e possui o efeito de aumentar ou diminuir a entrada da função de ativação, caso seja positivo ou negativo, respectivamente. Essa característica proporciona maior estabilidade à rede, aprimora sua capacidade de generalização e permite a consideração de um valor não relacionado aos fatores elencados pelos parâmetros de entrada (90).

Um neurônio k pode ser representado, matematicamente, por meio do seguinte par de equações:

$$u_k = \sum_{j=1}^m w_{kj} x_j \quad (2.10)$$

$$y_k = \varphi(u_k + b_k) \quad (2.11)$$

Onde:

- x_j = sinal de entrada correspondente à sinapse j ;
- w_{kj} = peso sináptico do neurônio k , na sinapse j ;
- u_k = saída do combinador linear do neurônio k ;
- b_k = *bias* do neurônio k ;
- $\varphi(\cdot)$ = função de ativação;
- y_k = sinal de saída do neurônio k .

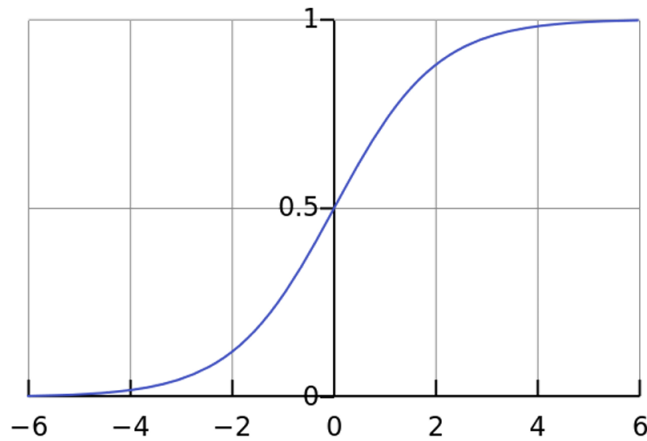
2.3.1 Funções de ativação

A função de ativação representa o efeito que a entrada interna e o estado atual de ativação exercem na definição do próximo estado de ativação da unidade. Quando propriedades dinâmicas

estão envolvidas na definição do estado de ativação, equações diferenciais (caso contínuo) ou a diferenças (caso discreto) são empregadas. Tendo em vista a simplicidade desejada para as unidades processadoras, geralmente se define seu estado de ativação como uma função algébrica da entrada interna atual, independente de valores passados do estado de ativação ou mesmo da entrada interna. Geralmente, esta função é monotonicamente não-decrescente e apresenta algum tipo de descontinuidade ou outros tipos especiais de não-linearidade.

Uma função matemática com as características de uma curva sigmoide é chamada de função sigmoide. As combinações da função sigmoide não são lineares, pois esse tipo de função de ativação é não linear. A Figura 8 representa uma demonstração dessa função.

Figura 8 – Demonstração da Função de Ativação Sigmoide.



Fonte: Szandala, T. (2021). Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. Wroclaw University of Science and Technology.

Assim, faz sentido empilhar camadas. Isso também se aplica a ativações não binárias. A função possui um gradiente suave, o que a torna adequada para redes rasas, como simulações de funções lógicas. A equação [Eq. 2.12] abaixo mostra como a função sigmoid pode ser usada em aprendizado profundo.

$$\sigma(x) = \frac{1}{1 + e^{-x}} \quad (2.12)$$

A partir da equação acima, pode-se observar que os valores de Y são muito íngremes quando os valores de X estão entre -2 e $+2$. Assim, haverá uma mudança significativa em Y para qualquer pequena variação em X dentro desse intervalo (91). Isso faz com que a função tenha uma tendência de levar os valores de Y para qualquer uma das extremidades da curva, aprimorando o classificador ao tornar mais claras as distinções de predição. Isso a torna ainda mais adequada para redes rasas, como simulações de funções lógicas (92).

$$\frac{d}{dx}\sigma(x) = \sigma(x)(1 - \sigma(x)) \quad (2.13)$$

Além disso, outra vantagem dessa função é que seu intervalo é $(0, 1)$, ao contrário da função linear, cujo intervalo é $(-\infty, \infty)$. Portanto, o intervalo de ativação é limitado, o que facilita evitar o “estouro” durante as ativações. Por isso, essa é a função de ativação mais amplamente utilizada e tem sido aplicada há muito tempo (93).

No entanto, pesquisas mostram que a função *sigmoid* não é perfeita, pois os valores de Y tendem a responder apenas levemente às mudanças nos valores de X nas extremidades da função. Isso implica que haverá um gradiente pequeno nesses pontos. Assim, isso leva ao problema do *vanishing gradient* (gradiente que desaparece), que ocorre nas partes quase horizontais da curva da função de ativação em ambos os lados. Consequentemente, o gradiente desaparece ou diminui de tamanho e não consegue provocar mudanças significativas, uma vez que seu valor é mínimo. Nesse caso, a rede pode aprender muito lentamente ou até parar de aprender; isso dependerá de como ela é usada, até que o gradiente se aproxime dos limites do ponto flutuante (93).

Outro problema da função *sigmoid* é sua centralidade não nula (94); em outras palavras, ela sempre produz um resultado positivo. Considere uma unidade *sigmoid* y com entradas x_1 e x_2 , ambas provenientes do neurônio *sigmoid*, com pesos w_1 e w_2 , respectivamente. Pela definição da função *sigmoid*, fica claro que x_1 e x_2 são sempre positivos, assim como y . Isso gera o sinal de erro que entra em y durante a retropropagação: o gradiente do erro será positivo tanto para w_1 quanto para w_2 , ou ambos negativos para w_1 e w_2 (93).

Liu et al. (2017). (95) demonstraram que, quando o vetor de pesos ótimo requer um aumento em w_1 e uma diminuição em w_2 , o procedimento de retropropagação não consegue ajustar ambos os pesos simultaneamente em um único passo, pois deve aumentá-los ou reduzi-los juntos. Assim, todo o processo de treinamento leva mais etapas para convergir do que levaria com uma função de ativação melhor.

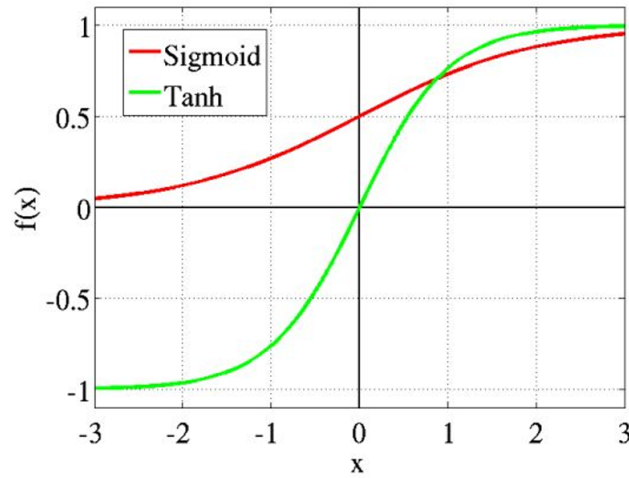
O travamento da rede neural no valor limite pode ocorrer se usarmos apenas a função de ativação sigmoide. Portanto, precisamos aplicar a função hiperbólica como alternativa, que também é conhecida como função tanh. A Figura 9 mostra uma comparação entre a função sigmoide e a função tanh (93).

Mollahosseini, Chan e Mahoor (96) sugeriram que a função hiperbólica é sigmoideal, assim como no caso das funções sigmoides. Contudo, para a função tanh, os valores de saída variam entre -1 e 1 , o que corresponde essencialmente à curva da função sigmoide estendida. Assim, entradas negativas da função hiperbólica serão mapeadas para saídas negativas, e valores de entrada próximos de zero também serão mapeados para valores de saída próximos de zero. Portanto, devido a essas características, a rede não fica presa durante o treinamento (93).

$$\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}} \quad (2.14)$$

Outra razão pela qual a função tanh é preferida à função sigmoide é que suas derivadas são significativamente maiores próximas de zero. No entanto, em grandes conjuntos de dados,

Figura 9 – Diagrama de Comparação entre as Funções Sigmoid e Tanh.



Fonte: Szandała, T. (2021). Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. Wroclaw University of Science and Technology.

um indivíduo normalmente tem dificuldade em encontrar rapidamente o mínimo local ou global, o que pode ser considerado uma característica benéfica da derivada da função tanh. Em outras palavras, espera-se que a função de custo seja minimizada mais rapidamente se a tanh for usada como função de ativação. Por fim, tanh reduz o problema de oscilação da função sigmoide, que produz apenas valores positivos (97).

$$\frac{d}{dx} \tanh(x) = 1 - \tanh^2(x) \quad (2.15)$$

Contudo, um caso frequentemente utilizado é a classificação binária. Assim, usar sigmoide ou tanh na camada final produz uma quantidade que pode ser escalada de 0 a 1. Dessa forma, é adequado utilizar uma função que exija probabilidades como entrada, como a entropia cruzada. Portanto, tanh, assim como a sigmoide, sofre com o problema do gradiente (93).

O retificador é uma função de ativação que atribui zero a valores negativos e o próprio valor quando este é maior que zero (98). Isso também é conhecido como função Ramp e é análogo à retificação de meia onda na engenharia elétrica. No entanto, a função de ativação foi introduzida pela primeira vez em uma rede dinâmica com fortes motivações biológicas e justificativas matemáticas.

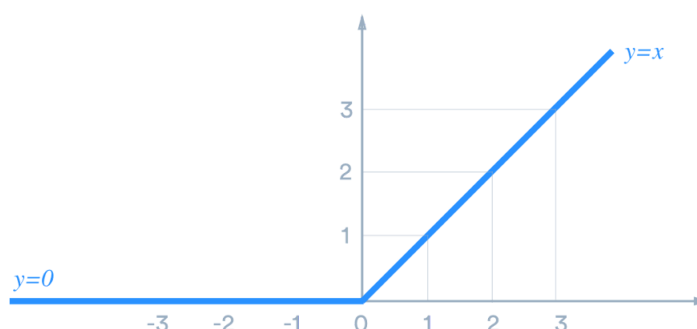
$$f(x) = x^+ = \max(0, x) \quad (2.16)$$

Além disso, foi comprovado pela primeira vez em 2011 que ela permite o treinamento mais eficiente de redes mais profundas quando comparada a funções de ativação amplamente usadas antes de 2011, como a sigmoide ou a tangente hiperbólica. Ping et al. (93) sugerem que

o retificador era a função de ativação mais popular para redes neurais profundas, de acordo com pesquisas realizadas em 2018.

Por fim, calcular o resultado da função e o gradiente é uma tarefa simples, pois as etapas de propagação direta e retropropagação são rápidas. Resultados de pesquisas mostram que a ReLU é seis vezes mais rápida do que outras funções de ativação bem conhecidas (99). Observa-se que, para problemas simples de regressão com redes neurais, a função tanh pode ser superior à ReLU (100).

Figura 10 – Função de ativação Ramp ou ReLu.



Fonte: Szandala, T. (2021). Review and Comparison of Commonly Used Activation Functions for Deep Neural Networks. Wroclaw University of Science and Technology.

Contudo, qualquer função aproximada usando ReLU indicado na figura 10 é sempre linear por partes. Além disso, são necessárias muitas funções lineares por partes para ajustar uma função suave, como o seno. Enquanto isso, a função tanh é muito suave e não requer tantas primitivas tanh para construir algo que se assemelhe a uma onda senoidal. Assim, para problemas de classificação, especialmente utilizando múltiplas camadas convolucionais, a ReLU e suas pequenas variantes são difíceis de superar (93).

2.4 Redes Neurais Multilayer Perceptron

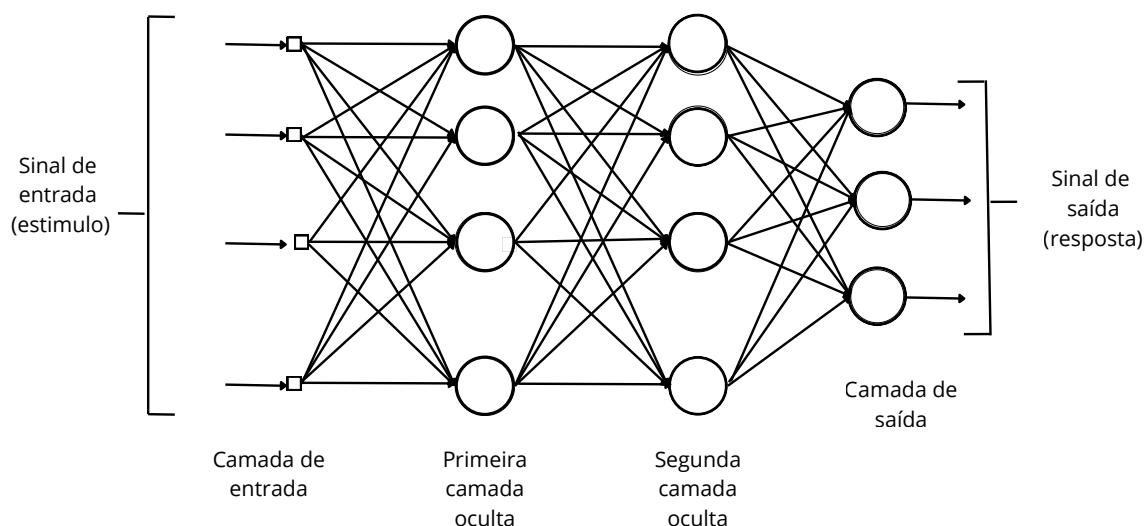
A MLP é um tipo de rede neural *extitfeed-forward* (FFNN) que contém uma ou mais camadas ocultas, onde cada camada possui um ou mais neurônios (101, 102). Ela é uma extensão da rede perceptron e talvez seja o modelo de rede neural mais amplamente utilizado.

Uma MLP com uma única camada oculta é chamada de rede neural rasa (*extitshallow neural network*); com um número suficiente de neurônios ocultos, uma MLP de camada única pode fornecer uma aproximação universal para quase qualquer problema com dados tabulares. Quando a MLP contém mais de uma camada oculta, ela é chamada de rede neural profunda (*extit deep neural network*).

Adicionar mais camadas ocultas pode trazer poucos benefícios, mas aumenta significativamente o custo computacional, pois o número de parâmetros treináveis cresce, tornando a rede mais propensa ao *exitoverfitting* (103).

Na Figura 11 temos a estrutura da rede neural do tipo perceptrons multicamadas (MLP). A rede é composta por três camadas principais: camada de entrada, camada oculta e camada de saída. A camada de entrada é formada por 410 neurônios, correspondentes aos pares de dados normalizados (x,y). Então, a camada de saída é composta por 15 neurônios, responsáveis pela predição dos parâmetros gaussianos normalizados associados ao ajuste dos espectros Raman.

Figura 11 – Representação da rede Multiperceptron.



Fonte: O autor.

Ela consiste em três camadas de nós: uma camada de entrada, uma camada oculta e uma camada de saída. Os dados são transmitidos apenas na direção progressiva (*exitforward*), das unidades de entrada, passando pelas ocultas até as unidades de saída. Exceto pelos nós de entrada, cada nó é um neurônio, incluindo um neurônio de viés, que realiza cálculos utilizando uma função de ativação não linear (102, 104).

2.4.1 Definição da camada densa

Em uma rede neural do tipo Multilayer Perceptron (MLP), as camadas são constituídas por neurônios organizados de forma totalmente conectada. Isso significa, conforme descrito no texto original, que “existe um alto grau de conectividade entre os neurônios, o que significa que um neurônio de qualquer camada da rede está conectado a todos os neurônios da camada

anterior”, caracterizando uma arquitetura totalmente conectada típica das redes Multilayer Perceptron. (105).

Esse padrão de conectividade caracteriza o que se denomina camada densa (ou *fully connected layer*) (105), pois cada neurônio recebe como entrada todas as ativações produzidas pela camada anterior. Como consequência, qualquer alteração estrutural — como adicionar ou remover neurônios — afeta diretamente o número total de conexões sinápticas e seus respectivos pesos. Esse tipo de camada é fundamental no MLP, pois permite que a rede combine informações de forma global, produzindo transformações lineares ou não lineares (quando associada a funções de ativação) (105).

2.4.2 Empilhamento (entrada–ocultas–saída)

O MLP apresenta uma arquitetura sequencial composta tipicamente por três regiões funcionais: a camada de entrada, uma ou mais camadas ocultas e a camada de saída (105). No texto, isso é evidenciado pelo trecho que descreve que o sinal de entrada “propaga-se de camada em camada a partir da camada de entrada, chegando à camada de saída”, caracterizando o fluxo de dados típico de redes do tipo *feed-forward* (105).

Além disso, destaca-se que o MLP “contém uma ou mais camadas ocultas, que são diferentes da camada de entrada e de saída”. Essas camadas intermediárias são responsáveis por possibilitar que a rede aprenda representações internas mais abstratas e resolva problemas complexos. Assim, o empilhamento entrada e ocultas saída fornece a estrutura hierárquica essencial que permite que o modelo aprenda relações não lineares a partir dos dados por meio de sucessivas transformações realizadas em cada camada (105).

2.4.3 Fluxo feed-forward e por que isso resolve mapeamentos não lineares de espectros

No perceptrons multicamadas (MLP), o processamento inicia-se pelo fluxo feed-forward, no qual “o sinal de entrada propaga-se de camada em camada a partir da camada de entrada, chegando à camada de saída”. Nesse percurso, cada camada transforma o sinal recebido até a obtenção da saída, posteriormente comparada ao alvo desejado (105).

A capacidade do MLP de modelar mapeamentos não lineares, como os presentes em dados espectrais, decorre de três elementos destacados no texto original. O primeiro é o uso de funções de ativação não lineares, que possuem “curvatura suave” e tornam o modelo apto a representar relações complexas. O segundo é a presença de camadas ocultas, responsáveis pela aprendizagem de padrões mais elaborados. O terceiro é o alto grau de conectividade (105), que garante que cada neurônio receba informações de todos os neurônios da camada anterior,

permitindo combinações ricas entre as variáveis de entrada.

Assim, o fluxo feed-forward (105), aliado à não linearidade das ativações, às camadas ocultas e às conexões densas, possibilita que o MLP represente relações não lineares típicas de dados espectrais.

2.5 Definição de *Overfitting*

Em aprendizado de máquina supervisionado, existe um problema inevitável. O modelo não generaliza bem dos dados observados para dados não vistos, o que é chamado de *overfitting* (106). Devido à existência do *overfitting*, o modelo apresenta desempenho perfeito no conjunto de treinamento, mas se ajusta mal ao conjunto de teste. Isso ocorre porque o modelo sobreajustado tem dificuldade em lidar com partes das informações no conjunto de teste, que podem ser diferentes daquelas presentes no conjunto de treinamento. Por outro lado, modelos sobreajustados tendem a memorizar todos os dados, incluindo o ruído inevitável no conjunto de treinamento, em vez de aprender o padrão subjacente escondido nos dados.

As causas desse fenômeno podem ser complexas. Geralmente, podemos categorizá-las em três tipos:

1. **Aprendizado de ruído no conjunto de treinamento:** quando o conjunto de treinamento é pequeno, possui dados pouco representativos ou contém muito ruído de natureza aleatória. Essa situação faz com que o ruído tenha grandes chances de ser aprendido e, posteriormente, sirva como base para previsões. Portanto, um algoritmo bem projetado deve ser capaz de distinguir dados representativos de ruídos (107).
2. **Complexidade da hipótese:** o *trade-off* em complexidade, um conceito-chave em estatística e aprendizado de máquina, é um compromisso entre *variância* e *viés*. Refere-se ao equilíbrio entre precisão e consistência. Quando os algoritmos possuem muitas hipóteses (ou muitas entradas), o modelo se torna mais preciso em média, mas com menor consistência (107). Essa situação significa que os modelos podem se comportar de maneira muito diferente em conjuntos de dados distintos. Como consequência, o modelo apresenta alta variância e menor capacidade de generalização, podendo resultar em *overfitting*.
3. **Procedimentos de múltiplas comparações,** que são onipresentes em algoritmos de indução, assim como em outros algoritmos de Inteligência Artificial (IA) (108). Durante esses processos, sempre comparamos múltiplos itens com base em pontuações de uma função de avaliação e selecionamos o item com a pontuação máxima. No entanto, esse processo provavelmente escolherá alguns itens que não irão melhorar, ou até reduzir, a acurácia da classificação (109). Dessa forma, são necessários ajustes para evitar o sobreajuste e

melhorar a capacidade de generalização do modelo, os quais são realizados por meio da definição adequada de hiperparâmetros, como taxa de aprendizado, número de épocas, arquitetura da rede e técnicas de regularização.

3 MATERIAIS E MÉTODOS

3.1 Preparação das amostras

As amostras foram preparadas utilizando o método de fusão e resfriamento nas dependências do grupo de Propriedades Eletrônicas e Vibracionais em Materiais (GPEViM) da Universidade Estadual de Ponta Grossa (UEPG).

Como reagentes, foram utilizados TeO_2 (99% SIGMA-ALDRICH), Li_2CO_3 (99% LAFAN), ZnO (99,99% ALFA AESAR) e Cr_2O_3 (99% ALFA AESAR). Foi utilizada a matriz vítrea com a composição de 65 TeO_2 -15 Li_2O -20 ZnO , e a porcentagem do dopante Cr_2O_3 foi 0,1; 0,25; 0,35; 0,50; 0,75; 1,0 e 2,0 em mol%. A Tabela 2 apresenta a composição e a nomenclatura das amostras obtidas.

Tabela 2 – Nomenclatura e composição das amostras preparadas.

Amostra	TeO_2 (mol%)	Li_2O (mol%)	ZnO (mol%)	Cr_2O_3 (mol%)
TLZ	65,00	15,00	20,00	–
Cr-0,10	64,97	14,97	19,97	0,10
Cr-0,25	64,92	14,92	19,92	0,25
Cr-0,35	64,88	14,88	19,88	0,35
Cr-0,50	64,83	14,83	19,83	0,50
Cr-0,75	64,75	14,75	19,75	0,75
Cr-1	64,66	14,66	19,66	1,00
Cr-2	64,33	14,33	19,33	2,00

Fonte: o autor.

As massas dos reagentes foram pré-calculadas; após esse processo, foram aferidas em uma balança analítica com precisão de 0,1 mg e, então, misturadas e homogeneizadas em almofariz de ágata por aproximadamente 50 minutos. A mistura obtida foi colocada em um cadinho de platina e levada ao forno para o processo de fusão.

A temperatura do forno foi elevada a uma taxa de 3 °C/min até 400 °C, onde permaneceu por 60 minutos. Depois disso, a temperatura foi elevada a uma taxa de 6 °C/min até 870 °C e mantida por 30 minutos para a homogeneização do material fundido.

Em seguida, a amostra foi retirada do forno e submetida ao processo de choque térmico, que consiste em verter o material fundido em um molde previamente aquecido a uma temperatura de 20 °C abaixo da temperatura de transição vítrea do material. Para isso, foi utilizado um molde de latão pré-aquecido. Após o rápido resfriamento do material fundido, obtém-se a amostra vítrea. Posteriormente, a amostra foi levada ao segundo forno, o qual foi aquecido para um tratamento térmico a 240 °C por um período de 240 minutos. Este tratamento teve como objetivo eliminar as tensões mecânicas geradas no processo de vitrificação.

3.2 Espectroscopia Raman

As medidas de espectroscopia Raman foram realizadas utilizando um espectrofotômetro da marca Bruker, modelo Senterra, pertencente ao C-Labmu/UEPG. Foi utilizado um laser de excitação com comprimento de onda de 532 nm, uma grade de difração com 1200 linhas/mm e um microscópio acoplado com lente objetiva de 50× e abertura numérica de 0,40. As medições foram feitas em amostras maciças (*bulk*), na faixa de 100 a 1500 cm^{-1} , com um tempo de integração de 5 segundos.

Na análise dos espectros Raman de vidros teluretos, os três parâmetros associados a cada banda vibracional são de particular interesse: a posição, a largura e a área. Cada modo vibracional é modelado como uma função gaussiana, cujo centro corresponde à posição Raman mais provável (ou número de onda médio) daquela vibração, enquanto a largura à meia altura da gaussiana está relacionada ao grau de desordem estrutural da rede vítrea, isto é, à sua amorficidade. Já a área sob a gaussiana é proporcional à contribuição relativa daquele modo para o espectro total, podendo ser interpretada como uma medida do número efetivo de unidades vibracionais associadas a esse modo em comparação com as demais contribuições do espectro.

Ao analisar o comportamento desses parâmetros em função da concentração, torna-se possível compreender de que forma a alteração composicional influencia a estrutura: se há deslocamento das posições centrais dos picos ou, no caso da área, se há mudança no número relativo de vibrações associadas a cada modo. Dessa forma, é possível inferir como essas variações afetam as proporções estruturais de TeO_3 , TeO_{3+1} e TeO_4 , as quais, por sua vez, estão diretamente relacionadas a diversas propriedades macroscópicas do material, como estabilidade térmica, densidade e energia de *band gap* óptico, entre outras (10). Assim, alterações na estrutura dos acoplamentos e nos agregados de partículas modificam as propriedades funcionais do material, pois afetam a organização estrutural em nível microscópico, refletindo-se em propriedades macroscópicas como estabilidade térmica, densidade e energia de *band gap*.

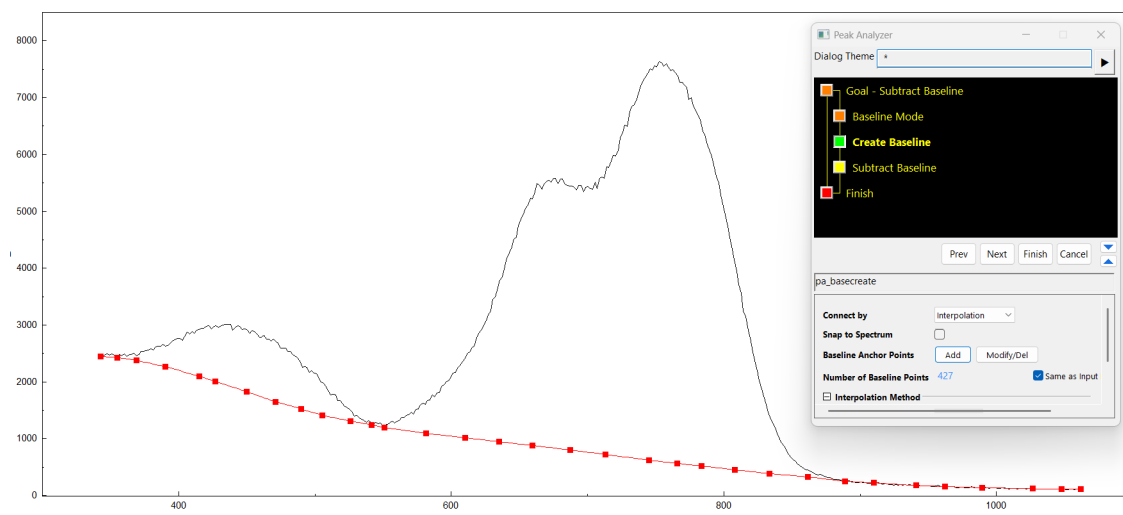
3.3 Pré-processamento - remoção da baseline e normalização pela área

Para a análise dos espectros Raman, tanto pelo método convencional de deconvolução gaussiana quanto pelo uso de redes neurais, foi necessário realizar previamente o pré-processamento dos dados experimentais. O primeiro procedimento consistiu na correção da linha de base, etapa fundamental para remover contribuições indesejadas provenientes do fundo do espectro, que podem interferir na identificação das bandas vibracionais características do material.

A remoção da linha de base foi realizada utilizando as ferramentas de ajuste disponíveis

no *software Origin*, conforme ilustrado na Figura 12. Esse procedimento consiste em ajustar e subtrair uma função matemática que representa a contribuição da linha do laser e do espalhamento Rayleigh, resultando em espectros contendo predominantemente as bandas associadas aos modos vibracionais do material. Essa etapa é particularmente importante em materiais vítreos, como os teluretos, pois o espalhamento Raman ocorre em regiões espectrais relativamente próximas à linha do laser.

Figura 12 – Obtenção do espectro Raman das vibrações modelulares pela remoção da linha de base.



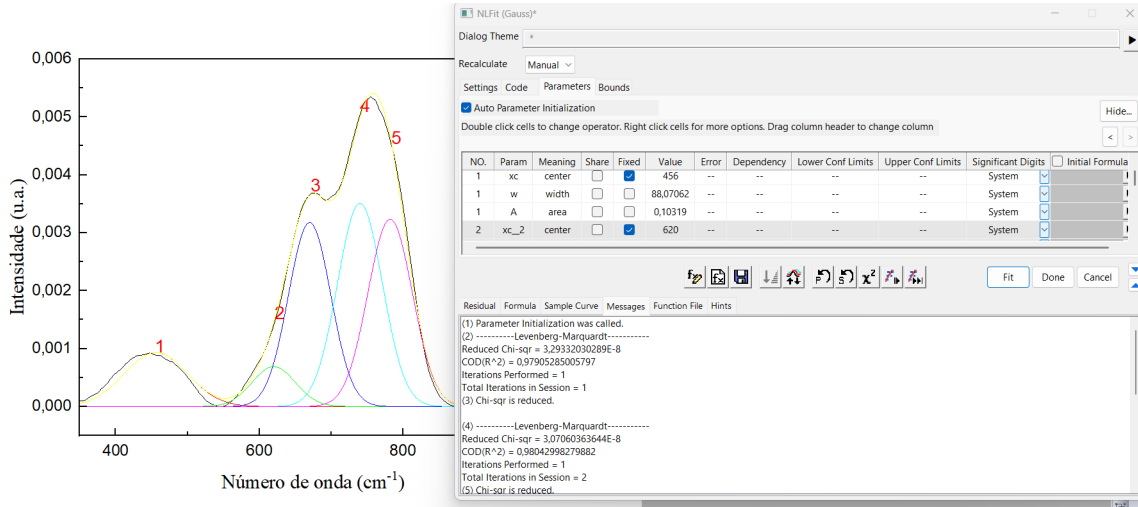
Fonte: O autor.

Após a etapa de correção, os espectros foram submetidos ao processo de normalização pela área total, de modo que as áreas sejam proporcionais em relação à área total das vibrações. Essa normalização assegura que as diferenças para o intervalo entre as amostras estejam relacionadas unicamente às variações estruturais e composicionais, e não a fatores experimentais, como a intensidade do feixe incidente ou pequenas variações nas condições de aquisição. Dessa forma, torna-se possível uma comparação direta e quantitativa entre os espectros obtidos por diferentes métodos de análise, garantindo maior confiabilidade nos resultados e permitindo uma avaliação mais precisa do desempenho das redes neurais na identificação e interpretação dos espectros Raman. Com os espectros devidamente normalizados, foi possível realizar a deconvolução na região espectral de interesse ($380\text{-}1000\text{ cm}^{-1}$).

3.3.1 Deconvolução gaussiana

A Figura 13 apresenta o processo de deconvolução do espectro experimental por meio do ajuste de funções gaussianas. O espectro obtido experimentalmente (linha preta) foi decomposto em cinco contribuições individuais (curvas coloridas), cuja soma é representada pela curva amarela. Cada gaussiana é descrita por três parâmetros principais: a posição central (x_c), a largura à meia altura (w) e a área sob a curva (A), utilizando o método de Levenberg-Marquardt.

Figura 13 – Exemplo de ajuste por deconvolução gaussiana no Origin.



Fonte: O autor.

A intensidade espectral de uma banda $I_i(x)$ para uma função gaussiana em função do número de onda em cm^{-1} , que é proporcional à mudança da energia do fóton espalhado com relação à energia do fóton incidente (Deslocamento Raman), aqui detonado por x , pode ser descrita por:

$$I_i(x) = \frac{A_i \sqrt{2}}{w_i \sqrt{\pi}} e^{-2 \left(\frac{x - x_{ci}}{w_i} \right)^2}, \quad (3.1)$$

onde A_i é a área, w_i a largura à meia altura e x_{ci} o centro, da i -ésima gaussiana. Assim, a curva espectral convoluída é:

$$I(x) = y_0 + \sum_{i=1}^N I_i(x) = y_0 + \sum_{i=1}^N \frac{A_i \sqrt{2}}{w_i \sqrt{\pi}} e^{-2 \left(\frac{x - x_{ci}}{w_i} \right)^2}. \quad (3.2)$$

É importante destacar que, neste procedimento, é necessário fixar os valores teoricamente conhecidos. Isso se deve ao fato de que, em um ajuste de uma curva composta por 5 curvas com 3 parâmetros ajustáveis, que totalizam 15 parâmetros (mais uma altura inicial y_0), no *Origin*, todos sendo livres, podem ocorrer vários erros, como largura acima dos valores reais ou até áreas negativas, entre outras variações que não têm significado físico real. Entretanto, mesmo com ajustes sem significado real, a curva interpolada se aproxima da curva experimental, apresentando um bom valor de coeficiente de correlação.

Para resolver esse problema, geralmente, a maioria dos valores, como centros (x_c) e larguras (w) das gaussianas, é mantida fixa com base na literatura para ajustar a área. Para melhorar e ter um ajuste fino, é necessária experiência em verificar quais parâmetros podem ser liberados para variar e permitir a obtenção de resultados satisfatórios com significado físico. No exemplo da Figura 13, os principais valores fixados para o ajuste foram os centros em $x_{c1} = 450$,

com w_1 livre, $x_{c2} = 615$ com $w_2 = 55$, $x_{c3} = 669$ com $w_3 = 60$, $x_{c4} = 735$ com $w_4 = 60$ e $x_{c5} = 735$ com $w_5 = 60$, todos em cm^{-1} .

A Figura 14 apresenta os resultados dos ajustes deconvoluídos para as oito amostras de TLZ dopadas com Cr_2O_3 . Nas figuras também são apresentados os valores encontrados pelos ajustes em cada amostra. É importante apresentar os resultados nesta seção, pois os valores obtidos aqui servirão como guia para a criação das bases de dados que servirão de treinamento para as redes neurais artificiais utilizadas neste trabalho. Em todos os casos, o espectro experimental é bem descrito como a soma de cinco gaussianas, com valores de R^2 entre 0,997 e 0,999, reforçando que o espectro na faixa de $380\text{--}1000\text{ cm}^{-1}$ pode ser compreendido em termos de cinco modos principais associados às unidades TeO_4 , TeO_{3+1} e TeO_3 discutidas na Seção 2.1.5. A Tabela 3 apresenta todos os intervalos espectrais obtidos para os vidros teluretos dopados com Cr_2O_3 .

Tabela 3 – Intervalos da espectroscopia Raman obtidos das deconvoluções pelo *Origin*.

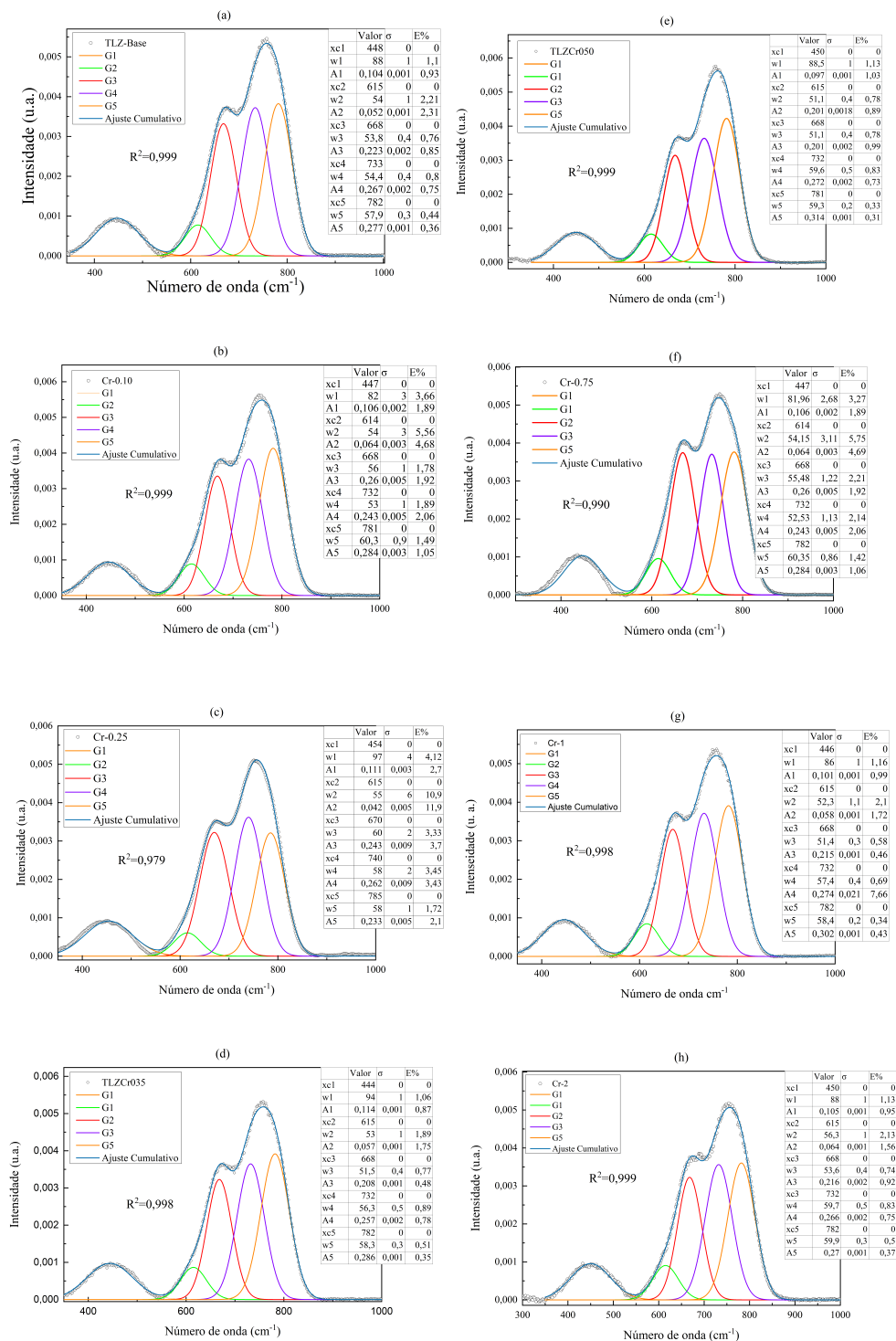
Intervalos	1	2	3	4	5
x_{ci}	444-454	614-615	668-670	732-740	781-785
w_i	82-97	52-56	51-60	52-59	58-60
A_i	0,097-0,111	0,042-0,201	0,201-0,260	0,243-0,274	233-0,314

Fonte: O autor.

Analisando os parâmetros de ajuste para as oito amostras, os centros x_{ci} das cinco bandas permanecem próximos aos valores de referência adotados com base na literatura (em torno de $450, 610\text{--}620, 660, 730$ e 780 cm^{-1}), com variações relativamente pequenas.

No que diz respeito às larguras à meia altura w_i , verifica-se que, para cada banda, os valores ajustados permanecem na faixa típica de algumas compatíveis com a natureza amorfa do material. As larguras não apresentam um crescimento descontrolado (como ocorreria em ajustes mal condicionados), o que confirma que o procedimento de manter centros e, em alguns casos, larguras parcialmente fixas conduziu a soluções fisicamente plausíveis. Por fim, as áreas A_i das cinco gaussianas são os parâmetros que mais claramente refletem as mudanças estruturais induzidas pela dopagem com Cr_2O_3 .

Figura 14 – Resultados dos ajustes pelo método clássico de deconvolução gaussiana dos espectros de intensidade normalizada dos Deslocamentos Raman pelo *Origin* para as oito amostras de TLZ dopadas com Cr_2O_3 .



Fonte: O autor.

3.4 Redes Neurais Artificiais

Para a criação de uma base de dados, para realizar o treinamento e a predição foi utilizado o *software Wolfram Mathematica*, que oferece uma ampla seleção de métodos de aprendizado de máquina para realizar regressão, classificação, agrupamento, redução de dimensionalidade, entre outros. Para verificar a eficiência das redes neurais e compará-las para a predição dos parâmetros gaussianos dos espectros Raman dos vidros teluretos, foram realizadas diferentes etapas: a) verificação da influência do número de neurônios em redes MLP com 1 e 2 camadas ocultas, sem novos processamentos dos resultados (sem normalização); b) normalização dos resultados de entrada e saída no intervalo de 0 a 1 e verificação da influência desse procedimento na predição; c) adição de ruídos aleatórios nos espectros teóricos de treinamento nos valores de 1%, 2% e 5% e adição de uma linha de base constante. Em cada etapa, foram seguidos os seguintes passos:

- 1º Criação de bases de dados para treinamento e validação das redes neurais. Cada curva utilizada para o treinamento é composta pelo par de dados $(x, I(x))$ no mesmo intervalo dos resultados experimentais de interesse, ou seja, de $300-1000 \text{ cm}^{-1}$. Como a rede será utilizada para predição, um fato importante é que as bases de dados e a curva que será utilizada devem ser um vetor de mesmo tamanho, isto é, conter o mesmo número de pares de dados. No caso em estudo foram 410 pares de dados, com passos de aproximadamente $1,7 \text{ cm}^{-1}$. Ainda, para poder reproduzir mais fielmente resultados experimentais onde o passo experimental do Deslocamento Raman (x) não é uniforme, uma variação aleatória no intervalo $(1,65 \text{ a } 1,75) \text{ cm}^{-1}$ considerando função randômica para gerar o passo de incremento no valor de x . Para cada curva da base de dados um novo conjunto de 410 pontos de x foi gerado, onde para a curva convoluída pela equação 3.2 era escolhido randomicamente dentre uma tabela de intervalos dos respectivos 15 parâmetros.
- 2º Curvas de teste teóricas para predição foram criadas. Com essas curvas pode verificar se as redes neurais conseguem prever com precisão os resultados calculando os erros percentuais com relação aos valores utilizados para gerar as curvas teóricas.
- 3º Determinação das arquiteturas das redes neurais e treinamento com as bases de dados. Nessa etapa foram escolhidos o número de camadas e números de neurônios em cada camada, para comparar o desempenho das redes com diferentes configurações. Ainda, foram testadas diferentes funções de ativação: ReLu (Ramp), Tanh e Sigmóide. As redes são treinadas com as entradas compostas pelas curvas de pares de dados, que estão associadas à saída: um vetor de 15 parâmetros, sendo os 3 parâmetros (centros, larguras e áreas) das 5 gaussianas que geraram a curva. Nessa etapa o Erro Quadrático Médio (MSE - *Mean Square Error*) é utilizado como parâmetro de análise do desempenho de treinamento em função das épocas de treino.

- 4º Posteriormente, as redes treinadas são utilizadas primeiramente para predição dos 15 parâmetros utilizados na geração das curvas teóricas, sendo calculados os erros percentuais, e a predição da convolução do espectro, sendo calculado o valor do coeficiente de correlação (R^2).
- 5º Por fim, as redes são utilizadas para prever os 15 parâmetros dos resultados experimentais, comparar com os resultados obtidos pelo *Origin* e observar a qualidade do ajuste pelo coeficiente de correlação R^2 .

3.4.1 Base de dados sem normalização

A primeira etapa apresentada é a geração da base de dados realizada a partir da definição de intervalos específicos para os três parâmetros: x_{ci} , w_i , A_i . Os intervalos foram estabelecidos de modo a representar faixas de variação observadas nos dados experimentais obtidos pelo *Origin*, aumentando, em média, em 1 cm^{-1} os intervalos para a possível predição de outros valores não fixados no método convencional. Os intervalos definidos são apresentados na Tabela 4:

Tabela 4 – Intervalos dos parâmetros utilizados para geração da base de dados para os convolução dos espectros de Deslocamento Raman sem normalização

Intervalos	1	2	3	4	5
x_{ci}	443 - 451	613 - 616	667 - 669	731 - 735	780 - 783
w_i	81 - 97	50 - 57	50 - 56	50 - 61	56 - 61
A_i	90 - 120	40 - 70	190 - 270	230 - 290	260 - 320

Fonte: O autor.

Com base nesses intervalos, uma função aleatória foi utilizada de modo que eram obtidos os valores x_{ci} , w_i , A_i para cada intervalo de uma das 5 gaussianas na geração do espectro convoluído, representando diversas possibilidades de espectro. Os valores do par de dados $(x, I(x))$ eram, então, salvos em um arquivo de dados de extensão .txt como entradas e os 15 parâmetros salvos em um segundo .txt como as saídas.

Após vários testes preliminares, com o objetivo de manter a coerência entre as escalas e facilitar a análise comparativa, os valores obtidos para as áreas das bandas foram divididos por 1000 na equação 3.1 e os valores dos intervalos na Tabela 4 foram multiplicados por 1000, de modo que todos os parâmetros (centro, largura e área) permanecessem dentro de uma mesma ordem de grandeza. Esse procedimento simplificou a visualização e interpretação dos resultados, evitando discrepâncias de escala que poderiam dificultar a comparação entre diferentes bandas e amostras.

Para verificar a influência do volume de dados no treinamento, foram geradas bases de dados com 500, 1000, 2000, 5000, 10000 e 20000 curvas. A primeira base de dados (500 curvas)

era utilizada unicamente para validação, enquanto as demais eram usadas para o treinamento das redes neurais. Esse fator é importante para a comparação entre as redes neurais, uma vez que o mesmo conjunto de curvas e parâmetros é utilizado para validação.

3.4.2 Base de dados com normalização

Na segunda etapa, com o intuito de melhorar a capacidade de predição da rede e reduzir o MSE tanto no treinamento quanto na validação, aplicou-se uma normalização no intervalo de 0 a 1 para todas as variáveis (entradas e saídas).

Uma possível fonte relevante de erro na primeira etapa é o desbalanço de escalas entre as variáveis: o deslocamento Raman x e os 15 parâmetros ajustados das gaussianas possuem valores da mesma ordem de grandeza, enquanto as intensidades $I(x)$ são da ordem de 10^{-3} desses valores. Esse descompasso faz com que a contribuição de $I(x)$ para a função custo seja relativamente pequena, dificultando o ajuste adequado dos pesos da rede. Ao normalizar todas as variáveis para a faixa comum $[0,1]$, esse desbalanço é mitigado, resultando em uma superfície de erro melhor condicionada e, em geral, em modelos com menor MSE. É amplamente relatado na literatura de redes neurais que o pré-processamento dos dados por meio de normalização ou padronização, trazendo todas as variáveis para faixas de valores comparáveis (por exemplo, 0–1 ou média zero e variância unitária), melhora o condicionamento do problema de otimização, acelera a convergência do gradiente e tende a reduzir o erro em problemas de regressão (110–114). Além disso, alguns intervalos também foram reduzidos para a redução de erros, conforme apresentado na Tabela 5.

Tabela 5 – Intervalos dos parâmetros utilizados para geração da base de dados para os convolução dos espectros de Deslocamento Raman normalizados

Intervalos	1	2	3	4	5
x_{ci}	0,443 - 0,451	0,613 - 0,616	0,667 - 0,669	0,731 - 0,735	0,780 - 0,783
w_i	0,84 - 0,94	0,50 - 0,54	0,50 - 0,56	0,57 - 0,61	0,57 - 0,61
$3 \times A_i$	0,090 - 0,120	0,040 - 0,070	0,170 - 0,270	0,230 - 0,290	0,260 - 0,320

Fonte: O autor.

Assim, os fatores ou pesos de normalização foram PA (peso da área), PW (peso da largura), PC (peso do centro e do eixo x do deslocamento Raman) e PI (peso da intensidade) a fim de que todos os valores do treinamento estivessem no mesmo intervalo. A intensidade de cada gaussiana foi calculada por:

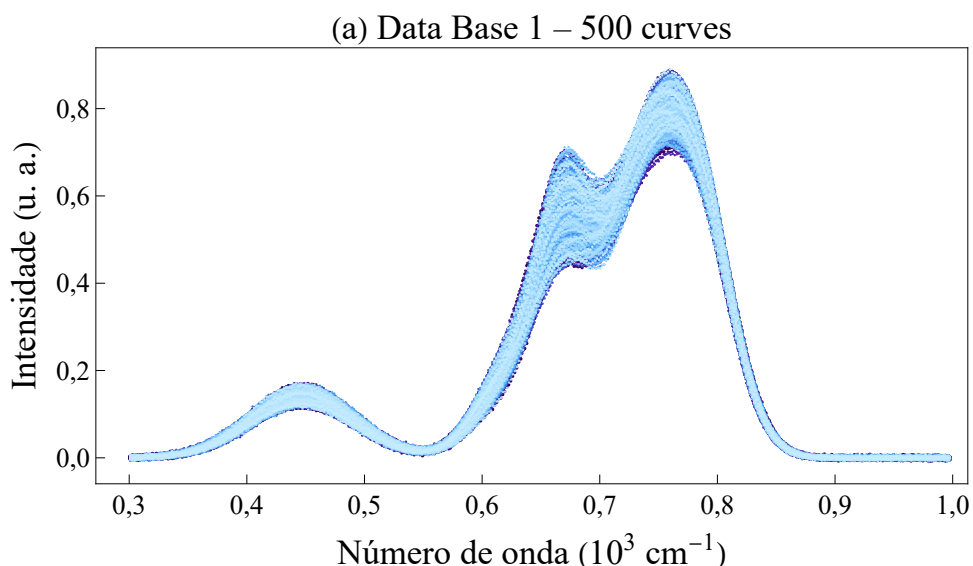
$$I_i(x) = PI \frac{\left(\frac{A_i}{PA}\right) \sqrt{2}}{\left(\frac{w_i}{PW}\right) \sqrt{\pi}} e^{-2 \left(\frac{\left(\frac{x-x_{ci}}{PC}\right)}{\left(\frac{w_i}{PW}\right)}\right)^2}, \quad (3.3)$$

com os valores dos pesos:

$$PA = 3 \quad ; \quad PW = \frac{1}{100} \quad ; \quad PC = \frac{1}{1000} \quad ; \quad PI = 150$$

Utilizando a equação 3.3 para cada gaussiana e os intervalos 5, foram criadas novas bases de dados com 500, 1000, 2000, 5000, 10000 e 20000 curvas para treinamento e validação das redes, além de salvar os dados de entrada e saída e treinar as redes neurais MLP. Para exemplificação, na Figura 15 são apresentadas todas as curvas geradas para as bases de dados de 500 curvas para validação. As primeiras curvas geradas são apresentadas com pontos maiores e tons mais escuros de azul, apresentando uma função decrescente no tamanho do ponto e tons para melhor visualização. Nota-se que, dentro do intervalo com 500 curvas, já há uma cobertura de diversas possibilidades de espectros. Com o aumento das curvas, a densidade e as possibilidades de variações de curvas aumentam.

Figura 15 – Exemplo de curvas de deslocamento Raman teóricas sem ruído geradas para base de dados para validação de redes neurais artificiais MLP.



Fonte: O autor.

3.4.3 Bases de dados normalizadas com adição de ruído

Após o treinamento das redes normalizadas e a predição dos parâmetros gaussianos, três novas bases de dados foram criadas, agora considerando ruídos na intensidade $I(x)$ por meio de uma função randômica nos valores de 1%, 2% e 5%. Após verificar a eficiência no treinamento das redes neurais em relação aos resultados experimentais, foi gerada uma sexta e última base de dados considerando uma baseline constante y_0 na base normalizada com 2% de ruído

Após o treinamento das redes com os dados normalizados e a predição dos parâmetros gaussianos, foram construídas três novas bases de dados com o objetivo de avaliar os modelos frente a incertezas experimentais. Nessas bases, introduziu-se ruído aleatório na intensidade $I(x)$

por meio de uma função randômica de amplitude relativa de 1%, 2% e 5%, respectivamente, em relação ao valor de $I(x)$ em cada ponto do espectro. Dessa forma, buscou-se simular diferentes níveis de ruído de medida típicos de experimentos reais, verificando como variações estatísticas na intensidade afetam a capacidade da rede de recuperar com precisão os parâmetros das cinco gaussianas.

Após verificar a eficiência do treinamento das redes neurais nessas condições ruidosas e comparar os resultados com os espectros experimentais, foi gerada uma sexta e última base de dados. Nessa base, considerou-se a inclusão de uma linha de base constante y_0 somada à intensidade, tomando como referência a base normalizada com 2% de ruído. Essa estratégia permite analisar, simultaneamente, o efeito combinado de ruído estatístico e de um *baseline* deslocado, situação frequentemente observada em espectros Raman experimentais devido à fluorescência residual ou contribuições instrumentais ou pré-processamento. Assim, a sexta base de dados foi utilizada como um teste mais rigoroso da capacidade de generalização da rede, especialmente quanto à correta predição dos parâmetros gaussianos na presença de ruído e de linha de base não nula.

3.4.4 Arquiteturas das redes neurais

O treinamento das redes neurais foi realizado no ambiente *Wolfram Language*, por meio da função `NetTrain`. A arquitetura da rede não é predefinida pelo ambiente, devendo ser explicitamente construída pelo usuário a partir de camadas e funções disponíveis na biblioteca. A seguir, apresenta-se um exemplo típico de comando utilizado para o ajuste dos parâmetros gaussianos a partir dos espectros Raman:

```
NetTrain[
  NetChain[{LinearLayer[150], Ramp, LinearLayer[15]}],
  <|"Input" -> entradasFormatadas1,
    "Output" -> saidasFormatadas1|>,
  LossFunction      -> MeanSquaredLossLayer[],
  TargetDevice      -> "GPU",
  MaxTrainingRounds -> 5000,
  BatchSize         -> 64,
  Method            -> {"ADAM", "LearningRate" -> 0.001},
  ValidationSet      -> <|"Input" -> entradasValidacao,
    "Output" -> saidasValidacao|>
]
```

Nesse exemplo, `NetChain[LinearLayer[150], Ramp, LinearLayer[15]]` é a es-

pecificação da arquitetura da rede. O comando `NetChain` define uma rede feedforward sequencial (MLP), em que a saída de cada camada é passada diretamente como entrada para a camada seguinte. A primeira camada oculta, `LinearLayer[300]`, é uma camada totalmente conectada (*fully connected*) com 300 neurônios. A dimensão de entrada dessa camada é inferida automaticamente a partir do formato de `entradasFormatadas1`, que contém os vetores de entrada (neste caso, os 410 pares $(x, I(x))$ dos espectros Raman normalizados). Em seguida, aplica-se a função de ativação Ramp, equivalente à função ReLU ($f(x) = \max(0, x)$), introduzindo não linearidade na rede. Por fim, a camada `LinearLayer[15]` gera um vetor de saída com 15 componentes, correspondentes aos parâmetros gaussianos a serem preditos (centros, larguras e áreas das cinco bandas).

Os dados de treinamento são fornecidos a `NetTrain` sob a forma de uma associação (*Association*) com as chaves "Input" e "Output": `<|"Input-> entradasFormatadas1, "Output-> saidasFormatadas1|>`. Em `entradasFormatadas1` estão os vetores de entrada x (pontos do espectro discretizado), enquanto `saidasFormatadas1` contém, para cada espectro, o conjunto de 15 parâmetros gaussianos. De modo análogo, o argumento `ValidationSet -> <|"Input-> entradasValidacao, "Output-> saidasValidacao|>` define um conjunto de validação separado, utilizado pelo Wolfram para monitorar o desempenho da rede durante o treinamento e evitar sobreajuste (*overfitting*).

A função `LossFunction -> MeanSquaredLossLayer[]` é a função de perda (*loss*). Essa escolha corresponde ao erro quadrático médio (MSE) entre os valores preditos pela rede e os valores-alvo dos parâmetros gaussianos, o que é apropriado para problemas de regressão contínua. Durante o treinamento, o *Wolfram* calcula automaticamente o gradiente desse erro em relação aos pesos da rede por meio de diferenciação automática e atualiza os parâmetros usando o otimizador indicado em `Method`.

No presente trabalho, foi utilizado o algoritmo de otimização ADAM, configurado como `Method -> {"ADAM", "LearningRate-> 0.001}`. O ADAM é um método de descida de gradiente estocástica com momentos adaptativos, que ajusta dinamicamente o passo de aprendizado (*learning rate*) para cada parâmetro, combinando estimativas de primeira e segunda ordens dos gradientes. O valor "LearningRate-> 0.001 foi escolhido como compromisso entre velocidade de convergência e estabilidade numérica do treinamento.

No contexto do `NetTrain` do *Wolfram Language*, o parâmetro `TrainingRounds` não corresponde diretamente ao número de épocas, mas sim ao número de atualizações de pesos realizadas durante o treinamento, isto é, ao número de *mini-batches* processados. Em problemas de aprendizado supervisionado, define-se uma *época* (*epoch*) como uma passagem completa de todo o conjunto de treinamento pela rede. Quando se utiliza treinamento em *mini-batches*, com um conjunto de treinamento contendo N_{samples} exemplos e um tamanho de lote dado por `BatchSize`, o número efetivo de épocas pode ser estimado pela relação

$$\text{epochs}_{\text{real}} \approx \frac{\text{TrainingRounds} \times \text{BatchSize}}{N_{\text{samples}}}. \quad (3.4)$$

Ou seja, cada *TrainingRound* corresponde ao processamento de um *mini-batch* com *BatchSize* exemplos, e a fração de época percorrida é dada pela razão entre o número total de exemplos já apresentados à rede e o número total de exemplos do conjunto de treinamento. O parâmetro *MaxTrainingRounds* fixa apenas o limite máximo de *TrainingRounds*, enquanto o número de épocas efetivamente percorridas depende simultaneamente de *BatchSize* e de N_{samples} , bem como de eventuais critérios de parada antecipada utilizados pelo *NetTrain*.

Por fim, o argumento *TargetDevice* -> "GPU" indica que o treinamento deve ser executado na unidade de processamento gráfico (GPU), quando disponível. O *Wolfram Language* gerencia automaticamente a transferência dos tensores de dados (entradas, saídas e pesos da rede) entre a memória da CPU e da GPU, explorando a paralelização massiva para acelerar as operações de multiplicação de matrizes e o cálculo de gradientes.

No trabalho, o número de camadas ocultas foi variado, sendo 1 ou 2 camadas lineares totalmente conectadas, com o número de neurônios em cada camada: 25, 50, 100, 150, 200 ou 400 neurônios. Para comparação com a função Ramp, em alguns casos, foram testadas as funções de ativação Tanh e *Logistic* Sigmoides.

Durante o processo de treinamento, os valores da função de perda (*Mean Squared Error* – MSE) para os conjuntos de treinamento e validação foram monitorados continuamente. Esses dados foram armazenados e utilizados para gerar um gráfico de acompanhamento da convergência da rede, possibilitando a análise do desempenho e da capacidade de generalização do modelo. O tempo total de treinamento foi registrado.

Para avaliar o desempenho dos diferentes modelos de redes neurais, foram utilizadas cinco bases de dados distintas, implementadas em uma plataforma específica para redes neurais, no ambiente *Wolfram Mathematica*. Os experimentos computacionais foram realizados em um computador pessoal do tipo notebook, modelo *Acer Aspire 3*, com processador *Intel Core i5-1135G7* (11ª geração), 4 núcleos de processamento e velocidade de 2,42 GHz, tendo memória RAM de 8 Gb. Cada uma dessas bases apresenta curvas características diferentes, as quais servem como referência para a parametrização individual de cada rede.

3.4.5 Nomenclatura das redes neurais

Neste estudo, foram utilizadas diferentes bases de dados e diversas arquiteturas de redes neurais para avaliar o comportamento dos modelos em função da quantidade de curvas e da estrutura interna da rede. As bases de dados foram divididas em cinco conjuntos distintos, variando conforme o número de curvas disponíveis: a Base 1-2 com 1000 curvas, a Base 1-3

possuía 2000 curvas, a Base 1-4 foi composta por 5000 curvas, a Base 1-5 com 10000 curvas e a Base 1-6 possuía 20000 curvas. Cada uma dessas bases representa um conjunto independente utilizado nos experimentos.

Para essas redes, foram empregadas diferentes funções de ativação, incluindo sigmoid, tanh e ReLU, com o objetivo de avaliar qual combinação entre arquitetura e função de ativação apresentava a melhor capacidade de predição.

As arquiteturas das redes incluíram inicialmente modelos com duas camadas ocultas, identificados pela notação “2C”. Nesse caso, “C” representa o número de camadas ocultas, de modo que 2C indica duas camadas. Foram avaliadas cinco variações: a rede 2C-1, com 100 neurônios na primeira camada e 100 na segunda; a rede 2C-2, com 100 neurônios na primeira camada e 50 na segunda; a rede 2C-3, composta por duas camadas de 50 neurônios; a rede 2C-4, com 50 neurônios na primeira e 100 na segunda; e a rede 2C-5, formada por 300 neurônios na primeira camada e 50 na segunda. Para as próximas redes, foi utilizada apenas a função ReLU, por ter apresentado o melhor desempenho preditivo entre as alternativas testadas.

Também foram avaliadas redes com apenas uma camada oculta, identificadas como “1C”, aplicadas a bases de dados previamente normalizadas entre 0 e 1. Essas redes foram marcadas com o sufixo “Ag”, indicando o uso de dados normalizados. As configurações analisadas foram: 1C-1Ag, com 400 neurônios; 1C-2Ag, com 100 neurônios; 1C-3Ag, com 200 neurônios; 1C-4Ag, com 50 neurônios; e 1C-11Ag, com 150 neurônios.

Além das variações estruturais, foi introduzido ruído artificial nas entradas dessas redes em três intensidades distintas: 1%, 2% e 5%. Por fim, também foi adicionada uma condição de baseline correspondente ao modelo treinado com 1% de ruído para efeito de comparação.

3.5 Curvas de teste teóricas

Foram gerados três conjuntos de dados distintos a partir da soma de cinco funções gaussianas, com o objetivo de representar diferentes condições experimentais e teóricas. O primeiro conjunto, denominado Resultado 1, é sem ruído e sem linha de base; o segundo conjunto, Resultado 2, sem ruído, corresponde ao sinal puramente teórico, obtido pela soma das cinco gaussianas, sem a adição de ruído ou de qualquer tendência de fundo. Um segundo conjunto de dados, denominado Resultado 2 – sem ruído. Os resultados para o teste foram gerados a partir da soma de cinco funções gaussianas com parâmetros definidos conforme a Tabela 6.

Em seguida, foi criado o conjunto denominado Resultado 1, com ruído e sem linha de base, no qual se adicionou uma perturbação aleatória simulando o ruído experimental normalmente presente em medições reais de 2%. Essa perturbação foi gerada a partir de uma

Tabela 6 – Parâmetros das funções gaussianas utilizadas na geração dos sinais teóricos.

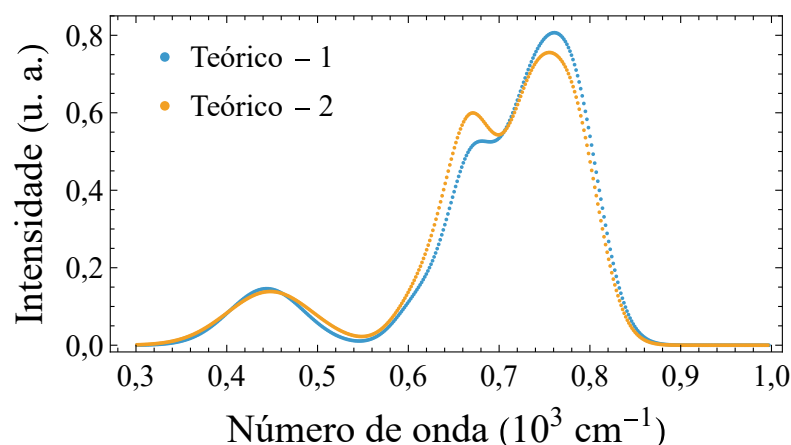
Resultado Teórica 1			Resultado Teórica 2		
x_{ci}	w_i	A_i	x_{ci}	w_i	A_i
444	82	100	448	95	110
614	53	50	615	56	60
669	52	190	668	55	250
732	60	270	733	55	240
782	60	300	781	58	270

Fonte: O autor.

distribuição normal com média zero e desvio padrão, sendo somada diretamente aos valores da função teórica. Dessa forma, o sinal manteve o mesmo perfil das curvas gaussianas originais, porém com pequenas flutuações de intensidade que refletem variações estocásticas.

Por fim, foi elaborado o conjunto Resultado 2 com ruído e com linha de base, no qual, além do ruído aleatório, adicionou-se uma linha de base constante com intensidade de 0,0003. Essa modificação teve o propósito de reproduzir o desvio de fundo frequentemente observado em espectros experimentais, causado por fatores instrumentais ou ambientais. O resultado final na figura 16 foi um conjunto de dados mais realista, contendo tanto o ruído quanto a linha de base.

Figura 16 – Representação gráfica dos resultados teóricos 1 e 2.



Fonte: O autor.

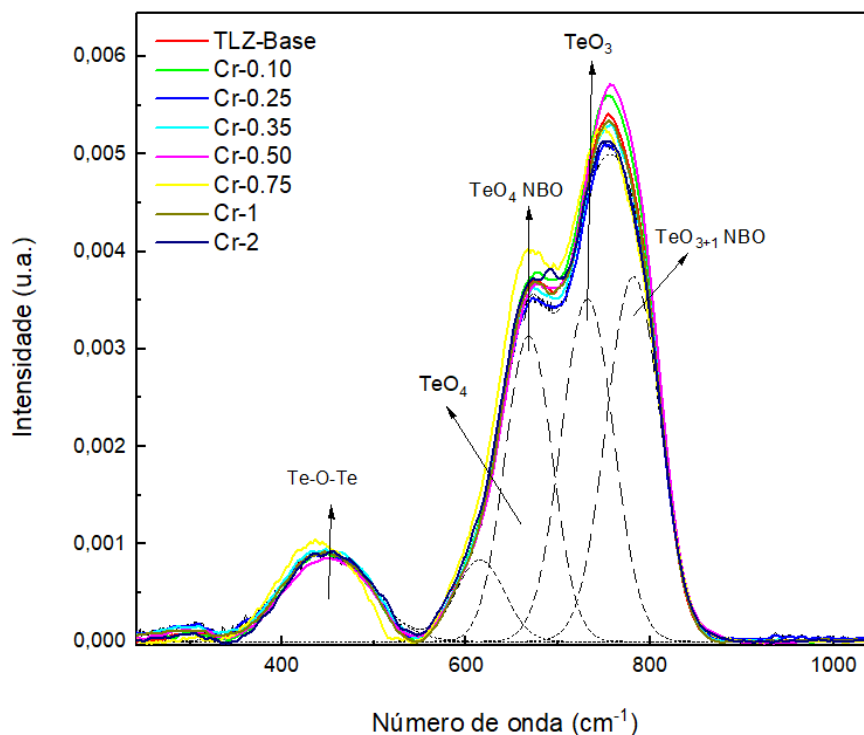
Dessa forma, foi possível obter o resultado referente ao parâmetro experimental 1, o qual corresponde à primeira curva teórica.

4 RESULTADOS E DISCUSSÕES

4.1 Espectros Raman

Figura 17 apresenta os espectros Raman obtidos na faixa de 400 a 1000 cm^{-1} para todas as amostras analisadas. Na literatura, são reportadas cinco bandas características do TeO_2 puro, localizadas em 456, 620, 670, 740 e 782 cm^{-1} , correspondentes aos modos vibracionais do TeO_2 (10, 27, 115). Essas bandas foram selecionadas com base em sua correlação com modos vibracionais específicos das unidades de TeO_2 na estrutura vítrea, conforme relatado em estudos anteriores (10, 116, 117). Aqui apresentamos novamente todos os espectros para compreender qualitativamente a influência do Cr_2O_3 .

Figura 17 – Espectroscopia Raman obtidos.



Fonte: O autor.

A banda em 456 cm^{-1} sugere a presença de ligações Te-O-Te ou O-Te-O , correspondendo ao primeiro pico mostrado na Figura 17. Essas ligações são formadas pela conexão dos vértices das unidades estruturais, típicas de TeO_4 , TeO_3^{+1} e TeO_4 (10, 73, 118, 119). Essa banda, em particular, serve como um indicador da presença de ligações Te-O-Te e O-Te-O na estrutura vítrea (10). Observa-se que essa banda permanece aproximadamente constante com a adição de Cr_2O_3 .

A banda em 615 cm^{-1} é atribuída à vibração das unidades TeO_4 sem oxigênios não

ligantes (NBO). Essa banda surge das vibrações antissimétricas das unidades TeO_4 na ausência de NBO (10, 120). Notavelmente, a intensidade dessa banda permanece constante mesmo com a introdução de Cr_2O_3 . Esse resultado sugere que a presença de Cr_2O_3 , até a concentração de 2 mol%, não influencia o comportamento vibracional das unidades TeO_4 (121).

A terceira banda relevante no espectro Raman, centrada em 668 cm^{-1} , é atribuída às vibrações antissimétricas das unidades TeO_4 (10, 115). Observa-se que a intensidade dessa banda aumenta com a adição de 0,75 mol% de Cr_2O_3 . Posteriormente, ocorre uma diminuição da intensidade à medida que a concentração de Cr_2O_3 aumenta, permanecendo então constante.

A banda centrada em 732 cm^{-1} é atribuída às vibrações das unidades estruturais TeO_{3+1} , formadas devido à quebra das ligações Te-O-Te ou O-Te-O (48, 121). Contudo, essa banda apresenta uma diminuição até a amostra Cr-0,25, seguida de um aumento até a Cr-0,50, permanecendo constante para maiores valores de Cr_2O_3 . Esse comportamento pode ser atribuído ao papel do Cr_2O_3 , que deixa de atuar como dopante e passa a funcionar como modificador de rede.

A banda em 782 cm^{-1} é atribuída às unidades de TeO_3 com NBO, mostrando uma diminuição na região Cr-0,25, seguida de um aumento na região Cr-0,50 com a adição de Cr_2O_3 . Assim, pode-se afirmar que a adição de Cr_2O_3 ao sistema vítreo, na região Cr-0,50, deixa de atuar como dopante e passa a atuar como modificador de rede, influenciando, conseqüentemente, a estrutura do vidro.

4.2 Predição de parâmetros e espectros Raman sem normalização

As redes neurais perceptrons multicamadas (MLP) foram treinadas para prever os espectros Raman de amostras de vidros teluretos dopados com Cr_2O_3 . Os resultados obtidos demonstram o desempenho da rede tanto para os espectros teóricos quanto para as amostras experimentais de Cr_2O_3 . Diversas redes foram treinadas; porém, apenas as principais, com os melhores resultados de predição, foram selecionadas para análise. Todos os resultados apresentados foram obtidos com o mesmo método, utilizando 5000 rounds, em lotes (*batch size*) iguais a 64 e uma taxa de aprendizado de 0,001.

Foram testados diferentes tipos de redes neurais utilizando bases de dados contendo 1.000, 2.000, 5.000, 10.000 e 20.000 curvas. Essa variação no tamanho das bases foi adotada porque, com o aumento da quantidade de dados, os ajustes dos modelos tendem a apresentar melhorias significativas no desempenho. Entretanto, para bases superiores a 10.000 curvas, observa-se um aumento expressivo do custo computacional, caracterizado pelo crescimento do tempo de treinamento à medida que o número de curvas aumenta. Em particular, o tempo de treinamento cresce de forma não linear com o tamanho da base de dados, não havendo uma

compensação proporcional em termos de ganho de desempenho para bases maiores.

Também foram realizados testes com diferentes funções de ativação, com o objetivo de identificar aquela que proporcionaria o melhor desempenho preditivo. Observou-se que, ao utilizar as funções Sigmoid e Tanh, as redes não foram capazes de realizar previsões satisfatórias, resultando em sobreajuste (*overfitting*). Por outro lado, as redes que empregaram a função ReLU apresentaram desempenho adequado, permitindo a predição dos espectros de espectroscopia Raman. As melhores redes foram treinadas com 10000 curvas de entrada. De maneira geral, as redes com duas camadas tiveram melhor desempenho do que as de uma camada oculta.

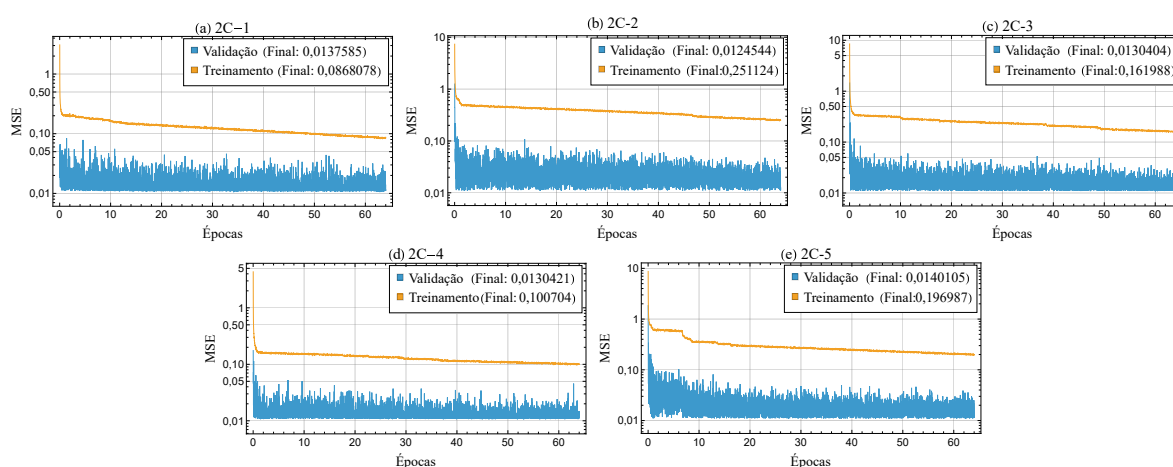
Tabela 7 – Resultados do treinamento, validação e R^2 das redes sem normalização com duas camadas. Todas as redes com função de ativação Ramp (ReLU).

Rede	Validação	Treinamento	R^2-1	R^2-2
2C-1	0,0137585	0,0868078	0,9947	0,9918
2C-2	0,0124544	0,251124	0,9970	0,9877
2C-3	0,0130404	0,161988	0,9938	0,9923
2C-4	0,0130421	0,10074	0,9940	0,9913
2C-5	0,0140105	0,196987	0,9950	0,9906

Fonte: O autor.

A Tabela 7 apresenta os resultados obtidos para diferentes redes de duas camadas, de 2C-1Ag até 2C-5Ag. Todas as redes foram treinadas usando a função de ativação Ramp e avaliadas em termos de erro e de ajuste em relação a dois modelos teóricos. Os resultados da função perda para validação e do treinamento estão na figura 18, onde as épocas correspondem ao número de vezes que um mesmo resultado foi utilizado para treinar a rede. A cada época, a rede ajusta seus pesos para tentar melhorar a predição; já o eixo Y (Erro Quadrático Médio) mostra o valor da função de custo (*loss*), ou seja, o quão distantes as previsões da rede estão dos valores reais.

Figura 18 – Resultados de validação e treinamento das redes neurais 2C-1 e 2C-5 na base 5.



Fonte: O autor.

Os resultados obtidos da Validação são próximos ($\sim 0,012$ – $0,014$), mostrando boa generalização independente da rede escolhida. Para os resultados do treinamento, o erro obtido

durante o aprendizado varia um pouco mais (0,086 a 0,25), indicando que algumas redes aprenderam os padrões de forma mais eficiente (ex.: 2C-4Ag com 0,10074). Para os resultados do R^2 -1 e R^2 -2, os coeficientes de determinação foram calculados comparando as saídas das redes com dois modelos teóricos diferentes. Os valores altos ($>0,98$) indicam que a rede conseguiu fazer com que os 15 parâmetros preditos gerassem uma curva próxima ao espectro teórico.

Para os resultados teóricos 1, nos quais se avalia a capacidade das redes neurais em realizar a predição proposta, apresentam-se abaixo os desempenhos correspondentes. São exibidos os resultados da arquitetura 2C-1, que obteve os menores valores de erro percentual, bem como os da arquitetura 2C-5, que apresentou o menor desempenho entre os modelos analisados. Na Tabela 8 são apresentados os valores que correspondem aos erros percentuais estimados para cada rede neural.

Tabela 8 – Erros percentuais dos resultados teóricos 1 das redes 2C-1 e 2C-5.

Gaussiana	Centro		Largura		Area	
	2C-1	2C-2	2C-1	2C-2	2C-1	2C-2
Gaussiana 1	0,738	0,995	9,455	8,036	5,965	3,570
Gaussiana 2	0,105	0,101	0,201	1,140	9,665	12,394
Gaussiana 3	0,150	0,161	2,637	2,182	15,832	14,864
Gaussiana 4	0,100	0,099	7,862	6,842	3,042	3,599
Gaussiana 5	0,086	0,139	1,530	2,266	3,590	2,392

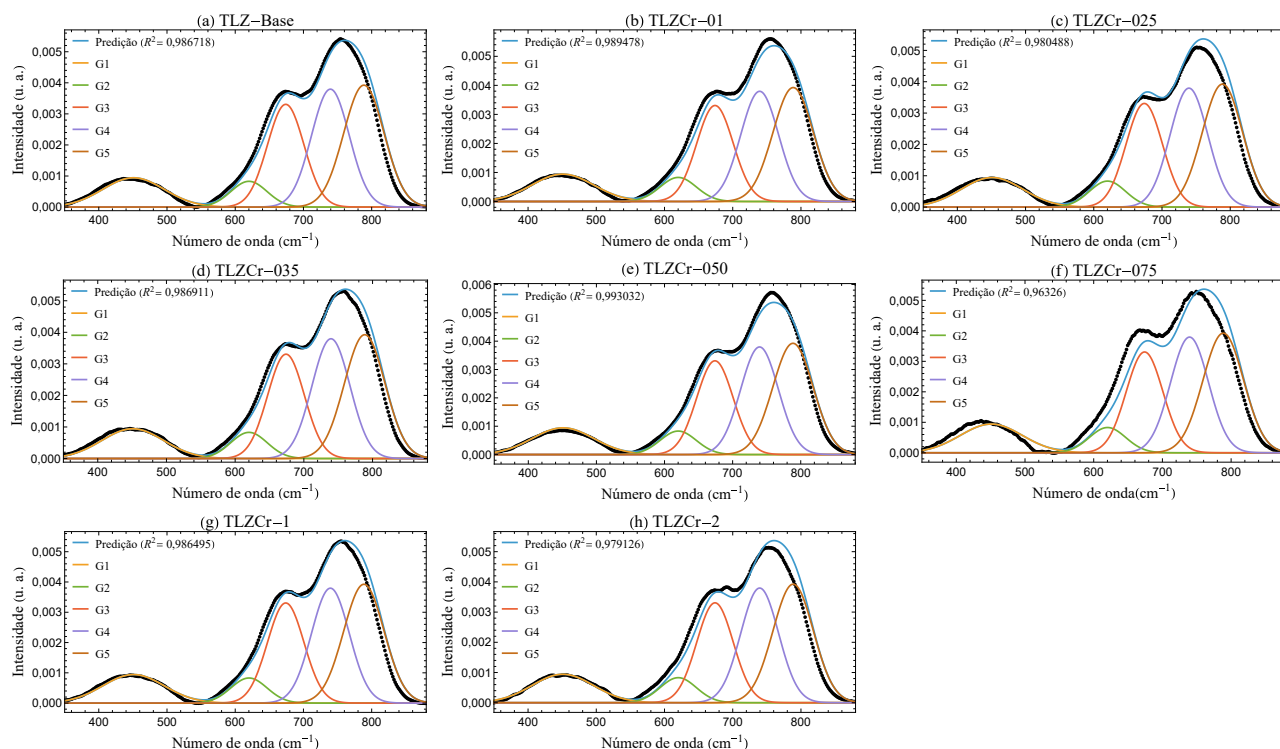
Fonte: O autor.

Conforme apresentado na Tabela 8, verifica-se que, entre todas as métricas de erro percentual analisadas, a variável área apresentou os maiores desvios relativos em ambas as arquiteturas avaliadas. Para a rede 2C-1, que obteve o melhor desempenho geral, o erro percentual associado à predição das áreas alcançou 15,832%, evidenciando maior sensibilidade do modelo a essa variável específica. De forma semelhante, a rede 2C-5, identificada como a de menor desempenho entre as arquiteturas testadas, apresentou erro percentual de 14,864% para a mesma variável. Esses resultados sugerem que a predição das áreas constitui um dos aspectos mais desafiadores para ambas as configurações de rede, indicando a possível necessidade de ajustes adicionais no conjunto de treinamento, no pré-processamento das entradas ou na própria estrutura das redes neurais.

4.2.1 Resultados das curvas experimentais da rede 2C-1

Dessa forma, foi possível fazer a predição para todas as amostras de Cr utilizando as redes 2C-1 a 2C-5. Um exemplo para a predição dos parâmetros da rede 2C-1 de um espectro Raman de uma amostra real de ^{65}Te - ^{15}Li - ^{20}Zn dopado de Cr (TLZCr) é mostrado na Figura 19.

Figura 19 – Resultados experimentais da rede na base 5 da rede MLP 2C-1.



Fonte: O autor.

Ao aplicar a rede sobre um espectro experimental da amostra TLZCr, observou-se uma leve redução no desempenho, com o coeficiente de determinação diminuindo para $R^2 = 0,959439$ para a amostra TLZCr-0,75. Apesar disso, o ajuste das curvas permaneceu visualmente inconsistente e não conseguiu representar adequadamente a forma espectral, similar à obtida pelo método convencional usando *Levenberg-Marquardt* no *Origin*. Tais discrepâncias podem estar relacionadas à presença de ruído, assimetrias experimentais ou interferências ópticas no espectro real, que não estavam presentes nos dados teóricos utilizados no treinamento. Esses resultados reforçam que a arquitetura 2C-1Ag não conseguiu realizar previsões confiáveis, uma vez que os valores obtidos para validação (0,00137585) e treinamento (0,251124) não foram satisfatórios, apresentando-se elevados mesmo diante das variações entre os dados simulados e reais. Então, separamos dois resultados da rede 2C-1Ag. Na Tabela 9, os resultados do erro percentual para a rede 2C-1Ag na predição do centro, largura e área.

O maior erro percentual para os centros foi de aproximadamente 1,341%, para as larguras foi de 10,02% e para as áreas de 14,75%, todos referentes à amostra Cr-0,75. Isso indica que a rede não conseguiu aprender com precisão os padrões espectrais, mesmo com pequenas variações nos parâmetros. A Gaussiana 3, por exemplo, apresentou o maior erro percentual em todos os parâmetros. As demais previsões para os resultados experimentais das redes de duas camadas da Tabela 9 são apresentadas no Apêndice A.

Tabela 9 – Erros percentuais dos parâmetros das gaussianas ajustadas para as redes 2C-1Ag, das amostras TLZCr-0.10 e TLZCr-0,75.

Gaussiana	Centro		Largura		área	
	Cr-0,10	Cr-0,75	Cr-0,10	Cr-0,75	Cr-0,10	Cr-0,75
Gaussiana 1	1,107	1,341	4,737	10,02	5,451	0,477
Gaussiana 2	0,083	0,981	1,270	2,096	5,091	13,98
Gaussiana 3	0,903	0,883	4,080	3,574	3,082	14,75
Gaussiana 4	1,023	1,016	8,411	5,542	3,782	8,492
Gaussiana 5	0,845	0,947	1,454	1,823	3,535	2,578

Fonte: O autor.

4.3 Predição de parâmetros e espectros Raman com normalização

Nesta seção, são apresentados e discutidos os resultados da predição dos parâmetros gaussianos e dos espectros Raman obtidos a partir de bases de dados teóricas normalizadas, construídas conforme o procedimento de pré-processamento descrito na seção 3.4.2.

Novamente, as melhores redes treinadas foram as com 10 mil e 20 mil curvas de entrada. Foram mantidas as mesmas configurações de treinamento com 5000 rounds, a uma taxa de aprendizado de 0,001, em lotes (*batch*) de 64 curvas de 410 valores de entradas e uma camada de saída com 15 neurônios, correspondendo aos 3 parâmetros (centro, largura e área) das 5 gaussianas.

Nas Tabelas 10 e 11 são apresentados os resultados de treinamento e validação das seis melhores redes e seus respectivos valores de R^2 para as predições das curvas teóricas 1 e 2 com as bases de dados de 10 mil e 20 mil curvas, respectivamente.

Tabela 10 – Resultados do treinamento, validação e R^2 das redes normalizadas com duas camadas para base de 10 mil curvas, todas com função de ativação Ramp.

Rede	Neurônios	Validação	Treinamento	R^2 -1	R^2 -2
1C-1Ag	400	0,0000387	0,0000402	0,999	0,999
1C-2Ag	100	0,0000401	0,0000403	0,999	0,999
1C-3Ag	200	0,0000352	0,0000408	0,999	0,999
1C-4Ag	50	0,0000396	0,0000449	0,998	0,995
1C-11Ag	150	0,0000353	0,0000366	0,999	0,999
2C-3Ag	100 e 100	0,0000423	0,0000410	0,999	0,999

Fonte: O autor.

Observa-se, nas Tabelas 10 e 11, que todas as arquiteturas analisadas apresentam erros de MSE de treinamento e validação da ordem de 10^{-5} , com valores muito próximos entre si para cada rede. É importante ressaltar que os valores de MSE das redes sem normalização apresentados na Tabela 9 eram de 10^{-1} , mostrando uma melhora significativa com a normalização. Essa proximidade entre os erros de treino e de validação indica que não há evidências de sobreajuste

Tabela 11 – Resultados do treinamento, validação e R^2 das redes normalizadas com duas camadas para base de 20 mil curvas, todas com função de ativação Ramp.

Rede	Função de ativação	Validação	Treinamento	R^2 -1	R^2 -2
1C-1Ag	400	0,0000361	0,0000388	0,999	0,999
1C-2Ag	100	0,0000383	0,0000373	0,999	0,999
1C-3Ag	200	0,0000348	0,0000368	0,999	0,999
1C-4Ag	50	0,0000414	0,0000390	0,999	0,999
1C-11Ag	150	0,0000353	0,0000366	0,999	0,999
2C-3Ag	100 e 100	0,0000407	0,00004365	0,999	0,999

Fonte: O autor.

(*overfitting*), sugerindo que as redes mantêm boa capacidade de generalização frente às curvas teóricas consideradas. Além disso, os coeficientes de determinação R^2 para as predições das curvas teóricas 1 e 2 são, em praticamente todos os casos, iguais ou superiores a 0,998, o que evidencia um ajuste satisfatório dos modelos às curvas de referência.

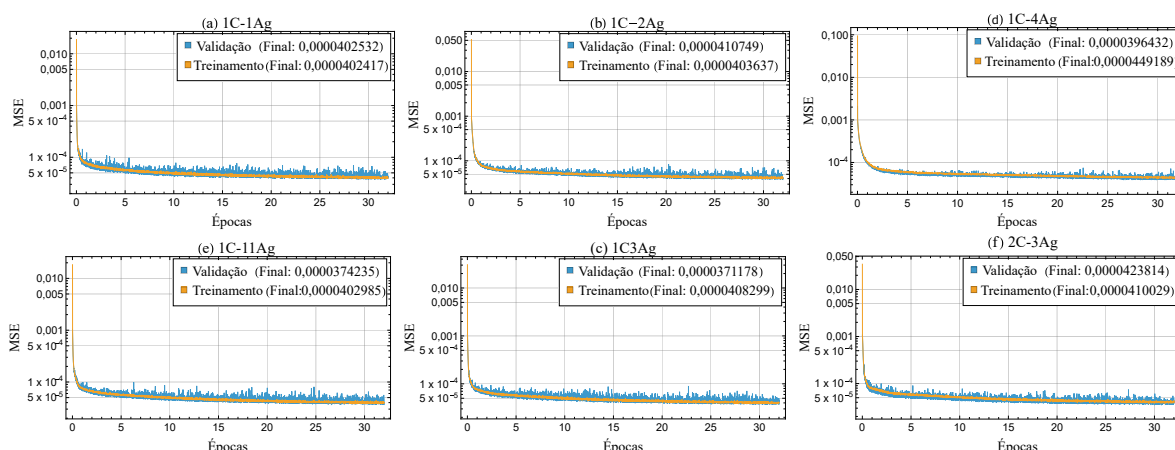
Comparando-se os resultados obtidos com bases de 10 mil e 20 mil curvas, nota-se que o aumento do tamanho da base proporciona apenas uma pequena melhoria nos erros de treinamento e validação. Para a maior parte das arquiteturas, a redução dos erros ao passar de 10 mil para 20 mil curvas é pequena e não altera de forma significativa os valores de R^2 , que já se encontram saturados em torno de 0,999. Mesmo nos casos em que há ligeira melhora (como nas redes 1C-3Ag e 1C-4Ag), o ganho numérico é discreto se comparado ao aumento de tempo e custo computacional associado ao treinamento com 20 mil curvas.

Dessa forma, os resultados indicam que o uso de 10 mil curvas já é suficiente para garantir um desempenho preditivo, com erros baixos e altos valores de R^2 , não havendo ganho substancial que justifique o aumento expressivo do tempo de treinamento ao se dobrar o tamanho da base para 20 mil curvas. Assim, do ponto de vista do compromisso entre custo computacional e qualidade do ajuste, a base de 10 mil curvas mostra-se mais viável.

A Figura 20 mostra os resultados do treinamento e da validação para as redes com a base de 10 mil curvas, onde o eixo X (*Epochs*) indica o número de iterações realizadas durante o treinamento a cada época, os pesos da rede são ajustados para melhorar a precisão da predição. O eixo Y (Erro) representa o valor da função de custo (*loss*), ou seja, a diferença entre as previsões da rede e os valores reais. Dentre as redes analisadas, destacam-se as arquiteturas 1C-2Ag e 1C-11Ag, as quais obtiveram os melhores resultados nos testes realizados, demonstrando maior capacidade de ajuste e precisão em relação às demais configurações. Na figura 20, temos os resultados da validação e dos treinamentos dessas duas redes.

Podemos observar que os resultados obtidos tanto na validação quanto no treinamento das redes analisadas apresentam uma tendência de diminuição contínua. Esse comportamento indica que as redes neurais estão conseguindo ajustar de forma adequada seus pesos, conseguindo realizar predições precisas dos espectros Raman. Além disso, a convergência dos erros em

Figura 20 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 de 10 mil curvas.



Fonte: O autor.

ambos os conjuntos sugere que não houve ocorrência de *overfitting* durante o processo de treinamento, evidenciando que as redes mantêm uma boa capacidade de generalização para dados não utilizados na aprendizagem. Portanto, os resultados obtidos reforçam a eficácia da normalização dos parâmetros para o treinamento das redes.

Na Tabela 12 são apresentados os resultados teóricos 1 referentes às redes normalizadas, nos quais se examina a capacidade dos modelos neurais em reproduzir adequadamente as variáveis que foram utilizadas para criação da curva teórica. Destacam-se os desempenhos obtidos pela arquitetura 1C-11Ag, identificada como a configuração com maior precisão preditiva, e pela arquitetura 2C-3Ag, que apresentou o menor desempenho entre as alternativas avaliadas.

Tabela 12 – Erros percentuais dos resultados teóricos 1 normalizados das redes 1C-11Ag (melhor) e 2C-3Ag (pior).

Gaussiana	Centro		Largura		Área	
	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag
Gaussiana 1	0,183	0,158	0,823	0,539	0,291	0,169
Gaussiana 2	0,081	0,079	0,534	0,752	0,282	0,888
Gaussiana 3	0,039	0,059	0,522	0,502	0,153	0,094
Gaussiana 4	0,081	0,015	0,602	0,554	0,650	0,502
Gaussiana 5	0,185	0,129	0,600	0,236	0,826	1,105

Fonte: O autor.

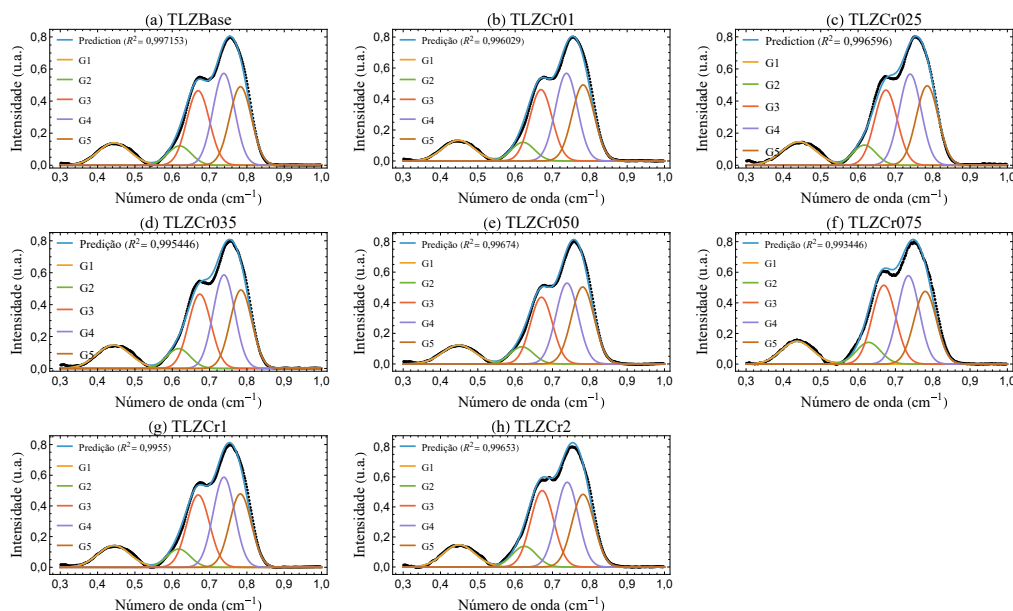
Observa-se que os erros percentuais obtidos nas previsões teóricas das redes neurais foram todos menores que 1%. Esse resultado indica que ambas as arquiteturas apresentaram elevada capacidade de generalização no contexto avaliado, reproduzindo com precisão os valores teóricos de referência, os quais não correspondem a dados experimentais. Abaixo temos os resultados da predição da cada rede.

4.3.1 Resultados das curvas experimentais da rede 1C-11Ag base 5

Na rede 1C-11Ag aplicada à base 5, os valores de validação e treinamento apresentaram um aumento em comparação às outras redes. No caso da rede 1C-11Ag, os valores obtidos para validação (0,0000374235) e treinamento (0,0000402985) apresentaram-se satisfatórios. Os resultados da predição dos espectros Raman, para todas as amostras de vidros teluretos dopados com Cr, indicaram que este modelo, assim como as demais redes que são apresentadas no Apêndice 6, foi capaz de realizar predições com elevada precisão.

A Figura 21 apresenta os espectros Raman experimentais e as curvas reconstruídas por convolução gaussiana a partir dos 15 parâmetros preditos (centro, largura e área das cinco bandas) para todas as amostras TLZ consideradas. Em todos os casos, observa-se que a curva obtida pela rede 1C-11Ag acompanha de forma muito fiel o perfil experimental, inclusive nas regiões de maior sobreposição entre as bandas, resultando em valores de R^2 superiores a 0,99 para todas as composições. As contribuições individuais das gaussianas G1–G5 permitem visualizar como a rede neural separa e recompõe as diferentes bandas que compõem o espectro total. Mesmo para amostras com perfis mais assimétricos ou com variações de intensidade entre as bandas, a soma das cinco gaussianas reproduz adequadamente a forma global do espectro, sem oscilações artificiais ou discrepâncias sistemáticas.

Figura 21 – Resultados do espectro predito pela rede 1C-11Ag treinada na base 5 para todos os resultados experimentais.



Fonte: O autor.

Na Tabela 13 são apresentados os dois piores casos de R^2 dos ajustes obtidos para a rede 1C-11Ag, correspondentes às amostras TLZCr-0,35 e TLZCr-0,75. Os valores listados representam as diferenças percentuais entre os parâmetros gaussianos preditos pela rede neural e

aqueles obtidos previamente pelo ajuste convencional realizado no *software Origin*, isto é, não se tratam de “erros absolutos” em relação a um valor experimental exato, mas sim de discrepâncias relativas entre dois procedimentos distintos de ajuste.

Observa-se que, para algumas bandas, especialmente nas gaussianas centrais (G2 e G3) e nas contribuições de área (G2, G3 e G4), as diferenças percentuais podem atingir valores da ordem de 15–20%. Esse comportamento é esperado nas regiões de maior sobreposição entre as bandas, nas quais o problema de deconvolução se torna mais mal condicionado e pequenas variações de forma espectral podem ser atribuídas a diferentes combinações de centro, largura e área. Além disso, no ajuste realizado no *Origin* foi necessário fixar diversos parâmetros para estabilizar a deconvolução, o que introduz uma forte dependência do resultado em relação às escolhas iniciais do usuário e pode levar a soluções artificialmente restringidas.

Por outro lado, mesmo nesses dois casos extremos, a predição experimental da rede 1C-11Ag apresenta coeficientes de determinação R^2 elevados e uma excelente reprodução visual dos espectros Raman, conforme mostrado na Figura 21. Isso indica que as diferenças percentuais listadas na Tabela 13 refletem, sobretudo, a discrepância entre dois esquemas de ajuste (rede neural e *Origin*) e não uma falha da rede em representar a resposta experimental. Considerando que a MLP ajusta simultaneamente todos os parâmetros a partir do espectro completo, sem necessidade de fixar valores a priori, a predição fornecida pela rede neural pode ser considerada mais robusta e fisicamente consistente do que o ajuste ponto a ponto obtido pelo método convencional no *Origin*.

Tabela 13 – Erros percentuais dos parâmetros das gaussianas ajustadas para as redes 1C-11Ag, das amostras TLZCr-0.35 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-0,35	Cr-0,75	Cr-0,35	Cr-0,75	Cr-0,35	Cr-0,75
Gaussiana 1	0,224	1,904	4,032	2,762	3,220	1,118
Gaussiana 2	0,309	2,258	17,897	19,463	14,034	19,858
Gaussiana 3	0,855	0,158	16,758	12,947	12,452	3,661
Gaussiana 4	0,895	0,408	3,033	5,542	11,428	16,153
Gaussiana 5	0,259	0,159	0,869	1,823	0,620	16,248

Fonte: O autor.

Na rede 1C-11Ag, os maiores desvios também foram observados na amostra TLZCr035, com valores de 17,897% na largura da gaussiana 2 e 19,858% na área da gaussiana 3. Para a mesma rede, mas agora na base 6, os resultados da validação e do treinamento foram muito próximos da rede 1C-11Ag na base 5 para os resultados experimentais da espectroscopia Raman, os maiores desvios foram nas amostras TLZCr035 e TLZCr075, e os resultados foram menos precisos em relação à base 5, com erro de 18,617% na largura gaussiana 3 e de 17,5082% na área para a amostra TLZCr035.

4.4 Redes neurais aplicadas a espectros normalizados com adição de 1% de ruído

Com o método de normalização de todos os parâmetros, obtiveram-se bons resultados para os dados teóricos. Entretanto, os ajustes aos espectros experimentais ainda apresentam certas discrepâncias. Isso pode estar relacionado, principalmente, à presença de erros aleatórios nas medidas e ao fato de a linha de base experimental não ser perfeitamente ajustada. Além disso, os resultados empregados na seção anterior para o treinamento das redes não consideravam ruído nem linha de base. Assim, o próximo passo consiste em incluir ruído nas bases de dados, a fim de avaliar se essa modificação prejudica ou não os processos de treinamento e validação e se contribui para aprimorar a predição dos parâmetros utilizados no ajuste dos espectros experimentais. Essa análise é apresentada nesta seção.

Para testar melhor a rede e aproximar as condições experimentais, foi adicionado ruído aleatório de 1% aos espectros simulados. Esse procedimento tem como objetivo melhorar o treinamento, evitando que a rede se ajuste apenas aos valores ideais e sofra *overfitting*. Ao introduzir pequenas variações nos dados, a rede é incentivada a aprender a lidar com as alterações inerentes aos espectros reais, aproximando seus resultados dos obtidos experimentalmente por espectroscopia Raman. Dessa forma, foram geradas novas bases de dados, com os mesmos intervalos das redes normalizadas da Tabela 5, porém com a adição de ruído. Novamente, foram treinadas redes com 10 mil e 20 mil curvas. O treinamento da rede neural foi realizado nas mesmas configurações, com 5.000 *rounds*, em lotes de 64 curvas.

Nas Tabelas 14 e 15 são apresentados os resultados de treinamento e validação das 6 melhores redes e seus respectivos valores de R^2 para as predições das curvas teóricas 1 e 2 com as bases de dados de 10 mil e 20 mil curvas, respectivamente.

Tabela 14 – Resultados do treinamento, validação e R^2 das redes normalizadas e 1% de ruído com duas camadas para base de 10 mil curvas, todas com função de ativação Ramp.

Rede	Neurônios	Validação	Treinamento	R^2 -1	R^2 -2
1C-1Ag	400	0,0000387	0,0000402	0,999	0,999
1C-2Ag	100	0,0000401	0,0000403	0,999	0,999
1C-3Ag	200	0,0000352	0,0000408	0,999	0,999
1C-4Ag	50	0,0000396	0,0000449	0,998	0,995
1C-11Ag	150	0,0000353	0,0000366	0,999	0,999
2C-3Ag	100 e 100	0,0000445	0,0000451	0,999	0,997

Fonte: O autor.

A partir das Tabelas 14 e 15, observa-se que a inclusão de 1% de ruído nas bases de dados não degradou o desempenho das redes neurais. Os erros de treinamento e validação permanecem da ordem de 10^{-5} para todas as arquiteturas analisadas, com valores muito próximos entre si, o que indica que a rede continua capaz de aprender o mapeamento entre espectros e parâmetros gaussianos, mesmo na presença de flutuações aleatórias nos dados. Além disso, os coeficientes

Tabela 15 – Resultados do treinamento, validação e R^2 das redes normalizadas e 1% de ruído com duas camadas para base de 20 mil curvas, todas com função de ativação Ramp.

Rede	Função de ativação	Validação	Treinamento	R^2 -1	R^2 -2
1C-1Ag	400	0,0000361	0,0000388	0,999	0,999
1C-2Ag	100	0,0000383	0,0000373	0,999	0,999
1C-3Ag	200	0,0000348	0,0000368	0,999	0,999
1C-4Ag	50	0,0000414	0,0000390	0,999	0,999
1C-11Ag	150	0,0000353	0,0000366	0,999	0,999
2C-3Ag	100 e 100	0,0000425	0,0000419	0,999	0,999

Fonte: O autor.

de determinação R^2 para as curvas teóricas 1 e 2 mantêm-se em torno de 0,999 em praticamente todos os casos.

Outro aspecto importante é que a proximidade entre os erros de treinamento e validação continua indicando a ausência de sobreajuste (*overfitting*), mesmo com a base ruidosa. Isso sugere que a perturbação de 1% atua como uma espécie de regularização implícita, incentivando a rede a aprender, em vez de memorizar detalhes específicos dos espectros ideais. Em outras palavras, a rede passa a ser treinada em um cenário mais próximo das condições experimentais reais, sem perda significativa de desempenho nas previsões teóricas.

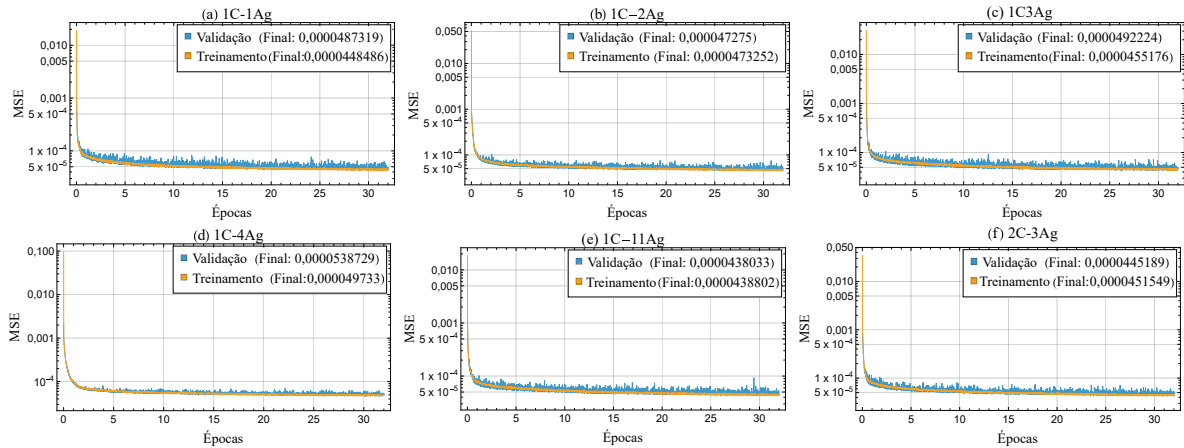
Comparando os resultados obtidos com 10 mil e 20 mil curvas ruidosas, nota-se novamente que o aumento do tamanho da base produz apenas melhorias marginais nos erros e nos valores de R^2 . Assim como no caso sem ruído, a base de 10 mil curvas mostra-se suficiente para garantir um ajuste de alta qualidade, de modo que a ampliação para 20 mil curvas não se justifica quando se leva em conta o custo computacional adicional. Dessa forma, pode-se concluir que a adição de 1% de ruído não compromete o treinamento das redes normalizadas e, ao mesmo tempo, contribui para tornar o modelo mais aderente às condições experimentais, sem a necessidade de aumentar o tamanho da base de dados.

A Figura 22 exibe a evolução do treinamento e da validação das redes com a base 5 de 10 mil curvas. Entre as arquiteturas analisadas, as configurações 1C-11Ag e 2C-3Ag apresentaram desempenho superior, com tendência de diminuição e menores valores finais de MSE.

A Figura 23 exibe a evolução do treinamento e da validação das redes com a base 6 de 20 mil curvas. Novamente, as configurações 1C-11Ag e 2C-3Ag apresentaram melhor desempenho, com uma tendência de diminuição e menores valores finais de MSE.

Pode-se notar que os resultados obtidos, tanto no treinamento quanto na validação das redes analisadas, apresentam uma tendência de redução constante. Esse comportamento indica que as redes neurais estão ajustando seus pesos de maneira eficiente, permitindo realizar previsões precisas dos espectros Raman. Além disso, a convergência dos erros em ambos os conjuntos sugere que não ocorreu *overfitting* durante o treinamento, evidenciando que as redes

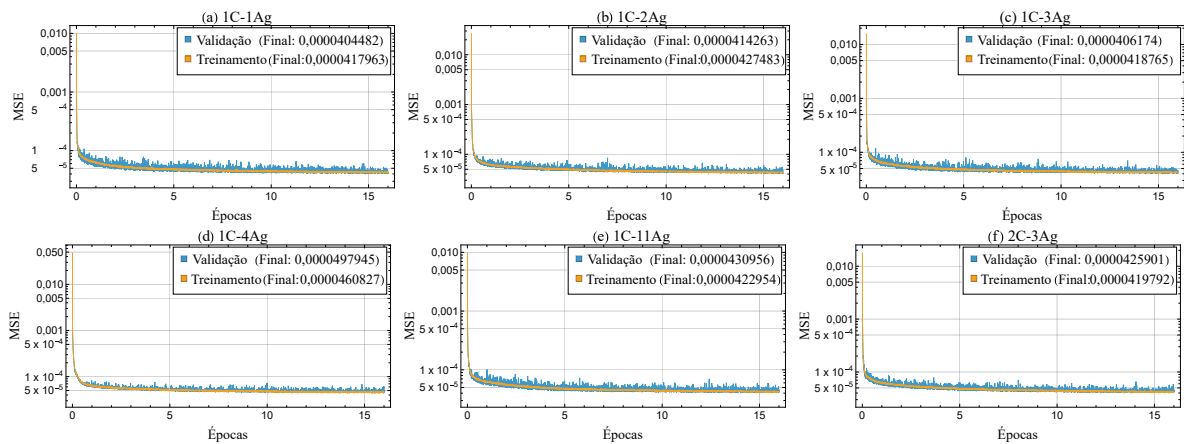
Figura 22 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com 1% de ruído.



Fonte: O autor.

mantêm uma boa capacidade de generalização para dados não utilizados na aprendizagem.

Figura 23 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com 1% de ruído.



Fonte: O autor.

Para os resultados teóricos 1 referentes às redes normalizadas, agora avaliadas sob a condição de 1% de ruído adicionado aos dados, podemos ver novamente a capacidade dos modelos neurais em reproduzir adequadamente as variáveis utilizadas para produzir os espectros, mesmo que estas não contenham ruídos. Observa-se que as arquiteturas 1C-11Ag e 2C-3Ag mantiveram o melhor desempenho preditivo. Os valores apresentados na Tabela 16 correspondem aos erros percentuais calculados para cada arquitetura.

Verifica-se que os erros percentuais obtidos pelas arquiteturas 1C-11Ag e 2C-3Ag, mesmo na presença de 1% de ruído nos dados de entrada, permaneceram, no máximo, em 1%. Esse comportamento evidencia que ambas as redes mantiveram uma elevada capacidade de generalização no cenário analisado, conseguindo reproduzir com alta precisão os valores teóricos de referência.

Nas seções seguintes, apresentamos os resultados para as duas melhores redes com

Tabela 16 – Erros percentuais dos resultados teóricos 1 normalizados com 1% de ruído das redes 1C-11Ag e 2C-3Ag.

Gaussiana	Centro		Largura		Área	
	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag
Gaussiana 1	0,096	0,059	0,721	0,428	0,291	0,400
Gaussiana 2	0,132	0,158	0,725	0,150	0,282	0,618
Gaussiana 3	0,087	0,047	0,450	0,571	0,153	0,228
Gaussiana 4	0,118	0,035	0,405	0,813	0,650	0,471
Gaussiana 5	0,059	0,044	0,182	0,890	0,826	1,003

Fonte: O autor.

resultados de treinamento: as configurações 1C-11Ag e 2C-3Ag. As demais redes que são apresentadas no Apêndice 6.

4.4.1 Resultados das curvas experimentais da rede 1C-11Ag com 1% de ruído

A rede 1C-1Ag, aplicada à base 5, revelou-se uma das mais promissoras, apresentando erros de validação e treinamento extremamente baixos, iguais a 0,0000438033 e 0,0000438802, respectivamente. A Figura 24 apresenta os espectros a partir da convolução das gaussianas obtidas pela predição dos 15 parâmetros. Nota-se uma evolução considerável em relação aos ajustes anteriores, conseguindo realizar uma boa predição para todas as amostras. Na Tabela 17 apresentam-se os resultados das diferenças percentuais entre os valores obtidos pelo *Origin* e pelo *Wolfram* para duas amostras de vidros teluretos que obtiveram, respectivamente, a melhor predição TLZCr050 e a pior predição TLZCr075, de modo a evidenciar os extremos de desempenho da rede no ajuste dos espectros Raman.

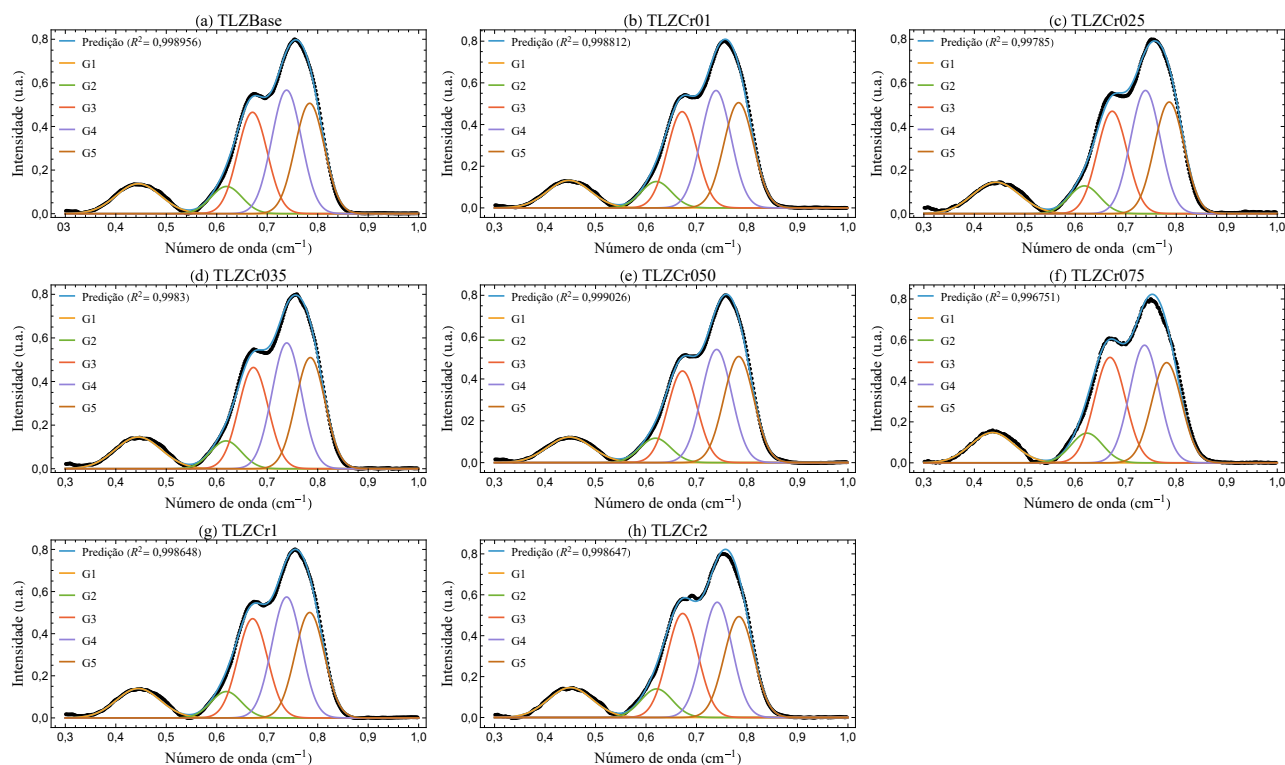
Tabela 17 – Parâmetros das gaussianas ajustadas para as redes 1C-11Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-01	Cr-075	Cr-01	Cr-075	Cr-01	Cr-075
Gaussiana 1	0,313	2,012	1,716	4,408	7,615	0,461
Gaussiana 2	0,604	1,445	14,779	15,071	8,390	17,808
Gaussiana 3	0,637	0,093	13,911	10,848	5,936	1,714
Gaussiana 4	1,075	0,698	0,148	12,385	0,788	16,547
Gaussiana 5	0,310	0,031	0,140	1,756	19,855	14,683

Fonte: O autor.

Os resultados da rede 1C-11Ag para a amostra TLZCr01 evidenciaram um erro máximo de 14,779% na largura da gaussiana 2 e de 19,855% na área da gaussiana 5. Apesar desses valores relativamente elevados em relação às demais gaussianas, a rede conseguiu realizar a predição de maneira satisfatória, demonstrando boa capacidade de ajuste para esta amostra. Por

Figura 24 – Resultados da predição 1C-11Ag na base 5.



Fonte: O autor.

outro lado, para a amostra TLZCr075, o maior erro registrado foi de 15,071% na largura e de 17,808% na área, ambos referentes à gaussiana 2.

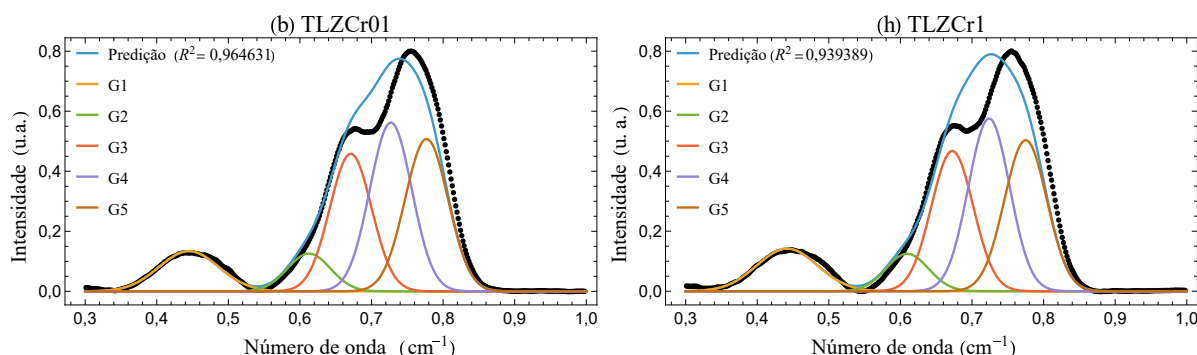
Cabe destacar que, diferentemente do ajuste convencional realizado no *Origin*, no qual foi necessário fixar diversos parâmetros para estabilizar a deconvolução, a rede neural ajusta simultaneamente todos os parâmetros a partir do espectro completo. Nesse sentido, mesmo quando as diferenças percentuais entre os dois métodos são mais elevadas, a consistência global do ajuste espectral e os altos valores de R^2 indicam que a predição fornecida pela rede neural é, em muitos casos, mais confiável do que aquela obtida pelo método convencional.

Para a rede 1C-11Ag na base 6, os valores de validação e treinamento apresentaram-se bastante próximos, 0,0000430956 e 0,0000422954, respectivamente. Essa proximidade indica que não houve ocorrência de *overfitting*, além de mostrar que os resultados foram bastante semelhantes aos obtidos na base 5, sem diferenças significativas. No que se refere às amostras, a rede também apresentou bom desempenho preditivo na base 6. Para a amostra TLZCr050, o maior erro observado foi de 15,201% na largura da gaussiana 3 e de 20,399% na área da gaussiana 5. Já para a amostra TLZCr075, os resultados permaneceram próximos aos encontrados na base 5, com erro máximo de 16,333% na largura e de 18,061% na área, ambos referentes à gaussiana 5. Esses resultados reforçam a consistência do modelo, demonstrando que sua capacidade de predição se manteve estável mesmo com a mudança da base de dados.

4.4.2 Resultados das curvas experimentais da rede 2C-3Ag com 1% de ruído

Para a arquitetura 2C-3Ag, os resultados obtidos não se mostraram satisfatórios do ponto de vista da aplicação aos dados experimentais, mesmo apresentando bons indicadores numéricos nas bases teóricas. Tanto na base 5 quanto na base 6 foram obtidos valores de validação (0,0000445189) e treinamento (0,0000451549) muito próximos, o que, à primeira vista, sugeriria um ajuste adequado. No entanto, ao se aplicar a rede aos espectros Raman experimentais, a predição não manteve o mesmo nível de concordância observado para as curvas simuladas, resultando em ajustes menos consistentes. A Figura 25 ilustra esse comportamento para duas amostras representativas, evidenciando que, apesar da boa qualidade global do ajuste teórico, a rede não reproduz com a mesma precisão os detalhes presentes nas medidas reais.

Figura 25 – Resultados da predição 2C-3Ag na base 5.



Fonte: O autor.

Esse contraste entre o excelente desempenho nas bases teóricas e a perda de qualidade nas predições experimentais é indicativo de um comportamento típico de sobreajuste às características específicas dos dados simulados: o modelo aprende muito bem o padrão ideal gerado numericamente, mas não generaliza para as variações estruturais, ruídos instrumentais e imperfeições de linha de base presentes nos espectros reais. Tal limitação pode estar associada à maior complexidade da rede 2C-3Ag em relação ao volume e à diversidade dos dados disponíveis, o que favorece o ajuste excessivo aos exemplos de treinamento. Além disso, a relativa semelhança entre os perfis espectrais das diferentes amostras reduz a variabilidade efetiva do conjunto de dados, restringindo o potencial da arquitetura de duas camadas em representar adequadamente novas instâncias do sistema vítreo analisado.

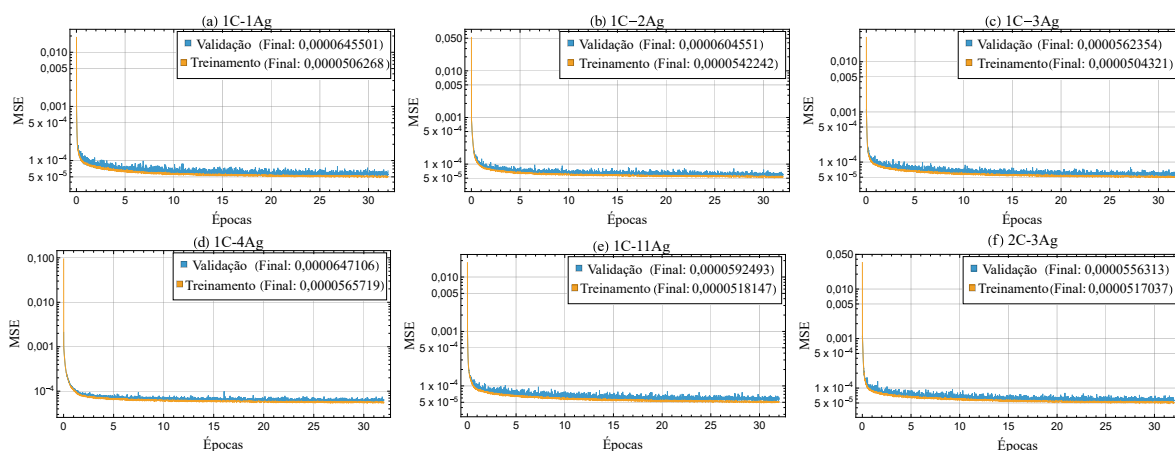
4.5 Redes neurais aplicadas a espectros normalizados com adição de 2% de ruído

Para identificar a avaliação da rede e simular ainda mais de perto as condições experimentais, foi adicionado ruído aleatório de 2% aos espectros simulados. Esse nível de ruído

mais elevado, em comparação ao de 1%, tem como finalidade testar a capacidade de generalização da rede diante de variações mais expressivas nos dados. Com a introdução desse ruído adicional, espera-se que a rede não apenas evite o sobreajuste (*overfitting*), mas também desenvolva uma maior tolerância a oscilações presentes em espectros obtidos experimentalmente por espectroscopia Raman.

Assim como no caso anterior, foram utilizadas as mesmas redes previamente normalizadas, preservando todas as configurações originais. A diferença está apenas na área do ruído aplicado, que neste caso corresponde a 2%. A Figura 26 apresenta a evolução do treinamento e da validação das redes.

Figura 26 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com 2% de ruído.



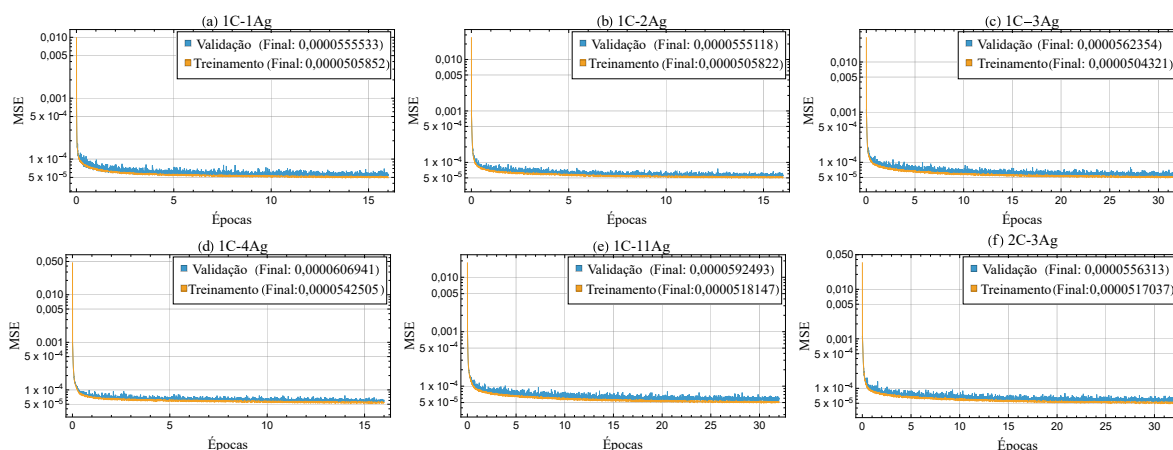
Fonte: O autor.

Dentre as arquiteturas avaliadas, as configurações 1C-11Ag e 2C-3Ag apresentaram os melhores resultados, demonstrando maior capacidade de ajuste e maior precisão em comparação com as demais. Pode-se observar que os resultados obtidos, tanto no treinamento quanto na validação das redes analisadas, apresentam uma tendência de redução contínua dos erros. Esse comportamento indica que as redes neurais ajustaram seus pesos de forma eficiente, permitindo realizar previsões consistentes dos espectros Raman. Além disso, a convergência entre os erros de treinamento e validação sugere a ausência de *overfitting*, evidenciando que os modelos preservam uma boa capacidade de generalização para dados não utilizados durante a aprendizagem.

Em comparação com a base contendo 1% de ruído, observa-se que os valores de erro de treinamento e validação foram ligeiramente superiores quando foi aplicado 2% de ruído. Esse comportamento indica que a rede apresentou maior dificuldade em aprender a partir de dados mais ruidosos, o que já era esperado em função do aumento da complexidade do conjunto de entrada. Para a base 6, verificou-se o mesmo padrão, conforme apresentado nos resultados na figura 51.

Ainda assim, os resultados confirmam a eficácia das redes na representação das características espectrais, demonstrando que os modelos são capazes de identificar padrões presentes

Figura 27 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com 2% de ruído.



Fonte: O autor.

nos espectros Raman simulados.

Para os resultados teóricos 1 referentes às redes normalizadas, agora avaliadas sob a condição de 2% de ruído adicionado aos dados, verifica-se que os modelos neurais continuam capazes de reproduzir adequadamente as variáveis alvo, demonstrando o aumento do nível de perturbação nos dados de entrada. A arquitetura 1C-11Ag manteve o melhor desempenho preditivo, enquanto a arquitetura 2C-3Ag apresentou o menor desempenho entre as configurações analisadas. Os valores apresentados na Tabela 38 correspondem aos erros percentuais estimados para cada arquitetura, permitindo uma comparação direta da eficiência preditiva dos modelos mesmo sob condições mais desafiadoras de ruído.

Tabela 18 – Erros percentuais dos resultados teóricos 1 normalizados com 2% de ruído das redes 1C-11Ag e 2C-3Ag.

Gaussiana	Centro		Largura		Área	
	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag
Gaussiana 1	0,049	0,0008	0,367	0,784	0,338	0,010
Gaussiana 2	0,027	0,531	0,334	0,151	0,474	0,676
Gaussiana 3	0,100	0,523	0,587	0,572	0,758	0,479
Gaussiana 4	0,065	0,602	0,176	0,817	1,134	0,975
Gaussiana 5	0,001	0,598	0,285	0,884	1,498	1,681

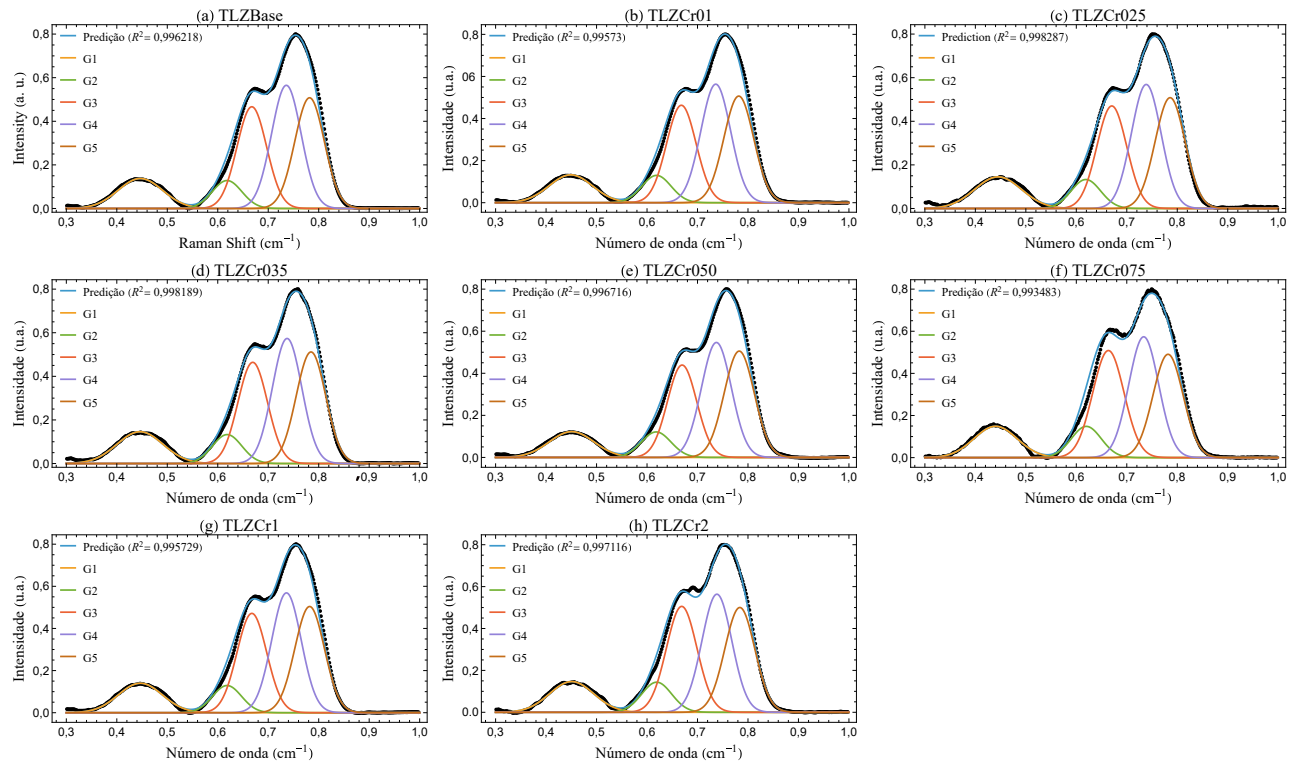
Fonte: O autor.

Verifica-se que os erros percentuais obtidos pelas arquiteturas 1C-11Ag e 2C-3Ag, mesmo na presença de 2% de ruído nos dados de entrada, apresentaram um ligeiro aumento em relação aos resultados teóricos com 1% de ruído, agora atingindo valores próximos a 2%. Esse comportamento demonstra que ambas as redes mantiveram elevada capacidade de generalização no cenário analisado, reproduzindo com precisão satisfatória os valores teóricos de referência os quais, convém destacar, não correspondem a dados experimentais. A seguir, são apresentados os resultados das predições obtidas para cada arquitetura avaliada sob a condição de 2% de ruído.

4.5.1 Resultados das curvas experimentais da rede 1C-11Ag com 2% de ruído

Nesta seção são apresentados os resultados obtidos para a melhor arquitetura após a inclusão de ruído aleatório de 2%, a rede 1C-11Ag. As demais configurações avaliadas encontram-se no Apêndice 6. Entre todas as redes testadas, a 1C-11Ag manteve-se como uma das mais promissoras, apresentando erros de validação de 0,0000592493 e de treinamento de 0,0000518147. Esses valores, próximos entre si, indicam estabilidade no aprendizado e ausência de indícios de *overfitting*. O melhor desempenho foi observado para a amostra TLZBase, enquanto a maior discrepância ocorreu para a amostra TLZCr2, conforme ilustrado na figura 28.

Figura 28 – Resultados da predição 1C-11Ag na base 5.



Fonte: O autor.

A Tabela 42 apresenta os parâmetros ajustados para duas amostras representativas (TLZBase e TLZCr2), de modo a evidenciar os limites superior e inferior de desempenho da rede após a adição de ruído.

Os resultados demonstram que a rede 1C-11Ag, aplicada à base 5 com ruído adicionado, foi capaz de reproduzir adequadamente os espectros Raman para todas as amostras analisadas. Para a amostra TLZBase, os maiores desvios observados foram de 9,382% na largura da Gaussiana 3 e de 20,311% na área da Gaussiana 2. No caso da amostra TLZCr2, o erro máximo registrado foi de 12,219% na largura da Gaussiana 3 e de 23,432% na área da Gaussiana 2. Apesar desses valores, os coeficientes de determinação R^2 permaneceram elevados, o que evidencia

Tabela 19 – Parâmetros das gaussianas ajustadas para a rede 1C-11Ag nas amostras TLZBase e TLZCr-2, incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	TLZBase	Cr-2	TLZBase	Cr-2	TLZBase	Cr-2
G1	0,300	0,159	0,00009	2,644	2,467	1,193
G2	0,559	0,673	8,540	9,479	20,311	23,432
G3	0,077	0,102	9,382	12,219	2,775	16,887
G4	0,263	0,875	2,133	0,504	3,628	9,109
G5	0,014	0,235	1,251	0,121	10,310	12,629

Fonte: O autor.

a capacidade da rede em manter ajustes consistentes mesmo diante de perturbações introduzidas nos dados simulados.

Em relação às análises anteriores, nas quais se utilizou ruído de apenas 1%, observa-se que a adição de ruído mais intenso não comprometeu significativamente a capacidade de generalização da rede. O padrão geral de desempenho foi mantido: as amostras TLZBase (menos dopada) continuam apresentando os melhores ajustes, enquanto amostras com maior concentração de Cr, como TLZCr2, seguem mais sensíveis às variações espectrais e exibem maiores discrepâncias. Entretanto, diferentemente dos casos com 1% de ruído, os desvios percentuais tornam-se ligeiramente mais acentuados, como observado nas áreas das gaussianas 2 e 3, o que era esperado, dada a maior complexidade dos dados.

Quando aplicada à base 6, a rede apresentou erros de validação (0,0000578309) e treinamento (0,0000503308) ainda menores do que os obtidos para a base 5, reforçando a ausência de *overfitting* e demonstrando a estabilidade do modelo. Entre as amostras avaliadas nessa base, a TLZCr035 mostrou o melhor ajuste, com desvios máximos de 12,898% na largura da Gaussiana 3 e 13,941% na área da Gaussiana 5. Em contrapartida, a amostra TLZCr1 apresentou o desempenho menos preciso, com erros de 13,041% na largura da Gaussiana 2 e 15,372% em sua área.

De modo geral, a análise comparativa mostra que, mesmo com a inclusão de ruído adicional, a rede 1C-11Ag manteve desempenho robusto e reproduziu adequadamente os espectros Raman experimentais. A leve piora observada em algumas bandas é compatível com a maior complexidade dos dados ruidosos, não comprometendo a confiabilidade global da predição. Esses resultados reforçam a adequação da arquitetura selecionada para problemas reais de ajuste espectral.

De maneira geral, a comparação direta entre os resultados com 1% e 2% de ruído evidencia que o aumento da perturbação introduzida nos dados simulados levou a uma leve piora no ajuste dos parâmetros, indicando que, embora a rede mantenha boa capacidade de generalização, sua performance é sensivelmente degradada à medida que o nível de ruído se afasta das condições reais observadas nos espectros experimentais.

4.6 Redes neurais aplicadas a espectros normalizados com adição de 5% de ruído

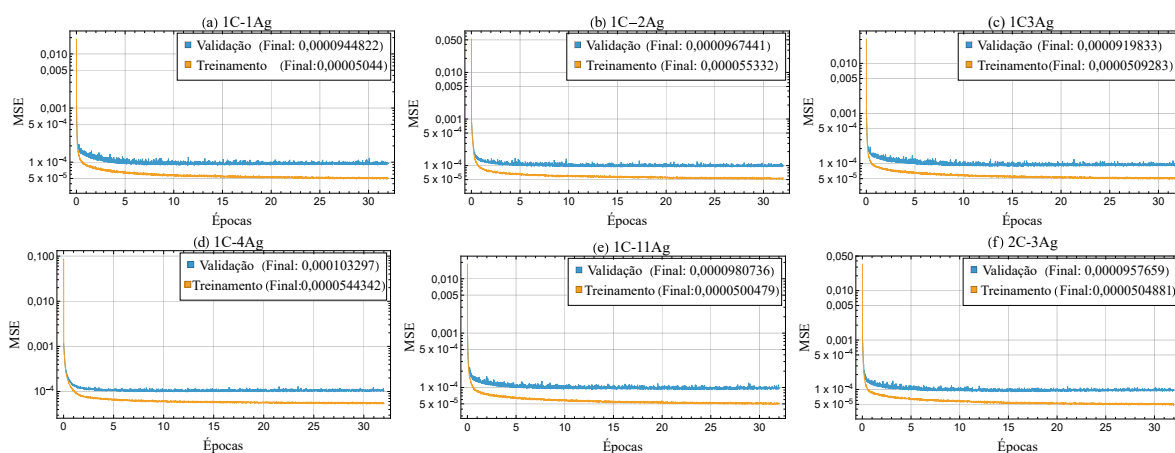
Para avaliar o desempenho da rede em condições ainda mais próximas da realidade experimental, foi adicionado ainda mais ruído aleatório de 5% aos espectros simulados. Esse nível de ruído, mais elevado em relação aos testes anteriores com 1% e 2%, tem como objetivo verificar a capacidade de generalização da rede diante de variações significativas nos dados, reforçando a análise.

Com a inserção desse nível adicional de ruído, espera-se que a rede apresente resultados menos precisos e com maior margem de erro, uma vez que se está intensificando as incertezas no conjunto de dados, o que também pode aumentar a probabilidade de sobreajuste (*overfitting*). Por outro lado, essa condição possibilita avaliar a capacidade do modelo em manter a robustez diante das flutuações típicas dos espectros experimentais de espectroscopia Raman. Assim, torna-se viável investigar se o desempenho da rede se sustenta mesmo em cenários de simulação mais desafiadores.

Para a construção da base de dados, foram empregados dois conjuntos de espectros teóricos utilizados como referência para a validação da metodologia. Mantiveram-se as mesmas configurações das bases Exp1 e Exp2, assim como o procedimento de normalização aplicado, sendo estas compostas por 10.000 e 20.000 curvas de entrada, respectivamente. Dessa forma, assegurou-se que a rede mantivesse as mesmas características adotadas nos testes anteriores.

Os resultados foram feitos para as mesmas redes 1C-1Ag e 2C-3Ag. Todas as arquiteturas foram avaliadas sob condições equivalentes, com a adição de 5% de ruído aos dados. Então, seguindo na figura 29, temos os resultados da validação e treinamento para cada rede analisada na base 5.

Figura 29 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com 5% de ruído.



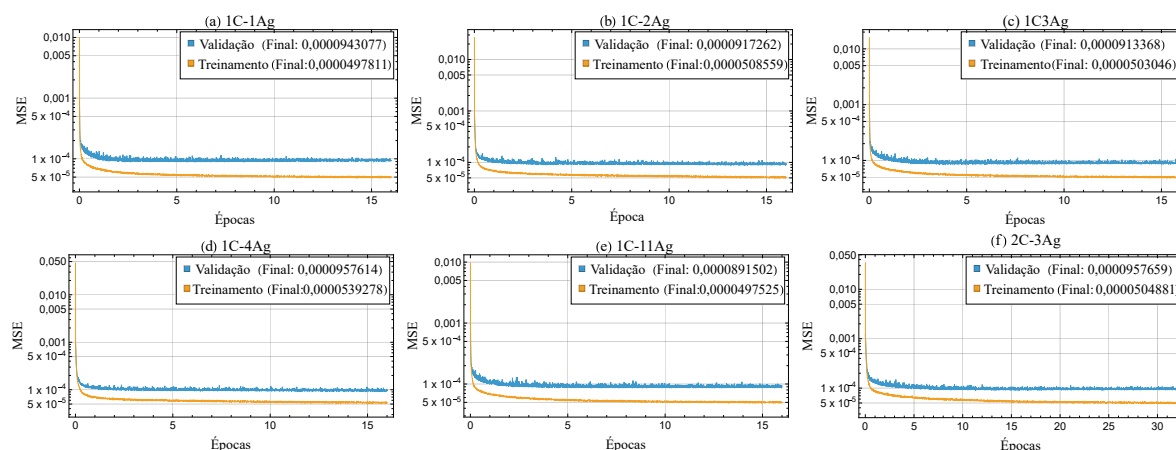
Fonte: O autor.

Nos resultados apresentados na Figura 30, observa-se a evolução do erro quadrático médio (MSE) para os conjuntos de treinamento e validação em diferentes arquiteturas de redes

neurais. Após a adição de 5% de ruído aos dados de entrada, verifica-se um aumento consistente nos valores finais de erro em todas as redes testadas, comportamento este intuitivamente esperado, dada a maior incerteza introduzida nos dados. Esse aumento indica que a presença de ruído dificultou o processo de aprendizagem e a capacidade de generalização dos modelos, refletindo-se em um desempenho preditivo inferior. Ainda assim, nota-se que, após poucas épocas, os erros tendem à estabilização, o que demonstra que as redes foram capazes de aprender padrões relevantes dos dados mesmo diante das perturbações introduzidas.

Em seguida na figura 30, foram adotados os mesmos procedimentos para as redes, mas agora aplicados à base 6, ou seja, da rede 1C-1Ag até a rede 2C-3Ag, incluindo também a adição de 5% de ruído na base de dados.

Figura 30 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com 5% de ruído.



Fonte: O autor.

Para os resultados de validação e treinamento na base 6, considerando todas as redes com 5% de ruído na base de dados, observa-se que os valores obtidos são muito semelhantes aos da base 5. Em outras palavras, o erro aumentou em função da maior adição de ruído, mas não foram verificadas alterações significativas entre as bases 5 e 6.

Para os resultados teóricos 1 referentes às redes normalizadas, agora avaliadas sob a condição de 5% de ruído adicionado aos dados, observa-se que os modelos neurais permaneceram capazes de reproduzir adequadamente as variáveis alvo. Mesmo com o aumento do nível do erro de perturbação nos dados de entrada, não foram verificadas alterações significativas no desempenho das redes. A arquitetura 1C-11Ag manteve o melhor desempenho entre as configurações avaliadas, enquanto a arquitetura 2C-3Ag apresentou o menor desempenho relativo. Os valores apresentados na tabela 20 correspondem aos erros percentuais estimados para cada arquitetura, possibilitando uma comparação direta da eficiência preditiva dos modelos, mesmo sob condições mais desafiadoras de ruído.

Verifica-se que os erros percentuais obtidos pelas arquiteturas 1C-11Ag e 2C-3Ag, mesmo na presença de 5% de ruído nos dados de entrada, apresentaram apenas um leve aumento,

Tabela 20 – Erros percentuais dos resultados teóricos 1 normalizados com 5% de ruído das redes 1C-11Ag e 2C-3Ag.

Gaussiana	Centro		Largura		área	
	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag	1C-11Ag	2C-3Ag
Gaussiana 1	0,038	0,160	0,752	0,599	0,299	0,381
Gaussiana 2	0,008	0,212	0,0007	0,296	0,150	0,075
Gaussiana 3	0,086	0,092	0,555	0,754	0,570	0,598
Gaussiana 4	0,057	0,109	0,399	0,537	0,820	1,250
Gaussiana 5	0,062	0,102	0,396	0,400	0,887	1,495

Fonte: O autor.

atingindo valores próximos de 2%, como nas últimas redes. Esse comportamento demonstra que ambas as redes mantiveram elevada capacidade de generalização no cenário analisado, reproduzindo com precisão satisfatória os valores teóricos de referência, os quais, convém salientar, não correspondem a dados experimentais.

Da mesma forma, na análise das predições das redes, a inclusão de 5% de ruído não resultou em melhorias expressivas em relação aos casos com 1% e 2% de ruído; ao contrário, os gráficos de treinamento e validação passaram a exibir uma diferença mais acentuada ao longo das épocas, indicando piora na capacidade de generalização e tornando esse nível de ruído inadequado para o ajuste dos espectros experimentais.

Essa diferença crescente entre as curvas de treinamento e validação indica que, com 5% de ruído, a rede passa a ajustar excessivamente as flutuações aleatórias presentes na base de treinamento, sem conseguir reproduzir com a mesma qualidade os dados do conjunto de validação. Em termos práticos, isso significa que o modelo começa a sofrer *overfitting* em função do ruído em vez de aprender o padrão espectral subjacente, elevando a variância do ajuste e comprometendo a capacidade de generalização; por isso, esse nível de ruído não é adequado para o refinamento das predições experimentais.

4.7 Aplicação de redes neurais com 1% em espectros normalizados com linha de base

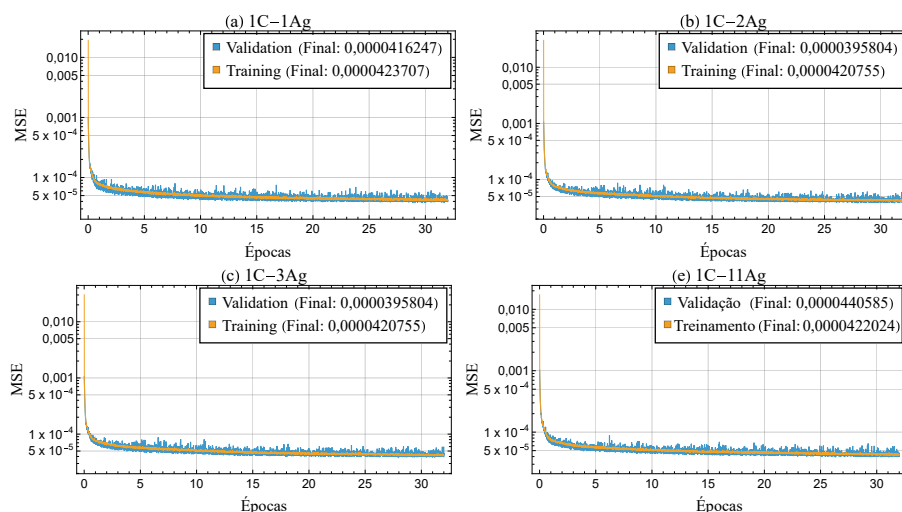
Dessa forma, para avaliar ainda mais o desempenho da rede neural, foi utilizada uma linha de base construída a partir da geração da base de dados. As arquiteturas empregadas com a linha de base possuem as mesmas configurações das redes previamente descritas, ou seja, todos os dados foram normalizados no intervalo de 0 a 1. Além disso, foi inserido um ruído aleatório de 1%, de modo a aproximar os resultados do comportamento experimental real, que foi o melhor resultado para representar os espectros experimentais.

A base de dados foi gerada por meio da soma de funções gaussianas com diferentes centros, larguras e áreas, às quais também foi acrescido o ruído de 1%, somando-se a um

pequeno valor de y_0 aleatório.

Os resultados das melhores configurações são apresentados na Figura 31 obtidos para as redes 1C-1Ag, 1C-2Ag, 1C-3Ag e 1C-11Ag. As redes que não estão incluídas nesta análise foram excluídas por não conseguirem realizar a predição da espectroscopia Raman.

Figura 31 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 5 com baseline na base de dados.



Fonte: O autor.

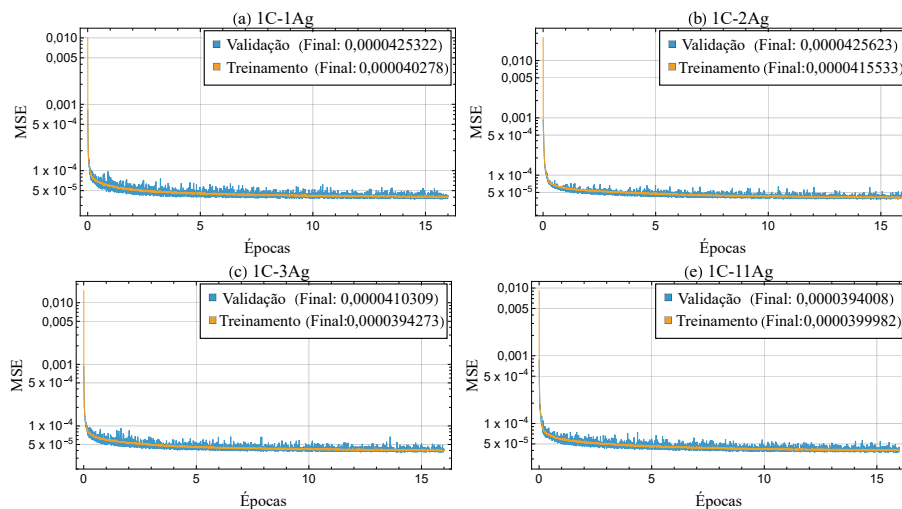
Pode-se observar a evolução do erro quadrático médio (MSE) para diferentes arquiteturas de redes neurais. Após a adição da baseline aos dados de entrada, verifica-se que a rede apresentou consistência nos valores finais de erro em todas as arquiteturas testadas. Esse comportamento indica a capacidade de generalização dos modelos, resultando em um desempenho bastante satisfatório na predição. Ainda assim, nota-se que, após poucas épocas, os erros tendem à estabilização, o que demonstra que as redes conseguiram aprender os padrões dos dados, mesmo diante das perturbações introduzidas.

Na Figura 32, são apresentados os resultados das mesmas redes na Base 6. Com a adição da baseline aos dados, observa-se uma ligeira melhora tanto na validação quanto no treinamento.

Nos resultados de validação e treinamento, observa-se que não há indícios de *overfitting*. As redes estão efetivamente aprendendo, uma vez que os valores apresentam tendência de queda. Esse comportamento indica que o processo de aprendizado está ocorrendo de forma adequada e que a taxa de aprendizagem está contribuindo positivamente para a convergência do modelo.

Os resultados teóricos 1 referentes às redes normalizadas, avaliadas tanto na condição baseline quanto com a adição de 1% de ruído aos dados, demonstram que os modelos neurais permaneceram capazes de reproduzir adequadamente as variáveis alvo. A arquitetura 1C-11Ag manteve-se como a de melhor desempenho. Os valores apresentados na Tabela 21 correspondem aos erros percentuais estimados para cada arquitetura, possibilitando uma comparação direta da eficiência preditiva dos modelos, mesmo sob condições mais desafiadoras de ruído.

Figura 32 – Resultados de validação e treinamento das redes neurais 1C-1Ag e 2C-3Ag na base 6 com baseline na base de dados.



Fonte: O autor.

Tabela 21 – Erros percentuais dos resultados teóricos 1 normalizados com baseline e 1% de ruído das redes 1C-11Ag e 2C-3Ag.

Gaussiana	Centro	Largura	Área
Gaussiana 1	0,062	0,157	0,493
Gaussiana 2	0,072	0,358	0,354
Gaussiana 3	0,025	0,693	0,365
Gaussiana 4	0,065	0,367	0,842
Gaussiana 5	0,029	0,332	1,315

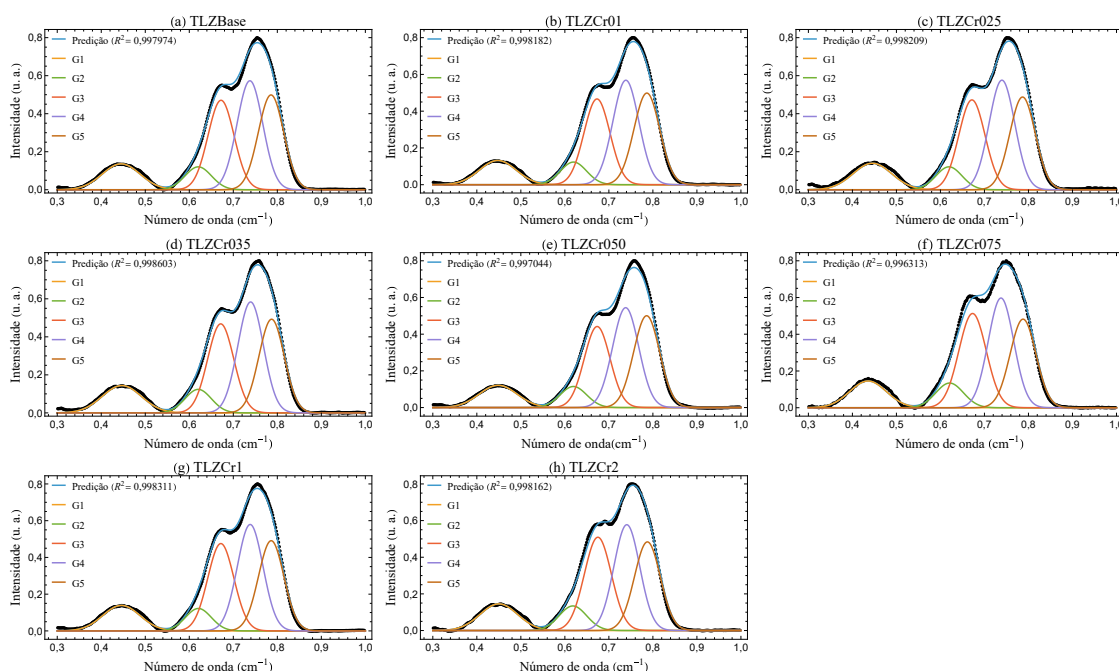
Fonte: O autor.

Os erros percentuais obtidos pelas arquiteturas 1C-11Ag e 2C-3Ag, tanto na condição baseline quanto com a adição de 1% de ruído aos dados de entrada, apresentaram apenas um leve aumento, mantendo-se próximos de 1%, comportamento semelhante ao observado nas análises anteriores. Esse resultado evidencia que ambas as redes preservaram elevada capacidade de generalização no cenário avaliado, reproduzindo com precisão satisfatória os valores teóricos de referência os quais, convém salientar, não correspondem a dados experimentais.

4.7.1 1C-1Ag base 5 – 1C-1Ag base 6 com baseline %

Iniciando a análise dos resultados das redes com baseline, a primeira arquitetura que apresentou desempenho promissor, sem indícios de *overfitting*, foi a rede 1C-1Ag. Essa rede obteve valores de validação de 0,0000416247 e treinamento de 0,0000423707, resultados bastante próximos aos das redes anteriores e igualmente livres de *overfitting*. Mesmo com a adição de 1% de ruído e a presença da baseline, a rede foi capaz de realizar a predição de forma satisfatória. Assim os resultados da predição estão na Figura 33.

Figura 33 – Resultados experimentais da rede 1C-1Ag com baseline na base de dados na base 5.



Fonte: O autor.

Foram selecionados os resultados correspondentes à amostra com a melhor predição TLZCr035 e à amostra com a menor predição TLZCr075, apresentados na Tabela 22.

Tabela 22 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-035 e TLZCr-050 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		área	
	Cr-035	Cr-075	Cr-035	Cr-075	Cr-035	Cr-075
Gaussiana 1	0,605	2,118	5,647	1,548	6,554	4,574
Gaussiana 2	0,855	1,116	12,302	16,041	8,139	10,115
Gaussiana 3	0,516	0,729	15,122	12,049	11,496	2,453
Gaussiana 4	1,008	0,756	3,185	10,606	10,256	19,282
Gaussiana 5	0,638	0,831	1,387	2,604	17,296	16,633

Fonte: O autor.

Observa-se que, para todas as amostras, a rede foi capaz de realizar a predição. A amostra com resultado mais promissor, TLZCr035, apresentou erro máximo de 15,122% na largura da Gaussiana 3 e de 17,296% na área da Gaussiana 5. Já a amostra com menor desempenho preditivo, TLZCr075, apresentou erro de 16,041% na largura da Gaussiana 2 e de 19,282% na área da Gaussiana 5. Mesmo com desempenho inferior, esses resultados ainda são considerados bastante promissores, indicando que a rede 1C-1Ag, aplicada à Base 5, foi capaz de realizar predições consistentes para todas as amostras analisadas.

Para a rede 1C-1Ag na Base 6, os valores de validação (0,0000425322) e de treinamento (0,000040278) apresentaram comportamento semelhante aos obtidos na Base 5, indicando resultados promissores e ausência de *overfitting*.

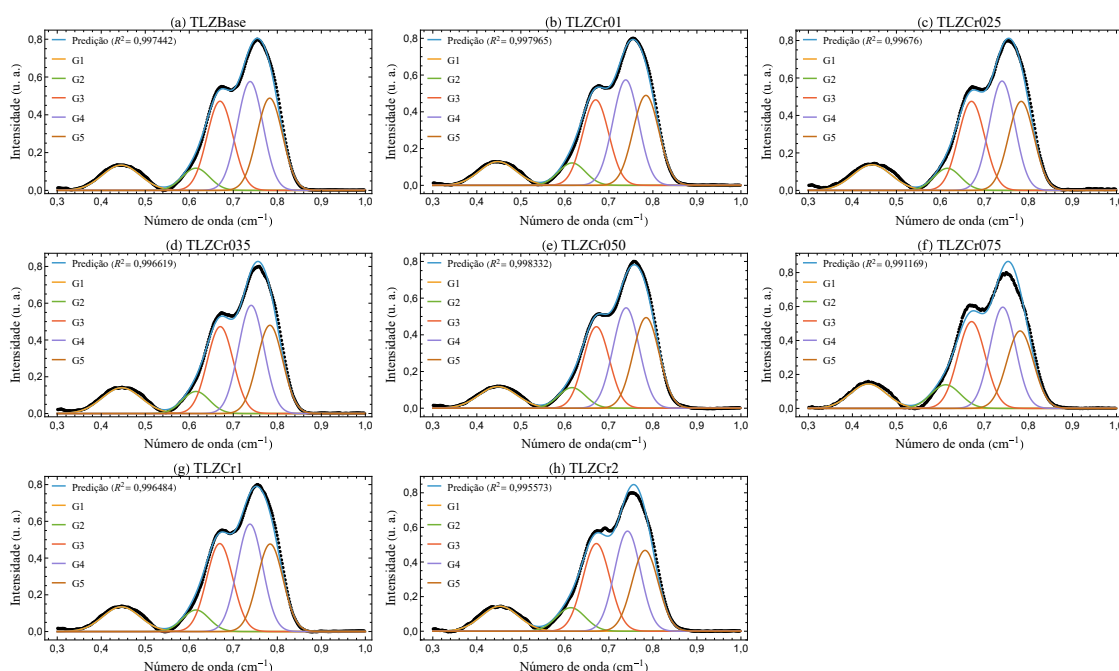
Na predição das amostras, o melhor desempenho foi obtido pela amostra TLZCr01, com erro de 16,193% na largura da Gaussiana 4 e 19,308% na área da Gaussiana 5. Já a amostra com menor desempenho preditivo, TLZCr2, apresentou erros de 14,483% na largura da Gaussiana 3 e 21,583% na área da Gaussiana 5. De modo geral, as predições dessa rede apresentaram erros bastante próximos entre as amostras, evidenciando consistência no desempenho do modelo.

4.7.2 1C-2Ag base 5 – 1C-2Ag base 6 com baseline

Para a rede 1C-2Ag na Base 5, os valores de validação 0,0000424255 e de treinamento 0,0000464611 não apresentaram piora em relação às redes anteriores, mesmo com a inclusão da baseline e a adição de ruído. Esses resultados indicam que os valores obtidos são promissores, evidenciando a ausência de *overfitting* na base de dados.

Dessa forma, a rede 1C-2Ag demonstrou boa capacidade de predição da espectroscopia Raman para a maioria das amostras de vidros teluretos dopados com Cr apresentados na Figura 34.

Figura 34 – Resultados experimentais da rede 1C-2Ag com baseline na base de dados na base 5.



Fonte: O autor.

Separaram-se os resultados das amostras com melhor e pior desempenho. A amostra TLZCr050 apresentou o melhor resultado, enquanto a amostra TLZCr075 obteve o menor desempenho. Entretanto, mesmo no caso da amostra TLZCr050, os resultados podem ser considerados bastante satisfatórios.

Conforme apresentado na Tabela 23, a amostra TLZCr050 apresentou um erro máximo

Tabela 23 – Parâmetros das gaussianas ajustadas para as redes 1C-2Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		área	
	Cr-050	Cr-075	Cr-050	Cr-075	Cr-050	Cr-075
Gaussiana 1	0,454	1,931	0,174	0,611	12,573	7,563
Gaussiana 2	0,215	0,668	14,322	23,534	3,655	21,416
Gaussiana 3	0,541	0,407	14,807	12,268	8,271	2,288
Gaussiana 4	0,981	1,313	0,491	11,054	0,315	19,558
Gaussiana 5	0,434	0,604	0,391	0,122	22,39	18,958

Fonte: O autor.

de 14,807% na largura da Gaussiana 3 e 22,39% na área da Gaussiana 5. Em contrapartida, a amostra TLZCr075 registrou erro de 23,534% na largura da Gaussiana 2 e 19,558% na área da Gaussiana 4.

Na Base 6, observou-se uma melhoria nos resultados, mesmo com a presença da baseline. Os valores de validação (0,0000425322) e treinamento (0,000040278) apresentaram desempenho superior ao obtido na Base 5, indicando que o modelo manteve estabilidade durante o aprendizado e não apresentou sinais de *overfitting*.

A amostra de melhor desempenho foi a TLZCr1, com erro de 14,562% na largura da Gaussiana 2 e 17,018% na área da Gaussiana 5. Já a amostra TLZCr075 obteve o menor desempenho, com erro de 16,074% na largura da Gaussiana 2 e 18,979% na área da Gaussiana 5. Esses resultados reforçam a boa capacidade preditiva da rede e a consistência dos valores obtidos entre as diferentes amostras.

4.7.3 1C-3Ag base 5 – 1C-3Ag base 6 com baseline %

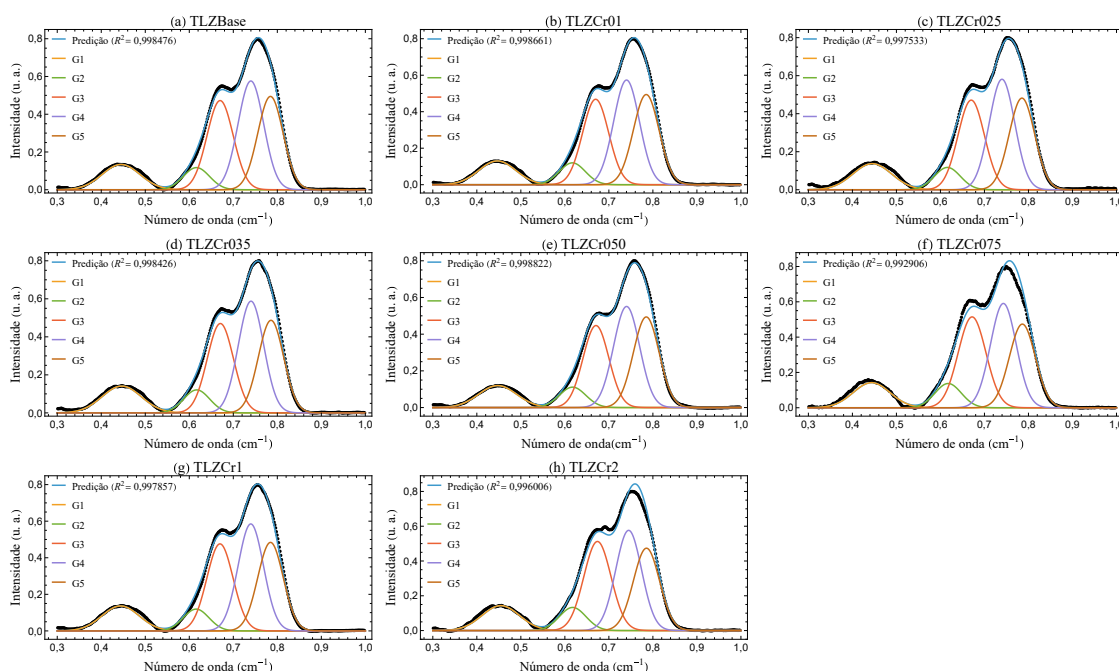
Essa rede também apresentou resultados bastante promissores, com valores de validação de 0,0000395804 e de treinamento de 0,0000420755, semelhantes aos observados nas redes anteriores, indicando a ausência de *overfitting*. Na Figura 35, são apresentados os resultados das predições para todas as amostras analisadas.

Foram destacados os resultados das amostras com desempenho extremo. A amostra TLZCr050 apresentou o desempenho mais elevado, enquanto a amostra TLZCr075 registrou o menor desempenho. Ainda assim, os resultados obtidos pela amostra TLZCr050 podem ser considerados bastante satisfatórios.

De acordo com a Tabela 24, o maior erro para a amostra TLZCr025 foi de 16,516% na largura da Gaussiana 3 e de 22,876% na área da Gaussiana 5. Já para a amostra TLZCr1, o maior erro foi de 15,899% na largura da Gaussiana 3 e de 17,344% na área da Gaussiana 5.

Na Base 6, os resultados também apresentaram melhora mesmo com a linha de base,

Figura 35 – Resultados experimentais da rede 1C-3Ag com baseline na base de dados na base 5.



Fonte: O autor.

Tabela 24 – Parâmetros das gaussianas ajustadas para as redes 1C-2Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		área	
	Cr-025	Cr1	Cr-025	Cr-1	Cr-025	Cr-1
Gaussiana 1	0,184	0,471	8,379	3,146	11,056	7,475
Gaussiana 2	0,126	0,093	10,648	13,095	1,051	7,136
Gaussiana 3	0,294	0,597	16,516	15,899	8,802	11,962
Gaussiana 4	1,073	0,579	2,303	2,491	0,251	7,701
Gaussiana 5	0,405	0,576	1,213	0,140	22,876	17,344

Fonte: O autor.

com valores de validação iguais a 0,00004103039 e de treinamento iguais a 0,0000394273. Esses valores foram ainda melhores do que os obtidos na Base 5, evidenciando que o modelo não sofreu com *overfitting*.

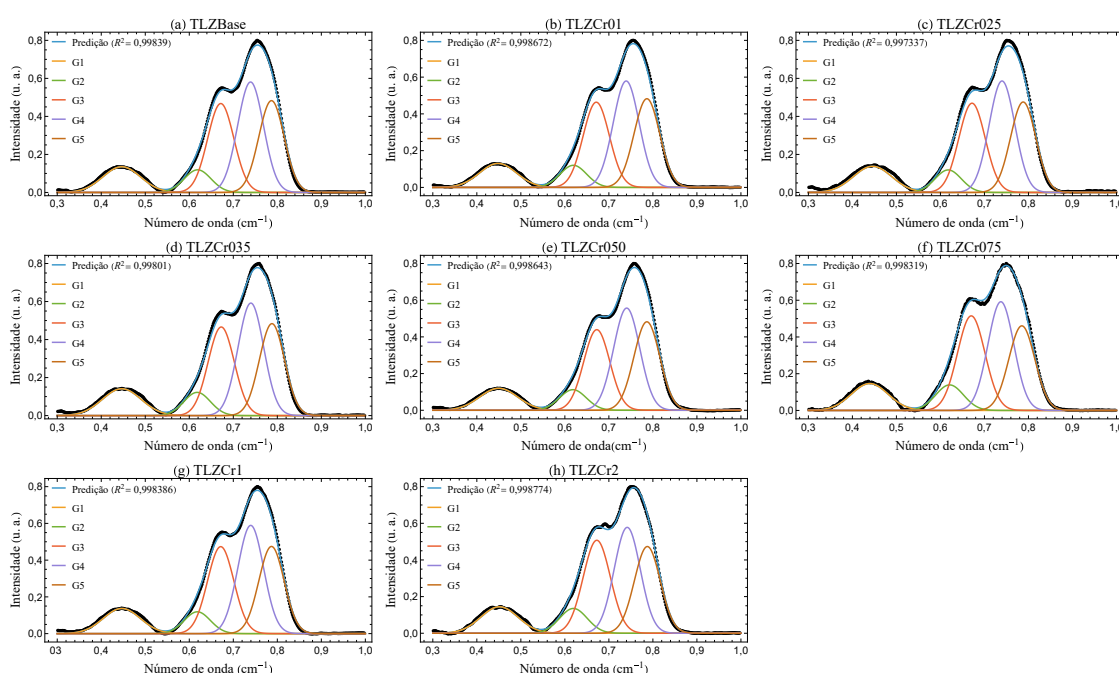
A amostra que apresentou o melhor resultado foi a TLZCr01, cujo maior erro foi de 15,114% na largura da Gaussiana 3 e 19,816% na área da Gaussiana 5. Já a amostra com pior desempenho de predição foi a TLZCr2, que apresentou erro de 15,362% na largura da Gaussiana 3 e 22,631% na área da Gaussiana 3.

4.7.4 1C-11Ag base 5 – 1C-11Ag base 6 com baseline %

A rede 1C-11Ag mostrou-se a mais promissora entre todas as arquiteturas testadas, apresentando resultados semelhantes aos obtidos pelas redes anteriores, com valores de validação (0,0000440585) e de treinamento (0,0000422024). Dessa forma, essa rede foi capaz de realizar a predição de todas as amostras.

Na Figura 36, apresentam-se os resultados das predições para todas as amostras de vidros teluretos dopados com Cr.

Figura 36 – Resultados experimentais da rede 1C-11Ag com baseline na base de dados na base 5.



Fonte: O autor.

Foram destacados os resultados das amostras com desempenho extremo. A amostra TLZCr025 apresentou o desempenho mais elevado, enquanto a amostra TLZCr2 registrou o menor desempenho. Ainda assim, os resultados obtidos pela amostra TLZCr2 podem ser considerados bastante satisfatórios. Mas todas as amostras tiveram uma boa predição; nenhuma delas ficou abaixo do esperado.

Conforme indicado na Tabela 25, o maior erro para a amostra TLZCr025 foi de 16,684% na largura da Gaussiana 3 e 20,802% na área da Gaussiana 3. Para a amostra TLZCr2, o maior erro registrado foi de 14,913% na largura da Gaussiana 3 e 20,802% na área da Gaussiana 5.

Para a mesma rede aplicada à Base 6, observou-se um comportamento semelhante ao da Base 5, mantendo-se entre as arquiteturas mais promissoras, com valores de validação (0,0000394008) e de treinamento (0,0000399982). Esses resultados foram superiores aos obtidos na Base 5, indicando que o modelo manteve estabilidade e não apresentou sinais de *overfitting*.

Tabela 25 – Parâmetros das gaussianas ajustadas para as redes 1C-11Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		área	
	Cr-025	Cr-2	Cr-025	Cr-2	Cr-025	Cr-2
Gaussiana 1	0,448	0,295	9,110	3,639	9,722	2,491
Gaussiana 2	0,426	0,706	12,328	11,231	7,361	12,376
Gaussiana 3	0,565	0,685	16,387	14,913	16,684	20,802
Gaussiana 4	1,083	1,321	3,381	0,843	15,339	7,523
Gaussiana 5	0,765	0,689	1,928	1,710	16,770	13,974

Fonte: O autor.

A amostra com melhor desempenho foi a TLZCr01, com erro máximo de 15,586% na largura da Gaussiana 3 e 19,504% na área da Gaussiana 5. Já a amostra com pior desempenho foi a TLZCr075, que apresentou erro de 15,705% na largura da Gaussiana 2 e 18,490% na área da Gaussiana 5.

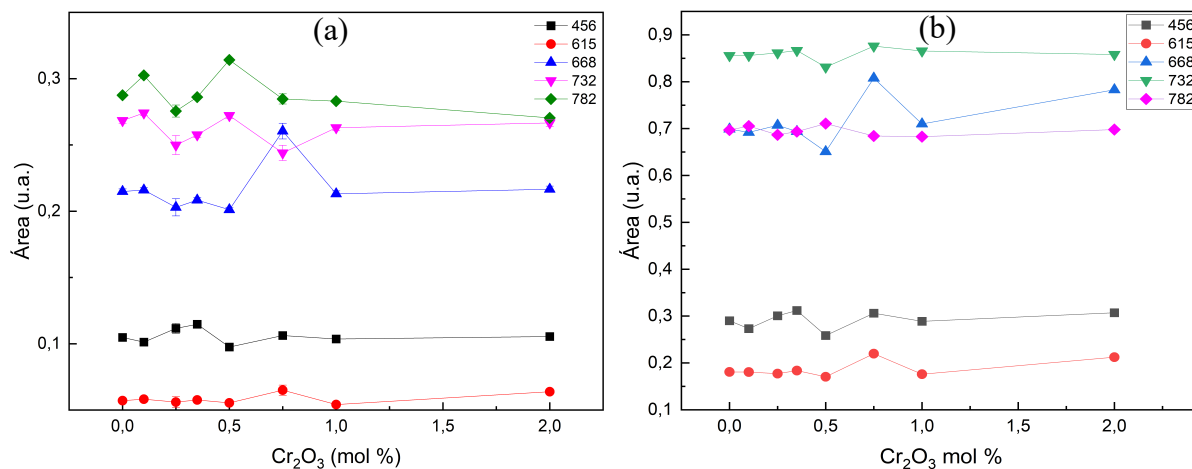
4.8 Comportamento das áreas

Uma das análises realizadas consiste na identificação do comportamento das áreas das bandas Raman em vidros à base de telureto dopados com Cr_2O_3 . Para isso, procede-se a uma comparação entre esses resultados e o desempenho da rede neural 1C-11Ag treinada com 1% de ruído (com linha de base), a qual possui 150 neurônios em sua arquitetura. Tal comparação fundamenta-se nos valores obtidos por meio do *software Origin*, permitindo avaliar a coerência e a robustez dos modelos adotados.

A primeira análise discutida é a obtida no *Origin*, na Figura 37(a), onde as áreas sob as curvas Gaussianas mostram como as bandas Raman variam em função da concentração de Cr_2O_3 . A banda em 456 cm^{-1} , associada às ligações Te-O-Te e O-Te-O, permanece praticamente inalterada com a adição de Cr_2O_3 . De modo semelhante, a banda em 615 cm^{-1} , relacionada a unidades TeO_4 sem oxigênios não ligantes, também não sofre mudanças significativas até 2 mol% de Cr_2O_3 . A banda em 668 cm^{-1} , atribuída a unidades TeO_4 contendo oxigênios não ligantes (NBO), inicialmente aumenta com 0,75 mol% de Cr_2O_3 , mas diminui em concentrações maiores, estabilizando-se posteriormente. Já a banda em 732 cm^{-1} , associada a unidades TeO_3 , apresenta inicialmente uma queda, seguida de aumento até 0,50 mol% e posterior estabilização. Por fim, a banda em 782 cm^{-1} , referente a unidades TeO_{3+1} com NBO, diminui em 0,25 mol%, aumenta em 0,50 mol% e então se mantém estável.

De modo geral, os resultados sugerem que o Cr_2O_3 atua inicialmente como dopante, mas a partir de aproximadamente 0,50 mol% passa a atuar como modificador de rede, alterando a estrutura vítrea.

Figura 37 – Áreas das bandas obtidas pela: (a) Deconvolução gaussiana no *Origin* e (b) Predição usando a rede neural 1C-11Ag.



Fonte: O autor.

É importante destacar que esse procedimento pode introduzir incertezas nos resultados, uma vez que o ajuste realizado no *Origin* exige a fixação das posições centrais das bandas em 456, 615, 668, 732 e 782 cm^{-1} . Caso ocorram deslocamentos reais desses picos em função da concentração de Cr_2O_3 , o ajuste forçado pode levar a distorções nas áreas calculadas, comprometendo a interpretação das variações estruturais observadas. Dessa forma, parte do comportamento obtido pode estar associada às limitações inerentes ao método de ajuste adotado.

A Figura 37(b) apresenta o comportamento das áreas obtidas pela rede neural em função da concentração de Cr_2O_3 . Ao comparar esses resultados com aqueles obtidos pelo ajuste no *Origin*, apresentados na Figura 37(a), observa-se que, no caso do *Origin*, as áreas exibem variações mais acentuadas e menor estabilidade entre as diferentes concentrações de Cr_2O_3 . Em contrapartida, os resultados fornecidos pela rede neural apresentam um comportamento mais suave e consistente, sugerindo maior robustez frente às limitações impostas pelo procedimento de ajuste tradicional.

Esses resultados indicam que a rede neural foi capaz de prever com precisão o comportamento das áreas associadas à influência do dopante na rede vítrea, produzindo um comportamento coerente esperado. Tal desempenho reforça a eficiência do modelo de rede neural em comparação ao método tradicional.

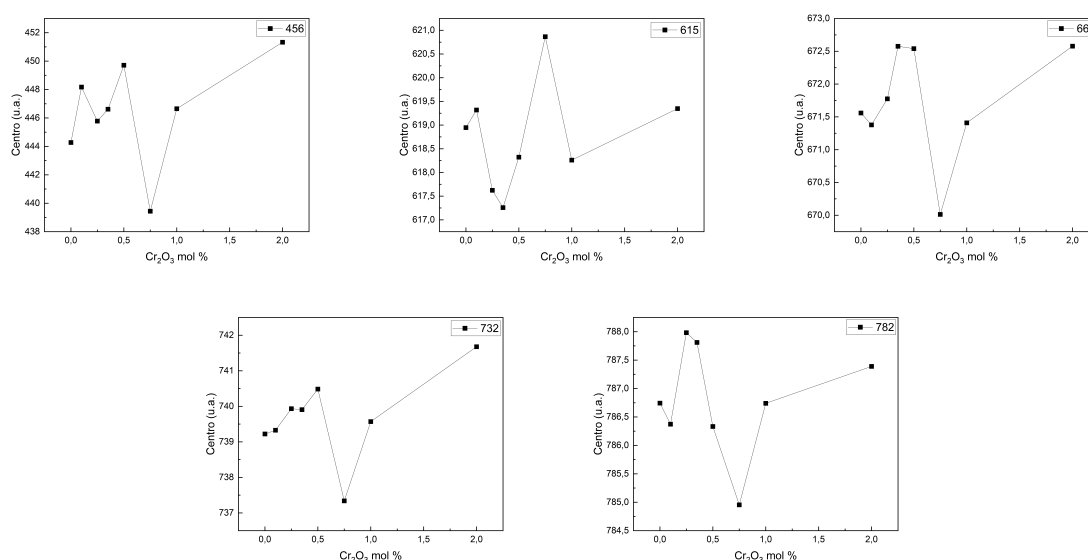
4.8.1 Comportamento dos centros pela 1C-11Ag

Neste momento, realizou-se a análise do comportamento dos centros das bandas Raman utilizando redes neurais. Diferentemente do procedimento tradicional de deconvolução, em

que os centros dos picos são previamente fixados, a abordagem baseada em redes neurais dispensa essa etapa, pois o próprio modelo é capaz de prever tanto as posições quanto as áreas correspondentes. Dessa forma, as redes neurais fornecem uma estimativa mais flexível e potencialmente mais precisa dos parâmetros espectrais, permitindo identificar variações sutis nos perfis das bandas que poderiam ser negligenciadas por métodos convencionais.

Conforme apresentado na Figura 38, verifica-se que os valores obtidos para cada banda mantiveram-se próximos aos reportados na literatura (456, 615, 668, 732 e 782 cm^{-1}). A amostra que apresenta maior variação na posição do centro é a que contém 0,75% de Cr, apresentando um deslocamento real, ou seja, maior que 3 cm^{-1} , que é a precisão do equipamento.

Figura 38 – Centros das bandas obtidas pela predição usando a rede neural 1C-11Ag na base de dados normalizada com 1% de ruído e linha de base.



Fonte: O autor.

4.8.2 Comportamento das áreas para todas as redes

Uma vez que todas as redes foram treinadas com a mesma base de dados e para gerar um perfil de confiabilidade dos resultados, assim como nos métodos tradicionais obtidos pelos *softwares*, como o *Origin*, foram calculadas as médias dos comportamentos das áreas para todas as redes neurais e os respectivos desvios padrão (erro) referentes às bandas Raman de vidros à base de telureto dopados com Cr₂O₃. Para isso, utilizaram-se os resultados das cinco melhores redes: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag, treinadas com 1% de ruído (com correção de linha de base). Os resultados do comportamento das áreas estão apresentados na Tabela 26. Nota-se que, mesmo tendo sido apresentadas diferenças entre as redes, o desvio padrão apresentado é relativamente pequeno, da ordem de 3 casas, com exceção de alguns

valores que obtiveram próximo à unidade na segunda casa.

Tabela 26 – Valores médios e respectivos erros para as áreas das gaussianas das amostras analisadas obtidos pelas cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.

Amostra	456	615	668	732	782
0	$0,230 \pm 0,002$	$0,181 \pm 0,006$	$0,701 \pm 0,004$	$0,845 \pm 0,009$	$0,71 \pm 0,01$
0,1	$0,273 \pm 0,002$	$0,183 \pm 0,005$	$0,692 \pm 0,005$	$0,846 \pm 0,006$	$0,66 \pm 0,01$
0,25	$0,303 \pm 0,002$	$0,180 \pm 0,004$	$0,709 \pm 0,004$	$0,85 \pm 0,01$	$0,69 \pm 0,01$
0,35	$0,316 \pm 0,003$	$0,186 \pm 0,006$	$0,698 \pm 0,006$	$0,86 \pm 0,01$	$0,70 \pm 0,01$
0,5	$0,259 \pm 0,002$	$0,171 \pm 0,003$	$0,652 \pm 0,005$	$0,820 \pm 0,008$	$0,73 \pm 0,01$
0,75	$0,302 \pm 0,005$	$0,22 \pm 0,02$	$0,801 \pm 0,008$	$0,880 \pm 0,009$	$0,69 \pm 0,01$
1	$0,290 \pm 0,003$	$0,180 \pm 0,005$	$0,710 \pm 0,004$	$0,86 \pm 0,01$	$0,69 \pm 0,01$
2	$0,305 \pm 0,008$	$0,21 \pm 0,02$	$0,77 \pm 0,04$	$0,85 \pm 0,01$	$0,70 \pm 0,01$

Fonte: O autor.

Para melhor visualização, o comportamento da média das áreas e seus respectivos erros são apresentados na Figura 39. As áreas sob as curvas Gaussianas evidenciam a pequena variação das bandas Raman em função da concentração de Cr_2O_3 para cada rede neural, como na analisada pela melhor rede. Dessa forma, observa-se que os resultados das áreas, obtidos a partir da comparação entre todas as redes neurais analisadas, apresentaram valores bastante próximos aos determinados pelo software *Origin*. Além disso, os erros associados às redes neurais mostraram-se baixos, indicando que os modelos foram capazes de realizar a predição das áreas de todas as bandas Raman de maneira consistente quando comparados aos resultados do *Origin*. Observa-se que as barras de erro dos resultados obtidos pela média das cinco melhores redes neurais ficaram bastante próximas e até menores do que aquelas determinadas pelo software *Origin*. Esse resultado demonstra que realizar a média e o desvio de resultados treinados por diferentes redes é um bom método para representar um intervalo de confiabilidade.

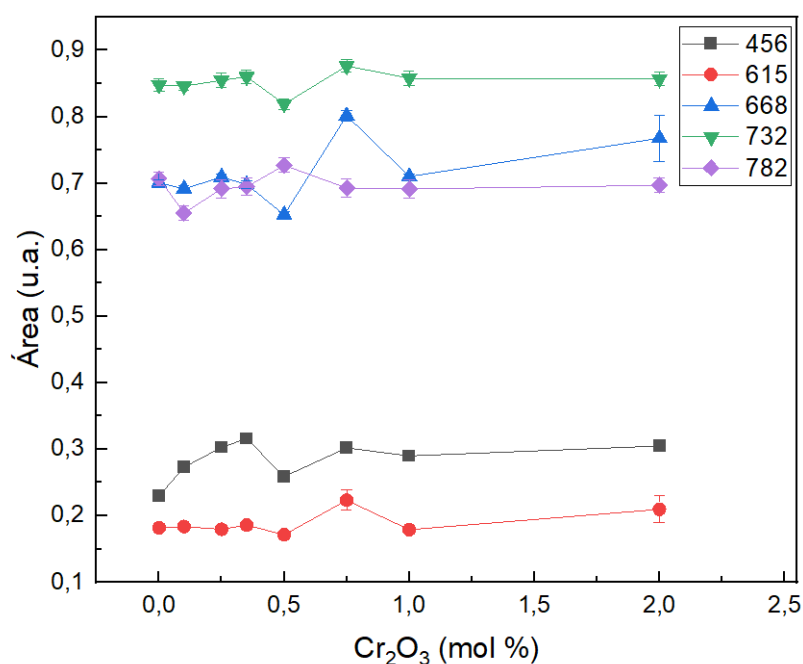
Além disso, ao considerar o conjunto das cinco melhores redes neurais, nota-se uma maior robustez estatística na estimativa das áreas, uma vez que a média reduz a influência de eventuais flutuações oriundas de inicializações aleatórias dos pesos, da ordem de apresentação dos dados de treinamento ou de mínimos locais da função de custo. Dessa forma, o procedimento de combinar múltiplos modelos permite mitigar vieses específicos de uma única rede e, consequentemente, fornecer uma predição mais estável e confiável.

Outro ponto relevante é que a boa concordância entre as áreas previstas pelas redes neurais e aquelas ajustadas pelo *Origin* indica que a abordagem de modelagem baseada em aprendizado de máquina é adequada para capturar a relação entre o espectro Raman e a concentração de Cr_2O_3 . Isso sugere que, uma vez treinadas, as redes podem ser utilizadas como ferramentas auxiliares na análise espectroscópica, reduzindo o tempo de processamento dos dados, bem como a dependência de ajustes manuais dos espectros.

Por fim, os resultados obtidos reforçam o potencial das redes neurais artificiais na

quantificação de parâmetros físico-químicos a partir de dados espectrais, demonstrando que, com um conjunto de treinamento adequadamente preparado e uma arquitetura bem ajustada, é possível alcançar resultados comparáveis, ou até superiores em termos de incerteza, aos métodos tradicionais de ajuste de curvas disponíveis em *softwares* amplamente utilizados na comunidade científica.

Figura 39 – Áreas das bandas obtidas pela predição usando as cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.



Fonte: O autor.

4.8.3 Comportamento das centros para todas as redes

Para verificar se houve deslocamento dos centros, também foram analisados os comportamentos dos centros das redes neurais e os respectivos desvios padrão (erro), utilizando as cinco melhores redes. A Tabela 27 apresenta os resultados dos centros para as cinco redes neurais, e a Figura 27 apresenta o gráfico com o comportamento da média e do desvio padrão.

Podemos ver que os resultados ficaram dentro do esperado quando comparamos os valores dos centros (456, 615, 668, 732 e 782 cm^{-1}). Como nos centros a predição possui menores erros relativos, temos um menor desvio padrão quando comparadas as diferentes redes neurais. Na Figura 27, nota-se que as barras de erro são pequenas o suficiente para ficarem sobrepostas ao tamanho do ponto. Além disso, observa-se que, à medida que nos afastamos dos centros em direção às extremidades do espectro, os erros relativos tendem a aumentar, o que se reflete em barras de erro ligeiramente maiores. Ainda assim, a sobreposição entre

as predições das diferentes redes indica boa reprodutibilidade do modelo, sugerindo que a arquitetura escolhida é robusta frente a pequenas variações na inicialização e no processo de treinamento.

Esse comportamento é consistente com o fato de que as regiões centrais concentram uma maior quantidade de informação espectral relevante e, consequentemente, são melhor aprendidas pelas redes neurais. Já nas regiões de borda, onde o sinal costuma ser mais fraco e ruidoso, a incerteza do modelo se torna mais evidente. Dessa forma, os resultados apresentados corroboram a adequação do método proposto para descrever com precisão os modos vibracionais principais, ainda que alguns ajustes adicionais possam ser necessários para aprimorar o desempenho nas regiões menos intensas do espectro.

Tabela 27 – Valores médios e respectivos erros para os centros das gaussianas das amostras analisadas obtidos pelas cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.

Amostra	456	615	668	732	782
0	443,6 ± 0,2	616,2 ± 0,3	669,8 ± 0,1	737,2 ± 0,1	785,4 ± 0,1
0,1	446,6 ± 0,1	616,8 ± 0,2	669,8 ± 0,1	737,8 ± 0,1	785,6 ± 0,1
0,25	444,4 ± 0,4	616,0 ± 0,2	669,6 ± 0,1	738,0 ± 0,1	786,0 ± 0,1
0,35	443,2 ± 0,6	616,2 ± 0,2	669,8 ± 0,2	738,4 ± 0,2	786,2 ± 0,1
0,5	447,4 ± 0,1	616,4 ± 0,2	671,2 ± 0,1	738,6 ± 0,1	785,4 ± 0,1
0,75	436,6 ± 0,4	615,6 ± 0,6	669,8 ± 0,2	738,0 ± 0,3	785,8 ± 0,1
1	444,4 ± 0,1	616,2 ± 0,3	669,0 ± 0,2	737,2 ± 0,2	785,6 ± 0,1
2	449,8 ± 0,3	616,2 ± 0,4	671,4 ± 0,1	740,2 ± 0,2	786,0 ± 0,1

Fonte: O autor.

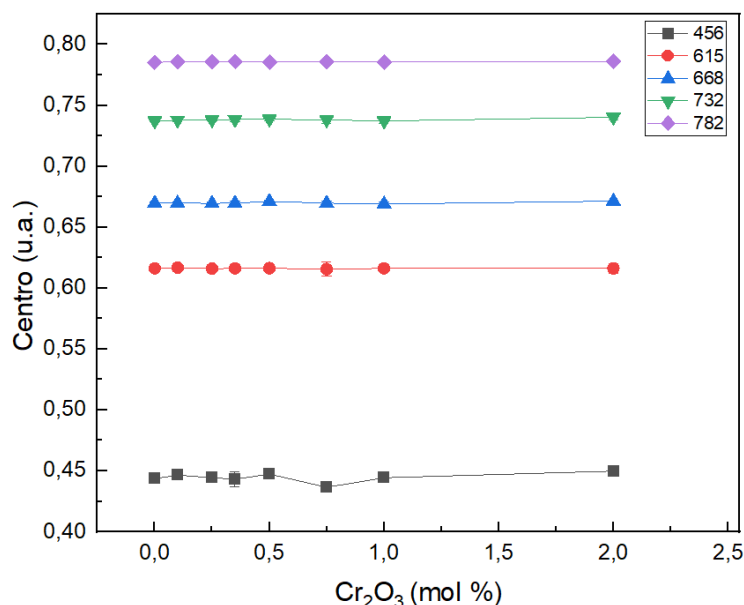
Assim, pode-se observar que os resultados dos centros apresentaram um comportamento totalmente linear, sem grandes variações entre eles. Isso indica que os erros associados aos resultados foram muito baixos, confirmando que os resultados obtidos para todas as redes neurais são mais confiáveis do que os apresentados pelo *Origin* e que as 5 melhores redes selecionadas conseguiram fazer a predição dos vidros teluretos dopados com Cr_2O_3 .

Como nas medidas experimentais a precisão do equipamento é de 3 cm^{-1} , pode-se dizer que não houve deslocamento dos centros, mesmo com o desvio padrão das redes sendo menor que 1 cm^{-1} . Esses resultados eram esperados em vidros com dopantes, uma vez que se espera que o dopante não influencie diretamente a estrutura.

4.8.4 Comportamento das larguras para todas as redes

Por fim, foram analisados os comportamentos das larguras das bandas, utilizando-se também as cinco melhores redes neurais treinadas com 1% de ruído, com correção de linha de base, por meio dos valores médios das larguras obtidas pelas redes neurais, bem como os

Figura 40 – Centros das bandas obtidas pela predição usando as cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.



Fonte: O autor.

respectivos desvios padrão (erro) para cada banda e amostra, que são apresentados na Tabela 28 na Figura 41.

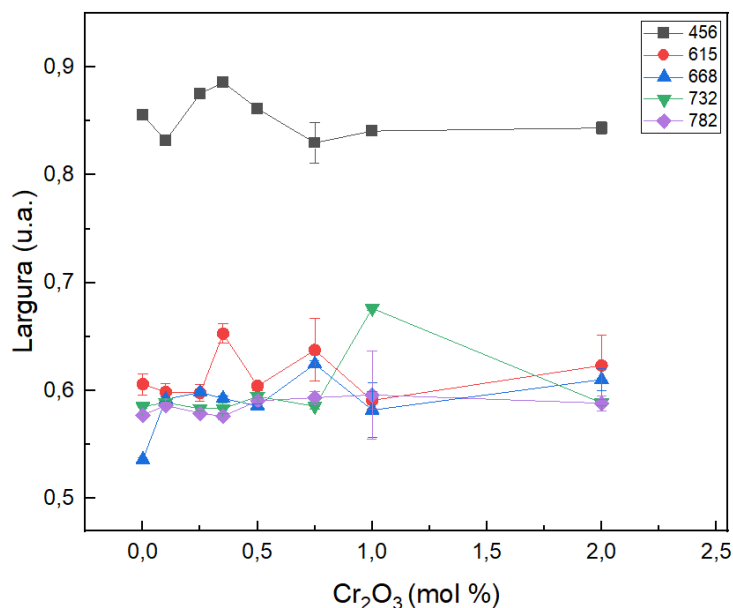
Tabela 28 – Valores médios e respectivos erros para as larguras das gaussianas das amostras analisadas obtidos pelas cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.

Amostra	456	615	668	732	782
0	85,5 ± 0,4	60,6 ± 0,1	53,6 ± 0,2	58,5 ± 0,1	57,7 ± 0,1
0,1	83,2 ± 0,5	59,9 ± 0,8	59,2 ± 0,2	58,9 ± 0,2	58,6 ± 0,1
0,25	87,5 ± 0,3	59,7 ± 0,8	59,8 ± 0,2	58,3 ± 0,2	57,9 ± 0,1
0,35	88,5 ± 0,3	65,2 ± 0,9	59,2 ± 0,2	58,3 ± 0,5	57,6 ± 0,3
0,5	86,1 ± 0,1	60,4 ± 0,4	58,6 ± 0,2	59,4 ± 0,1	59,06 ± 0,2
0,75	82,9 ± 0,2	63,7 ± 0,3	62,4 ± 0,3	58,5 ± 0,3	59,3 ± 0,6
1	84,0 ± 0,4	59,1 ± 0,8	58,2 ± 0,2	67,6 ± 0,2	59,6 ± 0,4
2	84,3 ± 0,6	62,3 ± 0,3	61,0 ± 0,1	58,8 ± 0,3	58,8 ± 0,7

Fonte: O autor.

Podemos observar que os resultados das larguras apresentaram os maiores erros em algumas amostras quando comparados com a área e os centros. Isso se deve às redes 1C-3Ag e 1C-4Ag, que, entre as cinco analisadas, foram as que apresentaram os maiores erros nas larguras para as amostras de 0,75% e 1%. Ou seja, essas redes tiveram erros de largura maiores quando comparadas às outras três redes, como, por exemplo, a rede 1C-11Ag. Ainda assim, os erros associados à largura permanecem muito baixos e dentro do esperado, menores que 3 cm⁻¹. Dessa forma, podemos afirmar que, mesmo com esses erros um pouco maiores, as redes neurais continuam sendo confiáveis quando comparadas ao *Origin*.

Figura 41 – Largura das bandas obtidas pela predição usando as cinco melhores redes neurais: 1C-1Ag, 1C-2Ag, 1C-3Ag, 1C-4Ag e 1C-11Ag.



Fonte: O autor.

Esses erros podem ser provenientes do pré-processamento dos dados pela remoção manual da linha de base. Nesse procedimento, pequenas variações na escolha dos pontos de ancoragem ou na forma como a linha é traçada podem introduzir discrepâncias sutis no formato dos picos, afetando de maneira mais sensível os parâmetros relacionados à largura do que aqueles associados à área ou à posição do centro.

Além disso, como a largura é diretamente relacionada ao ajuste da forma do pico, qualquer ruído residual ou tendência não completamente corrigida no espectro pode impactar o valor ajustado pelas redes neurais. Em contrapartida, a área e o centro tendem a ser mais robustos frente a esse tipo de variação, o que explica a melhor concordância observada nesses parâmetros.

Portanto, os resultados sugerem que uma etapa de pré-processamento mais automatizada e padronizada, por exemplo, com métodos de correção de linha de base sistemáticos por redes neurais, também pode contribuir para a redução ainda maior dos erros em largura. Mesmo assim, os valores já obtidos indicam que o desempenho das redes neurais é adequado para fins analíticos, reforçando seu potencial como alternativa confiável aos métodos tradicionais de ajuste, como o *Origin*.

5 CONCLUSÃO

Neste trabalho, foram aplicadas diferentes arquiteturas de redes neurais artificiais do tipo MLP para a predição de parâmetros gaussianos do espectro Raman relativos aos modos vibracionais das estruturas de TeO_2 de vidros TLZ dopados com Cr_2O_3 . As redes MLP foram treinadas com diferentes bases de dados, incluindo a chamada sem normalização, representando o espectro após o primeiro pré-processamento de remoção da linha de base; com normalização de todos os parâmetros de entrada e saída, adição de diferentes níveis de ruído (1%, 2% e 5%); e, por fim, 1% com linha de base constante.

As primeiras redes neurais MLP desenvolvidas de 2C-1Ag a 2C-5Ag apresentaram desempenho satisfatório na predição dos espectros teóricos, com baixos erros de validação, demonstrando boa capacidade de ajuste aos dados simulados. No entanto, os erros de treinamento mais elevados e a redução do coeficiente de determinação ($R^2=0,96$) ao aplicar os modelos em espectros experimentais indicam limitações na generalização das redes. Os maiores desvios foram observados nos parâmetros de largura e área, refletindo a dificuldade das redes em lidar com ruídos, variações de baseline e assimetrias presentes nos dados reais, ausentes no conjunto teórico de treinamento. Entre as arquiteturas testadas, a rede 2C-4Ag apresentou o melhor desempenho relativo, embora ainda seja insuficiente para predições confiáveis.

A normalização dos dados foi essencial para garantir o bom desempenho das redes neurais, um treinamento mais estável e a obtenção de resultados consistentes entre os conjuntos de treinamento e validação. Dentre as arquiteturas testadas, a rede 1C-11Ag destacou-se por apresentar os melhores desempenhos, com baixos valores de erro percentual de MSE e excelente capacidade de generalização, evidenciada pela ausência de *overfitting* e pela proximidade dos resultados de treinamento e validação. De modo geral, as redes foram capazes de realizar predições altamente precisas, mantendo os centros e áreas das bandas dentro de margens de erro aceitáveis e alcançando coeficientes de determinação superiores a 0,998, o que confirma os resultados dos modelos propostos.

As demais arquiteturas (1C-3Ag, 1C-4Ag e 2C-3Ag) também apresentaram resultados satisfatórios, embora com pequenas variações de desempenho entre a base 5 com 10 mil curvas de treinamento e a base 6 com 20 mil. Dessa forma, a base de 10 mil curvas mostrou-se adequada entre custo computacional e qualidade de ajuste, não havendo ganho que justifique o maior tempo de treinamento associado à base de 20 mil curvas.

Observou-se que o aumento do número de camadas de 1 para 2 ou de neurônios na camada 1 nem sempre resultou em melhorias significativas, podendo, em alguns casos, favorecer o sobreajuste dos dados, indicando que redes mais profundas não são ideais para o problema em questão. Assim, a arquitetura 1C-11Ag, com uma única camada oculta e um número moderado de neurônios, revelou-se a solução mais equilibrada, atendendo às exigências do problema sem introduzir complexidade.

De modo geral, as redes neurais desenvolvidas mostraram-se capazes de mapear com eficiência as características espectrais dos espectros Raman simulados e experimentais, validando a metodologia proposta. Assim, conclui-se que as arquiteturas testadas, especialmente a rede 1C-11Ag, apresentam grande potencial para aplicação em análises espectroscópicas, oferecendo uma ferramenta confiável e precisa para a identificação e caracterização de materiais vítreos dopados.

Comparando os resultados entre as arquiteturas, a rede 1C-11Ag destacou-se por apresentar os melhores desempenhos preditivos, sendo capaz de reproduzir de forma mais fiel as características das bandas Raman simuladas. As demais arquiteturas também exibiram resultados satisfatórios, com diferenças numéricas pequenas entre os valores de erro, o que reforça a consistência global do método empregado.

Em relação à adição de ruído nas bases de dados normalizadas, observou-se que a inclusão de 1% de ruído aleatório proporcionou o melhor resultado entre a estabilidade do treinamento e a capacidade de generalização. Nessa condição, os erros médios quadráticos permaneceram baixos, os coeficientes de determinação mantiveram-se elevados e não foram identificados indícios relevantes de *overfitting*, ao mesmo tempo em que as predições dos parâmetros gaussianos se aproximaram de forma consistente dos valores obtidos experimentalmente. Já para 2% de ruído, houve um leve aumento nos desvios percentuais e nenhuma melhora significativa na qualidade dos ajustes, enquanto a adição de 5% de ruído levou a uma diferença mais pronunciada entre as curvas de treinamento e validação, indicando perda de generalização e tornando esse nível de ruído inadequado para o refinamento das predições experimentais.

Complementarmente, a introdução de uma linha de base sintética na geração das curvas, combinada à normalização e à adição de 1% de ruído, mostrou-se uma estratégia eficiente para aproximar as condições de treinamento do cenário experimental. Nessa configuração, especialmente para as redes 1C-2Ag e 1C-11Ag, os espectros preditos reproduziram de forma mais coerente o comportamento das bandas Raman (centros, larguras e áreas), além de apresentarem tendências mais estáveis quando comparados aos ajustes realizados pelo *Origin*. Esses resultados indicam que o uso de baseline na base de dados, associado a um nível moderado de ruído, contribui para tornar o modelo de rede neural mais robusto frente às imperfeições dos dados reais e mais adequado para a análise automática de espectros Raman de vidros teluretos.

Em síntese, a análise comparativa entre os métodos demonstra que as redes neurais oferecem uma abordagem mais precisa e confiável para a determinação das áreas e posições das bandas Raman em vidros teluretos dopados com Cr_2O_3 . Enquanto o método tradicional implementado no *Origin* depende da fixação prévia dos centros das bandas, o que pode introduzir distorções e perda de informação, a rede neural é capaz de aprender diretamente dos espectros experimentais, ajustando-se às variações reais do sistema.

Além disso, a comparação dos parâmetros gaussianos preditos pelas cinco melhores arquiteturas, treinadas com 1% de ruído e com correção de linha de base, permitiu avaliar a

consistência dos resultados por meio das médias e dos desvios padrão entre redes. As áreas apresentaram baixa dispersão e barras de erro semelhantes ou menores que as do *Origin*. Os centros permaneceram nos valores esperados, com erros relativos pequenos, indicando alta estabilidade entre arquiteturas e ausência de deslocamentos relevantes, compatível com a precisão experimental e com a interpretação de que o dopante não alterou a estrutura do vidro. Já as larguras mostraram maior variação em algumas amostras, mas a análise estatística entre redes ajudou a delimitar a confiabilidade das estimativas e a identificar onde a variabilidade é mais significativa. Assim, ao utilizar a média e o desvio padrão obtidos a partir de diferentes redes, a metodologia fornece uma estimativa central acompanhada de uma medida objetiva de incerteza, aumentando a confiabilidade da predição e permitindo identificar parâmetros mais sensíveis à variação entre arquiteturas.

Os resultados obtidos indicam que o modelo neural não apenas reproduz de forma coerente os valores reportados na literatura e comparados ao método tradicional, como também apresenta menor sensibilidade a ruídos experimentais e maior estabilidade entre as diferentes concentrações de Cr_2O_3 . Dessa forma, a aplicação de redes neurais se mostra uma ferramenta promissora para o tratamento de dados espectroscópicos complexos, contribuindo significativamente para o avanço das análises estruturais em materiais vítreos.

6 PERSPECTIVAS FUTURAS

- Ajuste dos espectros Raman com remoção da linha de base realizada por rede neural sem necessidade de pré-processamento de dados;
- Utilização de redes neurais do tipo multilayer perceptron para a predição de espectros de infravermelho por transformada de Fourier (FTIR);
- Aplicação de redes neurais na análise de vidros teluretos com maiores variações de áreas, como ocorre com a adição de modificadores de redes;
- Aplicação de outros tipos de redes neurais, como as recorrentes e convolucionais, para a predição de espectros.

REFERÊNCIAS

- 1 HAYKIN, S. *Neural Networks and Learning Machines*. 3. ed. New York: Prentice Hall, 2009.
- 2 BABINI, M.; MARRANGHELLO, N. *Introdução às Redes Neurais Artificiais*. São José do Rio Preto, SP: Cultura Acadêmica Editora; Laboratório Editorial IBILCE, UNESP - Universidade Estadual Paulista, Instituto de Biociências, Letras e Ciências Exatas, 2007.
- 3 LI, L.-N. et al. A review of artificial neural network based chemometrics applied in laser-induced breakdown spectroscopy analysis. *Spectrochimica Acta Part B: Atomic Spectroscopy*, Elsevier, v. 180, p. 106183, 2021. Disponível em: <<https://doi.org/10.1016/j.sab.2021.106183>>. Acesso em: 24 abr. 2026.
- 4 ZHANG, L. et al. Decoding viscosity–microstructure relationships in the ternary $\text{cao-sio}_2\text{-fe}_x\text{o}$ system via integrated machine learning and multimodal characterization. *Journal of Materials Science & Technology*, Elsevier, v. 250, p. 299–308, 2026. Disponível em: <<https://doi.org/10.1016/j.jmst.2025.05.069>>. Acesso em: 24 abr. 2026.
- 5 POURSORKH, Z. et al. Detection of non-protein nitrogen (npr) milk adulterants using raman spectroscopy with machine learning. *Food Control*, Elsevier, Vancouver, BC, Canada, v. 181, p. 111737, 2026. Disponível em: <<https://doi.org/10.1016/j.foodcont.2025.111737>>. Acesso em: 24 abr. 2026.
- 6 ZHANG, H. et al. Multimodal deep learning method based on multiple spectra for lung cancer early diagnosis. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, Elsevier, v. 314, p. 124990, 2026. Disponível em: <<https://doi.org/10.1016/j.saa.2025.124990>>. Acesso em: 24 abr. 2026.
- 7 ZOU, H. H. et al. Spectral preprocessing across featureless nir and feature-rich mir: Machine learning classification of black plastics. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, Elsevier, v. 309, p. 124921, 2025. Disponível em: <<https://doi.org/10.1016/j.saa.2025.124921>>. Acesso em: 24 abr. 2026.
- 8 DUTTA, S. et al. Sustainable diesel engine performance enhancement using pyrolytic plastic oil blends: Experimental investigation and artificial neural network-based prediction. *Fuel*, Elsevier, v. 377, p. 129141, 2025. Disponível em: <<https://doi.org/10.1016/j.fuel.2025.129141>>. Acesso em: 24 abr. 2026.
- 9 BRANDT, J. et al. Deep learning for reconstructing low-quality ftir and raman spectra — a case study in microplastic analyses. *Applied Spectroscopy*, SAGE Publications, v. 75, n. 6, p. 665–674, 2021. Disponível em: <<https://doi.org/10.1177/00037028211009455>>. Acesso em: 24 abr. 2026.
- 10 NOVATSKI, A. et al. Thermal and optical properties of lithium-zinc-tellurite glasses. *Materials Chemistry and Physics*, v. 231, p. 150–158, 2019.
- 11 GONÇALVES, A. et al. Estudo da dependência composicional com as propriedades térmicas e espectroscópicas do sistema $\text{teo}_2\text{-li}_2\text{o-bao}$. UNIVERSIDADE ESTADUAL DE PONTA GROSSA, 2016.

- 12 GOMES, J. L. et al. Correlation between nonbridging oxygens and the thermal and optical properties of the $\text{teo}_2\text{-li}_2\text{o-moo}_3$ glassy system. *Journal of Materials Research*, v. 30, n. 16, p. 2417–2424, 2015.
- 13 FAN, M. et al. Deep learning-based component identification for the raman spectra of mixtures. *Analytica Chimica Acta*, Elsevier, v. 1081, p. 51–59, 2019. Disponível em: <<https://doi.org/10.1016/j.aca.2019.06.034>>. Acesso em: 04 maio 2026.
- 14 JOHNSON, J. R. et al. Towards optimizing neural network-based quantification for nmr metabolomics. *Analytica Chimica Acta*, Elsevier, v. 1309, p. 347054, 2025. Disponível em: <<https://doi.org/10.1016/j.aca.2024.347054>>. Acesso em: 04 maio 2026.
- 15 WANG, Z. et al. An iterative curve-fitting baseline correction method for raman spectra driven by neural network. *Vibrational Spectroscopy*, Elsevier, v. 125, p. 103560, 2023. Disponível em: <<https://doi.org/10.1016/j.vibspec.2023.103560>>. Acesso em: 04 maio 2026.
- 16 ALMEIDA, C. T. de et al. Aplicação de algoritmos de aprendizagem de máquinas para classificação mineralógica utilizando espectros de raman. *nome do periódico*, volume, ano.
- 17 EWUSI-ANNAN, E. et al. Automatic preprocessing of laser-induced breakdown spectra using partial least squares regression and feed-forward artificial neural network: Applications to earth and mars data. *Spectrochimica Acta Part B: Atomic Spectroscopy*, v. 171, p. 105930, 2020.
- 18 GHOSH, K. et al. Deep learning spectroscopy: Neural networks for molecular excitation spectra. *Advanced Science*, v. 6, p. 1801367, 2019. ISSN 2198-3844.
- 19 SCHULZE, H. G. et al. Characteristics of backpropagation neural networks employed in the identification of neurotransmitter raman spectra. *Applied Spectroscopy*, v. 48, n. 1, p. 50–57, 1994. Disponível em: <<https://doi.org/10.1366/0003702944027688>>. Acesso em: 04 maio 2026.
- 20 SCHULZE, H. G. et al. Artificial neural network and classical least-squares methods for neurotransmitter mixture analysis. *Journal of Neuroscience Methods*, v. 56, p. 155–167, 1995.
- 21 LEWIS, I. et al. Raman spectrometry and neural networks for the classification of wood types—1. *Spectrochimica Acta Part A: Molecular Spectroscopy*, v. 50, n. 11, p. 1943–1958, 1994. ISSN 0584-8539. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0584853994802076>>.
- 22 WU, J. et al. A raman spectroscopy algorithm based on convolutional neural networks and multilayer perceptrons: qualitative and quantitative analyses of chemical warfare agent simulants. *Analyst*, Royal Society of Chemistry, v. 150, p. 1823–1836, 2025. Disponível em: <<https://doi.org/10.1039/d5an00075k>>. Acesso em: 04 maio 2026.
- 23 RAHMAN, M. H. U. et al. Machine learning-assisted raman spectroscopy and applications: A review. *Sensors*, v. 12?, 2024.
- 24 IBTEHAZ, N. et al. Ramannet: A generalized neural network architecture for raman spectroscopy. *Neural Computing and Applications / arXiv (preprint)*, 2022.

- 25 ALHILFI, D. N.; AL-KABBI, A. S. The role of cdsete layer and cdcl treatment on the electrochemical applications of cdte nanowires. *Ceramics International*, Elsevier Ltd and Techna Group S.r.l., v. 51, p. 25531–25542, 2025. Disponível em: <<https://doi.org/10.1016/j.ceramint.2025.03.235>>. Acesso em: 04 maio 2026.
- 26 KUWIK, M. et al. Studies on structural and spectroscopic properties of germanate-phosphate glasses doped with eu³⁺ ions. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, Elsevier, v. 344, p. 126658, 2026. Disponível em: <<https://doi.org/10.1016/j.saa.2025.126658>>. Acesso em: 04 maio 2026.
- 27 EL-MALLAWANY, R. A. *Glasses Glasses*. 1st. ed. s.l.: Springer, 2002.
- 28 EL-MALLAWANY, R. *Tellurite glass smart materials: Applications in optics and beyond*. s.l.: Springer, 2018.
- 29 DREYFUS, C.; DREYFUS, G. A machine learning approach to the estimation of the liquidus temperature of glass-forming oxide blends. *Journal of Non-Crystalline Solids*, v. 318, p. 63–78, 2003.
- 30 BRAUER, D. S.; RÜSSEL, C.; KRAFT, J. Solubility of glasses in the system p₂o₅–cao–mgo–na₂o–tio₂: Experimental and modeling using artificial neural networks. *Journal of Non-Crystalline Solids*, v. 353, p. 263–270, 2007.
- 31 KRISHNAN, A.; others. Predicting glass transition temperatures using neural networks. *Acta Materialia*, v. 159, p. 249–256, 2018.
- 32 CASSAR, D. R.; CARVALHO, A. C. P. L. F. de; ZANOTTO, E. D. Predicting glass transition temperatures using neural networks. *Acta Materialia*, v. 159, p. 249–256, 2018.
- 33 GLASSNET, A. of. Glassnet: A multitask deep neural network for predicting many properties of glasses. (*see SciDirect / publicação relacionada*), 2023.
- 34 MARTIELLO, S. M. M. et al. Machine learning unveils composition–property relationships in chalcogenide glasses. *Acta Materialia / related* (2022), 2022.
- 35 ISAEV, I. et al. Joint application of raman and optical absorption spectroscopy to determine concentrations of heavy metal ions in water using artificial neural networks. In: *Saratov Fall Meeting 2019: Laser Physics, Photonic Technologies, and Molecular Modeling*. [S.l.]: SPIE, 2020. v. 11458.
- 36 FANUCCHI, R. Z.; OLESKOVICZ, M. A. R. I. O.; BARBOSA, D. A. N. I. E. L. Análise da detecção de faltas de alta impedância utilizando redes neurais artificiais com topologias baseadas em perceptron multicamadas e redes rbf. In: *Simpósio Brasileiro de Automação Inteligente (SBAI)*. Fortaleza: [s.n.], 2013.
- 37 QUEIROZ, A. A. A. E. d. *Espectroscopia Raman e redes neurais artificiais na identificação de minerais dos grupos do pirocloro e da microlita*. Dissertação (Mestrado) — Universidade de São Paulo, 2019.
- 38 IBTEHAZ, N. et al. Ramannet: a generalized neural network architecture for raman spectrum analysis. *Neural Computing and Applications*, v. 35, n. 25, p. 18719–18735, 2023.

- 39 LUO, R.; POPP, J.; BOCKLITZ, T. Deep learning for raman spectroscopy: a review. *Analytica*, v. 3, n. 3, p. 287–301, 2022. ISSN 2673-4532. Disponível em: <<https://www.mdpi.com/2673-4532/3/3/20>>. Acesso em: 04 maio 2026.
- 40 ALVES, O.; MAZALI, I.; VIDROS, I. Cadernos temáticos—química nova na escola. *Edição especial*, 2001.
- 41 GONÇALVES, A. et al. *Estudo da dependência composicional com as propriedades térmicas e espectroscópicas do sistema TeO₂-Li₂O-BaO*. [S.l.], 2016.
- 42 SHELBY, J. E. *Introduction to Glass Science and Technology*. Cambridge: Royal Society of Chemistry, 1997. 20 p.
- 43 ZANOTTO, E. D.; MAURO, J. C. The glassy state of matter: Its definition and ultimate fate. *Journal of Non-Crystalline Solids*, Elsevier, v. 471, p. 490–495, 2017.
- 44 ZACHARIASEN, W. H. The atomic arrangement in glass. *Journal of the American Chemical Society*, ACS Publications, v. 54, n. 10, p. 3841–3851, 1932.
- 45 IDALGO, E. et al. Effects of the particle size and nucleation temperature on tellurite 20Li₂O-80TeO₂ glass crystallization. *Materials Science and Engineering: A*, Elsevier, v. 434, n. 1-2, p. 13–18, 2006.
- 46 SABADEL, J. et al. Structural and nonlinear optical characterizations of tellurium oxide-based glasses: TeO₂-BaO-TiO₂. *Journal of Solid State Chemistry*, Elsevier, v. 132, n. 2, p. 411–419, 1997.
- 47 BOUBATA, N.; ROULA, A.; MOUSSAOUI, I. Thermodynamic and relative approach to compute glass-forming ability of oxides. *Bulletin of Materials Science*, v. 36, p. 457–460, 2013. ISSN 0250-4707.
- 48 NOVATSKI, A. *Vidro aluminossilicato de cálcio dopado com Ti³⁺ ou Ce³⁺ para geração de alta taxa de luminescência e de luz branca inteligente*. Dissertação (Mestrado) — Universidade Estadual de Maringá, 2009.
- 49 JR, W. D. C.; RETHWISCH, D. G. *Fundamentals of Materials Science and Engineering: An Integrated Approach*. [S.l.]: John Wiley & Sons, 2012.
- 50 ZACHARIASEN, W. H. The atomic arrangement in glass. *Journal of the American Chemical Society*, ACS Publications, v. 54, n. 10, p. 3841–3851, 1932.
- 51 SMEKAL, A. G. On the structure of glass. *Journal of the Society of Glass Technology*, v. 35, p. 392–395, 1951.
- 52 STANWORTH, J. Tellurite glasses. *Nature*, Nature Publishing Group, v. 169, n. 4301, p. 581–582, 1952.
- 53 EL-MALLAWANY, R. A. *Tellurite glasses handbook: physical properties and data*. [S.l.]: CRC press, 2016.
- 54 THOMAS, P. The crystal structure and absolute optical chirality of paratellurite, α -TeO₂. *Journal of Physics C: Solid State Physics*, IOP Publishing, v. 21, n. 25, p. 4611, 1988.

- 55 GUNHA, J. V. *Caracterização estrutural e cinética de cristalização do sistema vítreo TeO_2 – Na_2O – TiO_2* . Dissertação (Dissertação (Mestrado em Ciências)) — Universidade Estadual de Ponta Grossa, Ponta Grossa, 2020.
- 56 ARNAUDOV, M. et al. Infrared-spectral investigation of tellurites. *Materials Research Bulletin*, Elsevier, v. 17, n. 9, p. 1121–1129, 1982.
- 57 NOVATSKI, A. Vidro aluminosilicato de cálcio dopado com Ti^{3+} ou Ce^{3+} para geração de alta taxa de luminescência e de luz branca inteligente. *Universidade Estadual de Maringá*, 2009.
- 58 STANWORTH, J. Tellurite glasses. *Nature*, Nature Publishing Group, v. 169, n. 4301, p. 581–582, 1952.
- 59 BERTHEREAU, A. et al. Determination of local geometries around tellurium in TeO_2 – Nb_2O_5 and TeO_2 – Al_2O_3 oxide glasses by xanes and exafs: Investigation of electronic properties of evidenced oxygen clusters by ab initio calculations. *Journal of Solid State Chemistry*, Elsevier, v. 126, n. 2, p. 143–151, 1996.
- 60 MORAES, J. C. S. et al. Relation among optical, thermal and thermo-optical properties and niobium concentration in tellurite glasses. *Journal of Non-Crystalline Solids*, Elsevier, v. 356, n. 41-42, p. 2146–2150, 2010.
- 61 GUIMARÃES, A. V. A. et al. Espalhamento stokes-anti-stokes: uma abordagem pela função de onda do fóton. *Universidade Federal de Minas Gerais*, 2022. Trabalho acadêmico.
- 62 SALA, O. Fundamentos da espectroscopia raman no infravermelho. *Universidade Estadual Paulista*, 1996. [S.l.].
- 63 MOLLING, A. Trabalho de Conclusão de Curso (Bacharelado em Engenharia Física), *Classificação de Espectros Raman de Microplásticos com Técnicas de Inteligência Artificial*. Porto Alegre - RS: [s.n.], 2024.
- 64 LI-CHAN, E.; CHALMERS, J.; GRIFFITHS, P. *Applications of vibrational spectroscopy in food science*. [S.l.]: John Wiley & Sons, 2011.
- 65 REIS, F. D. dos. Síntese e estudos espectroscópicos e estruturais de esquarínas derivadas da isoniazida. *Trabalho acadêmico*, 2014.
- 66 TRICHES, D. M. et al. Caracterização de ligas binárias nanoestruturadas dos sistemas zn-sb, ge-sb, al-sb e co-sb produzidas por síntese mecânica. *Florianópolis, SC*, 2009. Trabalho acadêmico.
- 67 CARVALHO, L. V. de et al. O espalhamento raman correlacionado stokes-anti-stokes: um estudo da influência da largura temporal do pulso de excitação. *Universidade Federal de Minas Gerais*, 2022. Trabalho acadêmico.
- 68 PAULA, J. de; ATKINS, P. W. Physical chemistry for the life sciences. *S.l.: s.n.*, 2005. Livro consultado sem informações de local e editora.
- 69 LARKIN, P. Infrared and raman spectroscopy: Principles and spectral interpretation. *S.l.: s.n.*, 2011.

- 70 CHOWDHRY, B. Z.; JABEEN, S.; ALEXANDER, B. Raman spectroscopy in pharmaceutical analysis. *European Pharmaceutical Review*, Russell Publishing Limited, v. 5, p. 34–39, 2009.
- 71 BRADY, G. W. Structure of tellurium oxide glass. *Journal of Chemical Physics*, v. 27, n. 1, p. 300–305, 1957.
- 72 DIMITRIEV, Y.; DIMITROV, V.; ARNAUDOV, M. Ir spectra and structures of tellurite glasses. *Journal of Materials Science*, v. 18, n. 5, p. 1353–1358, 1983.
- 73 SEKIYA, T.; MOCHIDA, N.; OHTSUKA, A. Raman spectra of mo-teo₂ (m = mg, sr, ba and zn) glasses. *Journal of Non-Crystalline Solids*, v. 168, p. 106–114, 1994.
- 74 SEKIYA, T.; MOCHIDA, N.; OGAWA, S. Structural study of moo₃–teo₂ glasses. *Journal of Non-Crystalline Solids*, v. 185, p. 135–144, 1995.
- 75 GUNHA, J. V. et al. Thermal, structural and optical properties of teo₂–na₂o–tio₂ glassy system. *Journal of Materials Science: Materials in Electronics*, v. 30, p. 16695–16701, 2019.
- 76 SHALTOUT, I.; BADR, Y. Effects of sm³⁺/yb³⁺ co-doping and temperature on the raman, ir spectra and structure of [teo₂–geo₂–k₂o–sm₂o₃/yb₂o₃] glasses. *Physica B: Condensed Matter*, v. 381, p. 187–193, 2006.
- 77 UMAIR, M. M.; YAHYA, A. K. Elastic and structural changes of xna₂o–(35–x)v₂o₅–65teo₂ glass system with increasing sodium. *Materials Chemistry and Physics*, v. 142, n. 2–3, p. 549–555, 2013.
- 78 KRINS, N. et al. Structural and electrical properties of tellurovanadate glasses containing li₂o. *Solid State Ionics*, v. 177, p. 3147–3150, 2006.
- 79 ZHOU, Y. et al. Characterization of new tellurite glasses and crystalline phases in the teo₂–pbo–bi₂o₃–b₂o₃ system. *Journal of Non-Crystalline Solids*, v. 386, p. 90–94, 2014.
- 80 ZUBEN, F. J. V. *Redes neurais aplicadas ao controle de máquina de indução*. Tese (Tese de Doutorado) — [s.n.], 1993.
- 81 MCCULLOCH, W. S.; PITTS, W. A logical calculus of the ideas immanent in nervous activity. *Bulletin of Mathematical Biophysics*, v. 5, p. 115–133, 1943.
- 82 WIENER, N. *Cybernetics*. Cambridge, MA: MIT Press, 1948.
- 83 GALLANT, A. R.; WHITE, H. There exists a neural network that does not make avoidable mistakes. In: *IEEE Second International Conference on Neural Networks*. San Diego, CA: [s.n.], 1980. p. 657–664.
- 84 ROSENBLATT, F. *The Perceptron: A perceiving and recognizing automaton*. [S.l.], 1957.
- 85 WIDROW, B. Generalization and information storage in networks of adaline "neurons". In: YOVITZ, M.; JACOBI, G.; GOLDSTEIN, G. (Ed.). *Self-Organizing Systems*. Washington, DC: Spartan Books, 1962. p. 435–461.
- 86 MINSKY, M.; PAPERT, S. *Perceptrons: An Introduction to Computational Geometry*. Cambridge, Massachusetts: MIT Press, 1969.

- 87 THE ROYAL SWEDISH ACADEMY OF SCIENCES. *The Nobel Prize in Physics 2024*. 2024. Awarded to John J. Hopfield and Geoffrey E. Hinton for foundational discoveries and inventions that enable machine learning with artificial neural networks. Disponível em: <<https://www.nobelprize.org/prizes/physics/2024/press-release/>>. Acesso em: 04 maio 2026.
- 88 ROSENBLATT, F. *The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain*. Buffalo, NY, 1958.
- 89 HAYKIN, S. *Redes neurais: princípios e prática*. 2. ed. Porto Alegre, RS: Bookman, 2001. Tradução de Paulo Martins Engel.
- 90 JR., O. J. S. *Análise de dados de instrumentação de túneis do metrô de São Paulo: uma abordagem por Redes Neurais*. 160 p. Dissertação (Dissertação (Mestrado em Geotecnia)) — Universidade de São Paulo, Escola de Engenharia de São Carlos, São Carlos, SP, 2006.
- 91 HERTZ, J. A. *Introduction to the Theory of Neural Computation*. [S.l.]: CRC Press, 2018.
- 92 NEAL, R. M. Connectionist learning of belief networks. *Artificial Intelligence*, v. 56, n. 1, p. 71–113, jul. 1992.
- 93 SZANDAŁA, T. Review and comparison of commonly used activation functions for deep neural networks. Wrocław, Poland, 2021.
- 94 KARPATY, A. *Yes you should understand backprop*. 2016. Disponível em: <<https://medium.com/@karpathy/yes-you-should-understand-backprop-e2f06eab496b>>. Acesso em: 04 maio 2026.
- 95 LIU, W. et al. A survey of deep neural network architectures and their applications. *Neurocomputing*, v. 234, p. 11–26, 2017.
- 96 MOLLAHOSSEINI, A.; CHAN, D.; MAHOOR, M. H. Going deeper in facial expression recognition using deep neural networks. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. [S.l.: s.n.], 2016. p. 1–10.
- 97 LECUN, Y. et al. Efficient backprop. In: *Neural Networks: Tricks of the Trade*. Berlin, Heidelberg: Springer, 1998. p. 9–50.
- 98 HAHNLOSER, R. et al. Digital selection and analogue amplification coexist in a cortex-inspired silicon circuit. *Nature*, v. 405, p. 947–951, julho 2000.
- 99 HAHNLOSER, R. H. R.; SEUNG, H. S.; SLOTINE, J.-J. Permitted and forbidden sets in symmetric threshold-linear networks. *Neural Computation*, v. 15, n. 3, p. 621–638, 2003.
- 100 PING, W. et al. Deep voice 3: Scaling text-to-speech with convolutional sequence learning. *arXiv preprint arXiv:1710.07654*, 2017.
- 101 BATAINEH, A. A. A comparative analysis of nonlinear machine learning algorithms for breast cancer detection. *International Journal of Machine Learning and Computing*, v. 9, n. 3, p. 248–254, June 2019.
- 102 BATAINEH, A. A.; KAUR, D.; JALALI, S. M. J. Multi-layer perceptron training optimization using nature inspired computing. *IEEE Access*, v. 10, p. 36963–36977, 2022.

- 103 BATAINEH, A. A.; MAIRAJ, A.; KAUR, D. Autoencoder based semi-supervised anomaly detection in turbofan engines. *International Journal of Advanced Computer Science and Applications*, v. 11, n. 11, p. 1–7, 2020.
- 104 BATAINEH, A. S. A. A gradient boosting regression based approach for energy consumption prediction in buildings. *Advances in Energy Research*, v. 6, n. 2, p. 91–101, 2019.
- 105 MANZAN, J. R. G. *Análise de desempenho de redes neurais artificiais do tipo multilayer perceptron por meio do distanciamento dos pontos do espaço de saída*. 129 p. Tese (Tese (Doutorado em Engenharia Elétrica)) — Universidade Federal de Uberlândia, Uberlândia, 2016.
- 106 ELITEDATASCIENCE. *Overfitting in Machine Learning*. 2025. Disponível em: <<https://elitedatascience.com/overfitting-in-machine-learning>>. Acesso em: set. 2025.
- 107 PARIS, G.; ROBILLIARD, D.; FONLUPT, C. Exploring overfitting in genetic programming. In: *Artificial Evolution, International Conference, Evolution Artificielle (EA 2003)*. Marseilles, France: DBLP, 2004. p. 267–277.
- 108 JENSEN, D. D.; COHEN, P. R. Multiple comparisons in induction algorithms. *Machine Learning*, v. 38, n. 3, p. 309–338, 2000.
- 109 YING, X. An overview of overfitting and its solutions. *Journal of Physics: Conference Series*, v. 1168, p. 022022, 2019. Disponível em: <<https://iopscience.iop.org/article/10.1088/1742-6596/1168/2/022022>>. Acesso em: 04 maio 2026.
- 110 BISHOP, C. M. *Neural Networks for Pattern Recognition*. Oxford: Oxford University Press, 1995.
- 111 GOODFELLOW, I.; BENGIO, Y.; COURVILLE, A. *Deep Learning*. Cambridge, MA: MIT Press, 2016.
- 112 HUANG, L. et al. Normalization techniques in training deep neural networks: methodology, analysis, and application. *arXiv preprint*, arXiv:2009.12836, 2020. Disponível em: <<https://arxiv.org/abs/2009.12836>>. Acesso em: 04 maio 2026.
- 113 BROWNLEE, J. *How to Improve Neural Network Stability and Modeling Performance With Data Scaling*. 2020. Machine Learning Mastery. Disponível em: <<https://machinelearningmastery.com/how-to-improve-neural-network-stability-and-modeling-performance-with-data-scaling/>>. Acesso em: 22 de novembro 2025.
- 114 Google Developers. *Numerical Data: Normalization*. 2025. Google Machine Learning Crash Course. Disponível em: <<https://developers.google.com/machine-learning/crash-course/numerical-data/normalization>>. Acesso em: 22 de novembro 2025.
- 115 GUNHA, J. V.; AL. et. Influence of sm2o3 ion concentration on physical and thermal properties, and raman gain coefficient of teo2–li2o–zno vitreous system. *Journal of Materials Science: Materials in Electronics*, v. 35, n. 3, p. 240, 2024.
- 116 DIAS, D. et al. Raman gain coefficient of er3+doped teo2-li2o–zno glasses. *J. Mater. Sci. Mater. Electron.*, v. 30, p. 16917–16921, 2019.
- 117 SAGAR, S.; BABU, S.; RAO, K. Emission spectroscopy of sm3+ion activated zinc phosphate glass for reddish–orange lighting applications. *J. Mater. Sci. Mater. Electron.*, 2023.

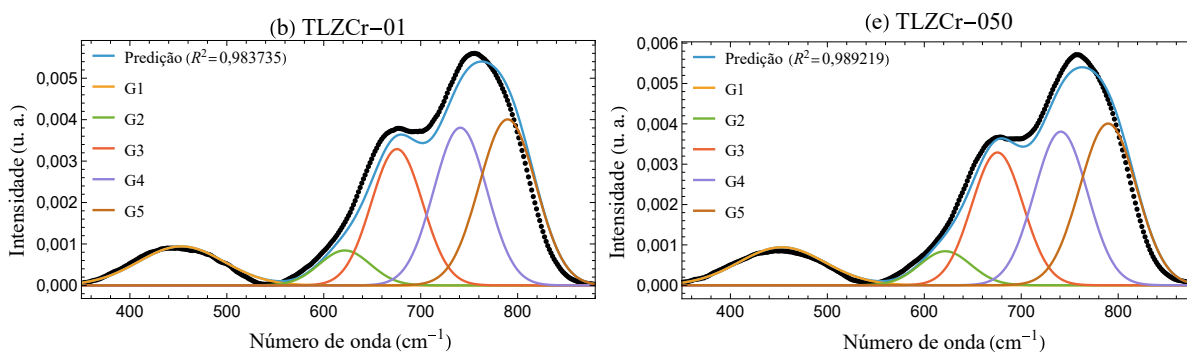
- 118 ELKHOSHKHANY, N.; MARZOUK, S.; SHAHIN, S. Synthesis and optical properties of new fluoro-tellurite glass within (teo₂–zno–lif–nb₂o₅–naf) system. *J. Non Cryst. Solids*, v. 472, p. 39–45, 2017.
- 119 CLERMONT-GALLERANDE, J. D. et al. New understanding of teo₂ e zno e na₂o ternary glass system. *J. Alloys Compd.*, 2021.
- 120 SQ, M. et al. Influence of sm₂o₃ ion concentration on structural and thermal properties. *J. Appl. Mech. Eng.*, 2016.
- 121 RDELEAN, I.; MURESAN, N.; PASCUTA, P. Ir and raman spectroscopic investigations of cr₂o₃-teo₂-b₂o₃-pbo glasses. *International Journal of Modern Physics B*, v. 18, n. 01, p. 95–101, 2004.

APÊNDICE A – RESULTADOS DA PREDIÇÃO DOS RESULTADOS EXPERIMENTAIS PELAS REDES SEM NORMALIZAÇÃO

• 2C-2

Para rede 2C-2Ag os valores de validação=0,0124544 e treinamento=0,251124 também não foram confiáveis, onde os valores da predição do espectroscopia também não foram confiáveis. Na Figura 42 e a tabela com os resultados do erro médio para a rede 2C-2Ag, considerando centro, largura e área.

Figura 42 – Resultados da predição 2C-2Ag na base 5.



Fonte: O autor.

Tabela 29 – Parâmetros das gaussianas ajustadas para as redes 2C-2Ag, das amostras TLZCr-01 e TLZCr-0,50 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		rea	
	Cr-0,10	Cr-0,75	Cr-0,10	Cr-0,75	Cr-0,10	Cr-0,75
Gaussiana 1	1,585	0,682	3,005	2,564	5,451	3,516
Gaussiana 2	1,083	1,083	0,641	1,096	5,091	4,074
Gaussiana 3	1,102	1,123	3,027	3,653	3,082	1,528
Gaussiana 4	1,211	1,212	7,095	8,140	3,782	4,679
Gaussiana 5	0,978	1,074	0,069	1,465	3,535	2,848

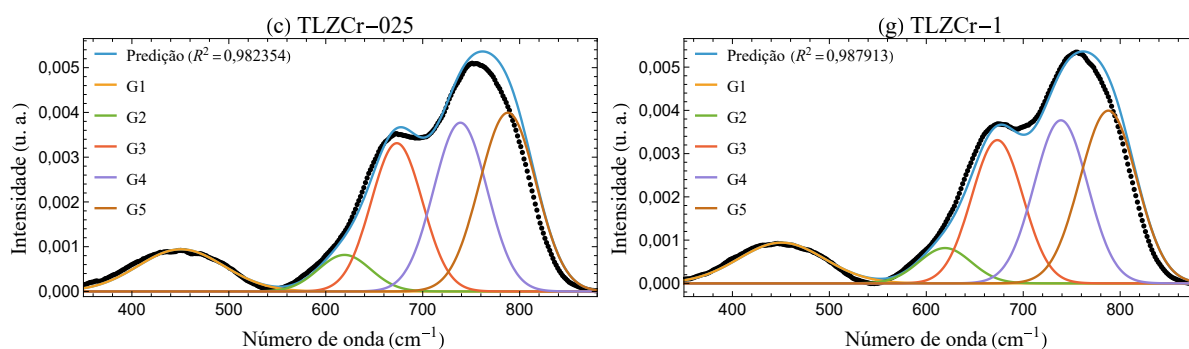
Fonte: O autor.

Os erros percentuais obtidos para os parâmetros ajustados foram de aproximadamente 1,585% para os centros da amostra TLZCr-01, 8,140% para as larguras da amostra TLZCr-050 e 5,451% para as áreas da amostra TLZCr-01. Esses resultados indicam que a rede não conseguiu aprender com precisão os padrões espectrais, mesmo diante de pequenas variações nos parâmetros.

• 2C-3

Para a rede 2C-3Ag, os valores de validação (0,0130404) e de treinamento (0,161988) foram semelhantes aos obtidos nas redes anteriores, indicando que essa configuração também não conseguiu realizar previsões satisfatórias. Na Figura 43, apresenta-se a os resultados do erro médio para a rede 2C-3Ag, considerando os parâmetros de centro, largura e área.

Figura 43 – Resultados da previsão 2C-3Ag na base 5.



Fonte: O autor.

Tabela 30 – Parâmetros das gaussianas ajustadas para as redes 2C-3Ag, das amostras TLZCr-0,25 e TLZCr-1 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-0,25	Cr-1	Cr-0,25	Cr-1	Cr-0,25	Cr-1
Gaussiana 1	1,029	1,029	7,844	1,558	5,931	0,400
Gaussiana 2	0,701	0,701	1,609	5,089	0,451	2,312
Gaussiana 3	0,794	0,774	2,449	2,688	8,726	3,598
Gaussiana 4	0,937	0,937	2,853	2,595	4,546	1,018
Gaussiana 5	0,742	0,724	0,629	1,074	6,197	3,561

Fonte: O autor.

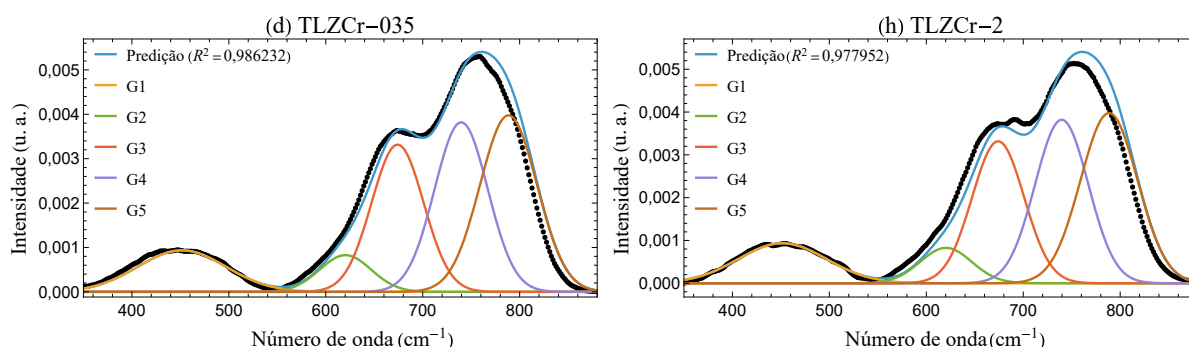
Os maiores erros percentuais obtidos para os parâmetros foram: 1,029% para os centros das amostras TLZCr-025 e TLZCr-1; 7,844% para as larguras da amostra TLZCr-025; e 8,726% para as áreas da mesma amostra (TLZCr-025). Esses resultados indicam que a rede não conseguiu aprender com precisão os padrões espectrais, mesmo diante de pequenas variações nos parâmetros.

• 2C-4

Os resultados da rede 2C-4Ag apresentaram valores de validação (0,0130421) e de treinamento (0,100704) abaixo do esperado, evidenciando que a rede não conseguiu realizar previsões com confiança. Na Figura 44, apresenta os resultados do erro médio para a rede 2C-4Ag, considerando os parâmetros de centro, largura e área.

Os maiores erros percentuais obtidos foram: 1,703% para os centros da amostra TLZCr-035; 5,089% para as larguras da amostra TLZCr-2; e 8,726% para as áreas da amostra

Figura 44 – Resultados da predição 2C-4Ag na base 5.



Fonte: O autor.

Tabela 31 – Parâmetros das gaussianas ajustadas para as redes 2C-4Ag, das amostras TLZCr-0,35 e TLZCr-2 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-0,35	Cr-2	Cr-0,35	Cr-2	Cr-0,35	Cr-2
Gaussiana 1	1,703	1,029	3,432	1,558	5,931	6,710
Gaussiana 2	0,852	0,701	0,028	5,089	0,451	3,454
Gaussiana 3	0,914	0,774	4,127	2,688	8,726	7,112
Gaussiana 4	1,050	0,937	2,668	2,595	4,546	6,154
Gaussiana 5	0,832	0,724	0,873	1,074	6,197	2,410

Fonte: O autor.

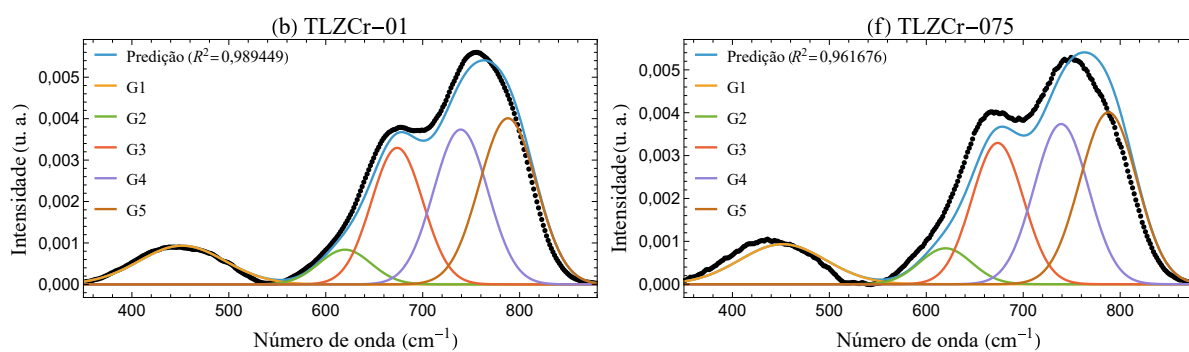
TLZCr-035. Esses resultados indicam que a rede não conseguiu aprender com precisão os padrões espectrais, mesmo diante de pequenas variações nos parâmetros.

• 2C-5

Para a última rede escolhida e analisada, 2C-5Ag, os resultados apresentaram valores de validação (0,0140105) e de treinamento (0,196987), também abaixo do esperado, evidenciando que a rede não conseguiu realizar predições com confiança. A seguir na Figura 45, apresenta os resultados do erro médio para a rede 2C-5Ag, considerando os parâmetros de centro, largura e área.

Os maiores erros percentuais para a última rede foram de 1,982% para os centros da amostra TLZCr-01 e 9,343% para as larguras da amostra TLZCr-01; e 15,605% para as intensidades da amostra TLZCr-075. Esses resultados indicam que a rede não conseguiu aprender com precisão os padrões espectrais, mesmo diante de pequenas variações nos parâmetros.

Figura 45 – Resultados da predição 2C-5Ag na base 5.



Fonte: O autor.

Tabela 32 – Parâmetros das gaussianas ajustadas para as redes 2C-4Ag, das amostras TLZCr-0,1 e TLZCr-75 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-0,1	Cr-0,75	Cr-0,1	Cr-0,75	Cr-0,1	Cr-0,75
Gaussiana 1	1,311	1,545	3,002	8,205	2,756	2,090
Gaussiana 2	0,783	0,933	2,422	0,983	3,003	12,090
Gaussiana 3	0,848	0,828	3,402	4,203	2,058	15,605
Gaussiana 4	1,982	0,975	9,343	6,449	4,448	7,740
Gaussiana 5	0,754	0,856	0,433	2,811	2,425	3,758

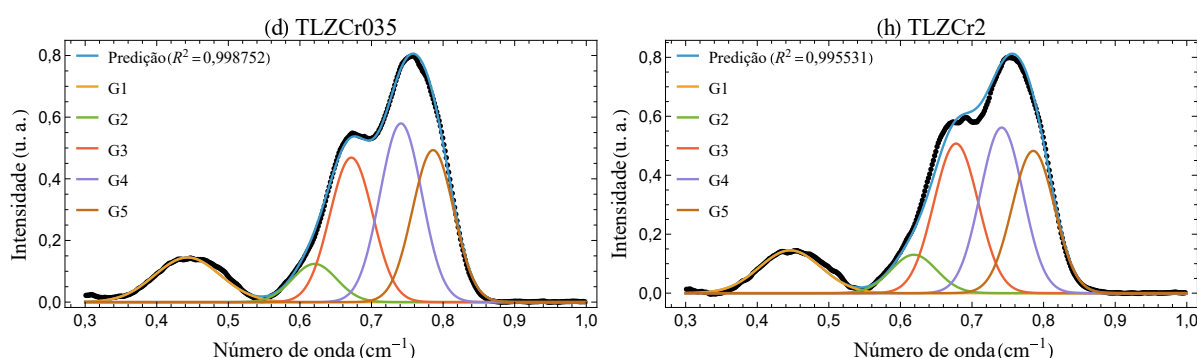
Fonte: O autor.

APÊNDICE B – RESULTADOS DA PREDIÇÃO EXPERIMENTAL PELAS REDES COM NORMALIZAÇÃO

• 1C-1Ag base 5

Para a rede 1C-1Ag na base 5, os valores de validação (0,0000402532) e treinamento (0,0000402417) apresentaram-se bastante satisfatórios, sem indícios de *overfitting*. Dessa forma, os resultados obtidos foram aplicados na análise dos espectros Raman de todas as amostras de vidros teluretos dopados com Cr. A rede foi capaz de realizar a predição de forma satisfatória para todas as amostras, apresentando um coeficiente de determinação $R^2 = 0,9989$ para a amostra TLZCr01. Dessa forma, o desempenho do modelo em relação aos dados experimentais pode ser considerado altamente positivo, mantendo os centros das bandas dentro de margens de erro aceitáveis e proporcionando um bom ajuste global do perfil espectral. Esses resultados reforçam a eficácia da arquitetura 1C-1Ag, demonstrando sua capacidade de realizar previsões confiáveis mesmo diante de variações entre dados simulados e experimentais. As previsões menos precisas da rede foram observadas para as amostras TLZCr035 e TLZCr2 na Figura 46.

Figura 46 – Resultados da predição 1C-1Ag na base 5.



Fonte: O autor.

Ainda assim, os resultados obtidos podem ser considerados bastante satisfatórios na tabela 17.

Tabela 33 – Resultados da rede 1C-1Ag na base 5.

Gaussiana	Centro		Largura		Área	
	Cr-0,35	Cr-2	Cr-0,35	Cr-2	Cr-0,35	Cr-2
Gaussiana 1	0,566	1,106	3,860	5,502	4,063	2,500
Gaussiana 2	0,185	0,509	19,551	18,786	10,137	17,580
Gaussiana 3	0,477	0,806	16,666	16,217	14,339	20,677
Gaussiana 4	1,310	0,906	3,882	1,046	4,588	6,724
Gaussiana 5	0,280	0,125	1,069	0,651	20,866	11,174

Fonte: O autor.

Na amostra TLZCr035, os maiores desvios foram observados na largura (16,226) da gaussiana 2 e na área foi de (16,972) na gaussiana 5, quando comparados aos valores obtidos pelo *software Origin*. Ainda assim, tais erros podem ser considerados baixos, demonstrando a boa precisão do método empregado. De forma semelhante, na amostra TLZCr2 os desvios foram próximos, com valores de 15,769 na largura da gaussiana 3 e 21,719 na área da gaussiana 3. Na rede 1C-1Ag, utilizando a base 6, os valores de validação (0,000361863) e treinamento (0,0000388303) apresentaram-se dentro do esperado, não indicando sobreajuste do modelo. Assim, os resultados obtidos foram considerados na análise dos espectros Raman. Nessa configuração, verificou-se um aumento nos erros de predição das amostras em comparação com a base 5; contudo, a rede manteve desempenho satisfatório, alcançando um coeficiente de determinação de $R^2 = 0,9980$ para a amostra TLZCr-025. Assim, embora o desempenho tenha apresentado ligeira redução em relação à base anterior, os resultados podem ser considerados altamente positivos, uma vez que os centros das bandas permaneceram dentro de margens de erro aceitáveis, garantindo um ajuste global consistente do perfil espectral. Os maiores erros foram observados na amostra TLZCr-2, com desvios mais significativos na largura (16,237) e na área (23,886) da Gaussiana 3, quando comparados aos valores obtidos pelo *software Origin*. Ainda assim, tais desvios podem ser considerados relativamente baixos, evidenciando a boa precisão do método empregado. De forma semelhante, para a mesma amostra, foram obtidos valores próximos, correspondentes a 15,769 para a largura e 21,719 para a área da Gaussiana 3.

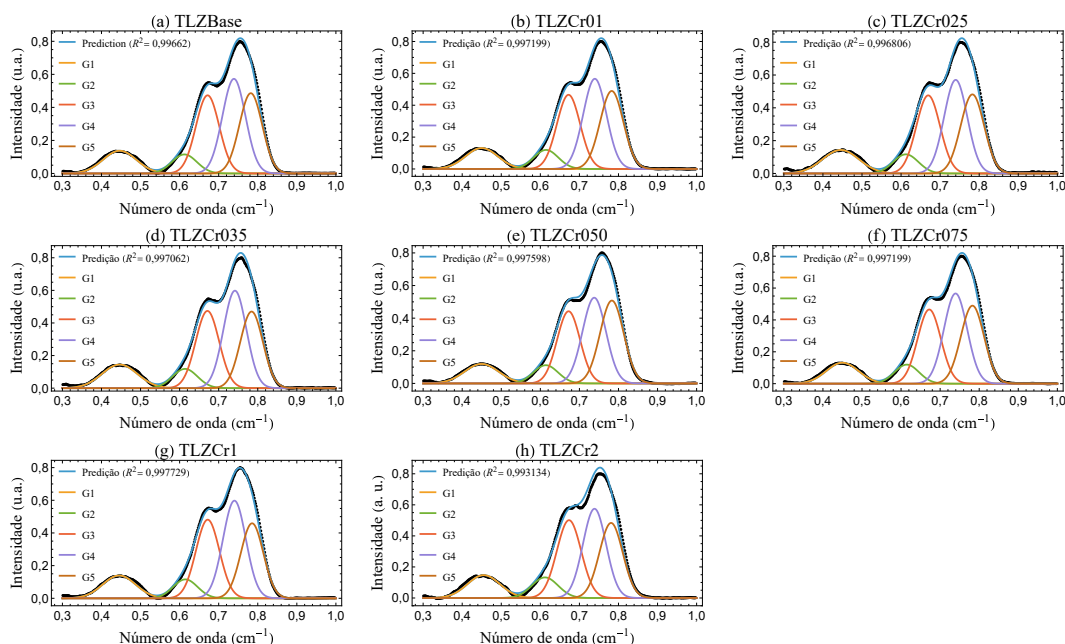
- 1C-2Ag base 5

A rede 1C-2Ag, utilizando a base 5, apresentou valores de validação (0,0000410749) e de treinamento (0,0000403637) extremamente baixos. Esses resultados evidenciam que a rede foi capaz de predizer de forma eficiente os dados provenientes da espectroscopia Raman. Além disso, a proximidade entre os valores de treinamento e validação demonstra que não houve ocorrência de *overfitting*, o que confirma a boa capacidade de generalização do modelo. Assim, a configuração da rede 1C-2Ag na base 5 revelou-se bastante eficiente para a tarefa proposta, proporcionando resultados consistentes e confiáveis no ajuste dos espectros analisados.

Na Tabela 35 com as amostras TLZCr01 e TLZCr1 da rede 1C-2Ag na base 5.

Na amostra TLZCr01, os maiores desvios foram observados na largura (19,195) e, para a amostra TLZCr1, o valor foi de 21,041, ambos referentes à gaussiana 2. Em relação à área, verificou-se um desvio de 19,511 para TLZCr01 e de 21,901 para TLZCr1, ambos associados à gaussiana 5, quando comparados aos valores obtidos no *software Origin*. Ainda assim, tais desvios podem ser considerados baixos, evidenciando a boa precisão do método empregado.

Figura 47 – Resultados da predição 1C-2Ag na base 5.



Fonte: O autor.

Tabela 34 – Resultados da rede 1C-2Ag na base 5.

Gaussiana	Centro		Largura		Área	
	Cr-0,1	Cr-1	Cr-0,1	Cr-1	Cr-0,1	Cr-1
Gaussiana 1	0,866	0,169	4,291	3,400	10,478	5,626
Gaussiana 2	0,059	0,048	19,195	21,041	7,353	11,411
Gaussiana 3	0,672	0,559	15,778	16,833	7,596	14,156
Gaussiana 4	1,017	1,133	16,388	3,700	2,688	11,337
Gaussiana 5	0,086	0,447	1,810	0,513	19,511	21,901

Fonte: O autor.

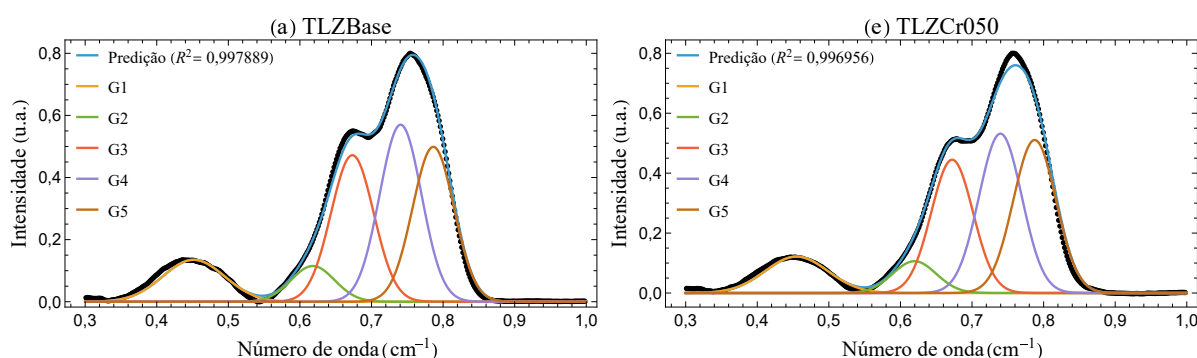
Para mesma rede 1C-2Ag na base 6 os valores de validação=0,0000383517 e treinamento=0,0000371371 também foram como o esperado a rede conseguiu fazer a predição. Para a amostra TLZCr01, a rede também foi capaz de realizar a predição, apresentando como maiores erros 16,172% na gaussiana 2 e 21,7013% na área, referente à gaussiana 5. Logo para a amostra TLZCr1, os erros foram de 19,653% na gaussiana 2 e 21,890% na área, igualmente na gaussiana 5. Em relação às redes 1C-2Ag na base 5, os valores obtidos mostraram-se bastante semelhantes, não sendo observada uma melhora significativa na base 6.

- 1C-3Ag base 5

Para a rede 1C-3Ag na base 5, os valores de validação e treinamento apresentaram-se bastante satisfatórios, uma vez que o erro sofreu uma queda acentuada, indicando que a rede rapidamente captou as principais características dos dados. Além disso, os erros

finais de treinamento (0,0000408299) e validação (0,0000352661) permaneceram muito próximos, evidenciando a ausência de sobreajuste (*overfitting*). Essa proximidade entre as curvas confirma que a rede não apenas aprendeu os padrões presentes no conjunto de treinamento, mas também foi capaz de generalizar de forma adequada para novos dados. Nos resultados experimentais, a rede obteve boa predição para todas as amostras, sendo apresentados a seguir na Figura 48 os valores dos erros para TLZBase e TLZCr050 da rede 1C-3Ag base 5.

Figura 48 – Resultados da predição 1C-3Ag na base 5.



Fonte: O autor.

Assim temos a Tabela 35 com as amostras TLZBase e TLZCr-050 da rede 1C-3Ag na base 5.

Tabela 35 – Parâmetros das gaussianas ajustadas para as redes 1C-3Ag, das amostras TLZBase e TLZCr-050 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	TLZBase	Cr-050	TLZBase	Cr-050	TLZBase	Cr-50
Gaussiana 1	0,807	1,096	2,499	0,711	6,540	8,730
Gaussiana 2	0,489	0,607	13,001	14,067	12,432	2,030
Gaussiana 3	0,783	0,635	10,826	14,093	5,328	7,760
Gaussiana 4	0,913	1,043	1,797	0,213	4,237	2,420
Gaussiana 5	0,552	0,801	0,071	0,665	12,929	18,885

Fonte: O autor.

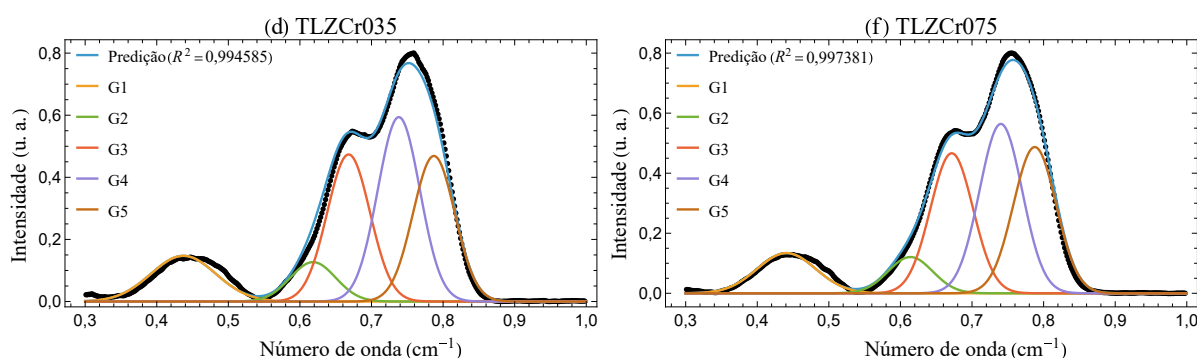
Para a amostra TLZBase, os maiores desvios observados foram de 13,001% na largura da Gaussiana 2 e de 12,929% na área da Gaussiana 5. Para a rede TLZCr050, o maior erro relacionado à largura foi de 14,090% na Gaussiana 3, enquanto para a área registrou-se 18,885% na Gaussiana 5, em comparação com os valores obtidos pelo *software Origin*. Apesar dessas diferenças, os desvios podem ser considerados relativamente modestos, indicando uma boa precisão do método utilizado. Na configuração 1C-3Ag aplicada à base 6, os valores de validação (0,0000348507) e de treinamento (0,0000368163) mostraram-se bastante próximos, indicando que a rede conseguiu realizar a predição dos resultados da espectroscopia Raman sem ocorrência de *overfitting*. Para a amostra TLZBase, os maiores

desvios foram de 14,253% na largura da Gaussiana 2 e de 13,951% na área, resultados muito semelhantes aos obtidos na base 5. De forma análoga, para a amostra TLZCr050, o erro na largura foi de 15,350%, enquanto na área atingiu 18,592%, valores também próximos aos verificados na rede correspondente à base 5.

- 1C-4Ag base 5

Na rede 1C-4Ag aplicada à base 5, os valores de validação e treinamento apresentaram um aumento em comparação às redes anteriores, indicando que a capacidade de aprendizado da rede foi ligeiramente inferior. Ainda assim, os resultados podem ser considerados satisfatórios. Observa-se que os erros finais de treinamento (0,0000396432) e validação (0,0000449189) permaneceram muito próximos, o que confirma a ausência de sobreajuste (overfitting). Nos resultados experimentais, a rede apresentou boa capacidade de predição para todas as amostras. Na Figura 49, são apresentados os valores de erro referentes às amostras TLZCr035 e TLZCr075, obtidos com a rede 1C-4Ag na base 5.

Figura 49 – Resultados da predição 1C-4Ag na base 5.



Fonte: O autor.

Temos na Tabela 36 com as amostras TLZBase e TLZCr-050 da rede 1C-3Ag na base 5.

Tabela 36 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0.35 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr035	Cr-075	Cr035	Cr-075	Cr035	Cr-075
Gaussiana 1	1,525	0,056	3,464	0,711	2,120	13,448
Gaussiana 2	0,533	14,908	19,004	14,067	17,603	1,974
Gaussiana 3	0,004	7,244	16,1108	7,244	13,739	10,853
Gaussiana 4	0,845	13,496	4,356	13,496	5,238	15,607
Gaussiana 5	0,661	0,920	0,0202	0,920	20,083	14,332

Fonte: O autor.

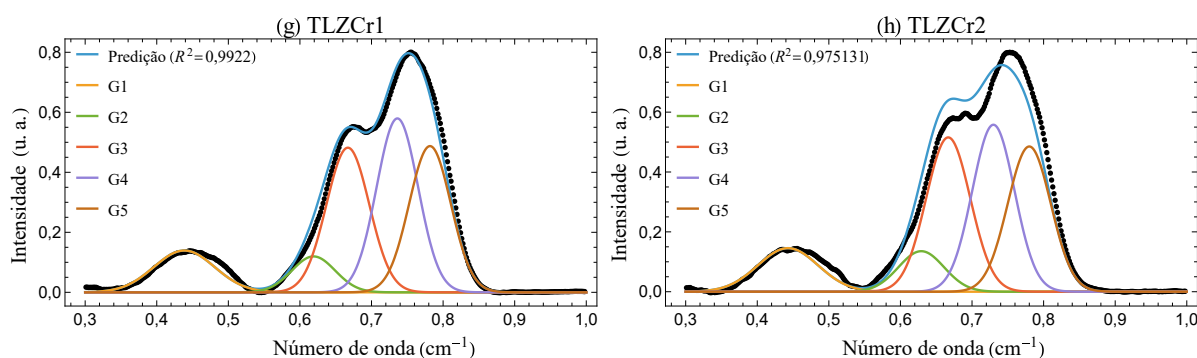
Então para a rede 1C-4Ag na base 5, teve seu erro na largura de 19,004 da Gaussiana 2 e de 20,038% na área também na Gaussiana 5. Para a rede TLZCr075, o maior erro relacionado

à largura foi de 13,496% na Gaussiana 4, enquanto para a área registrou-se 15,607% na Gaussiana 5, em comparação com os valores obtidos pelo *software Origin*. Apesar dessas diferenças, os desvios podem ser considerados relativamente modestos, indicando uma boa precisão do método utilizado. Na rede 1C-4Ag aplicada à base 6, os valores de validação (0,0000414596) e treinamento (0,0000390563) permaneceram bastante próximos, o que evidencia a capacidade da rede em realizar previsões confiáveis dos espectros Raman sem ocorrência de *overfitting*. Em comparação com a base 5, observa-se uma melhoria no desempenho da rede. Para a amostra TLZCr035, os maiores desvios observados foram de 16,914% na largura da Gaussiana 3 e de 21,066% na área, valores bastante próximos aos registrados na base 5. De forma semelhante, na amostra TLZCr075, os erros foram de 15,731% na largura e 15,279% na área, confirmando que a rede apresentou resultados superiores em relação à base 5.

- 2C-3Ag base 5

Foram também testadas redes com duas camadas, porém a maioria apresentou resultados indicativos de *overfitting*, o que compromete a sua confiabilidade. No entanto, a rede 2C-3Ag destacou-se por não apresentar *overfitting*, sendo a arquitetura de duas camadas com maior potencial. Os valores obtidos de validação (0,0000423814) e treinamento (0,0000410029) confirmam a ausência de *overfitting* e a consistência do modelo.

Figura 50 – Resultados da predição 2C-3Ag na base 5.



Fonte: O autor.

Nos resultados experimentais da espectroscopia Raman das amostras de vidros teluretos, a amostra TLZCr1 apresentou o melhor desempenho, com erro de 17,6487% na largura da Gaussiana 2 e de 15,9587% na área da Gaussiana 5. Para a amostra TLZCr2, a rede não conseguiu realizar a predição de forma adequada, apresentando erros mais elevados e indicando uma elevada probabilidade de ocorrência de *overfitting*. No caso da rede 2C-3Ag aplicada à base 6, os resultados evidenciaram a ocorrência de *overfitting*, impossibilitando a predição confiável das amostras, o que demonstra a limitação dessa arquitetura.

Tabela 37 – Parâmetros das gaussianas ajustadas para as redes 2C-3Ag, das amostras TLZCr-0.35 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-1	Cr-2	Cr-1	Cr-2	Cr-1	Cr-2
Gaussiana 1	1,701	1,572	4,032	2,651	3,220	5,084
Gaussiana 2	0,648	2,289	17,897	17,648	14,034	11,923
Gaussiana 3	0,171	0,184	16,758	16,465	12,452	13,971
Gaussiana 4	0,565	0,616	3,033	2,333	11,428	6,553
Gaussiana 5	0,039	0,443	0,869	0,826	0,620	15,958

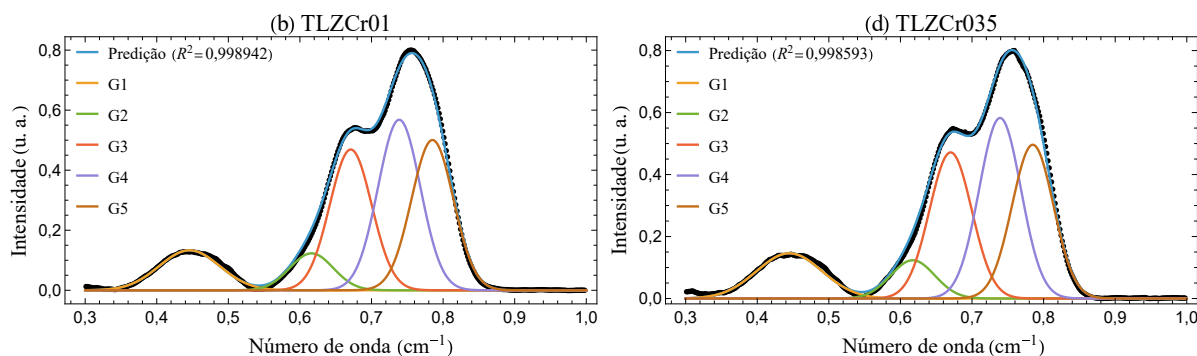
Fonte: O autor.

APÊNDICE C – RESULTADOS DAS DA PREDIÇÃO DOS RESULTADOS EXPERIMENTAIS PELAS REDES COM NORMALIZAÇÃO E 1% DE RUÍDO

- 1C-1Ag base 5 – 1C-1Ag base 6 com um 1% de ruído

Para a primeira rede 1C-1Ag com 1% de ruído os valores da validação= 0,0000487319 e treinamento= 0,0000448486 não teve overfitting mas os valores aumentaram isso quer dizer que a rede aprendeu um pouco menos o que é normal já que estamos adicionando ruído no resultados teóricos. Assim temos a Figura 51 os erros das amostras TLZCr01 e TLZCr035 do centro, largura e área.

Figura 51 – Resultados da predição 1C-1Ag na base 5.



Fonte: O autor.

Tabela 38 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0.1 e TLZCr-035 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-01	Cr-035	Cr-01	Cr-035	Cr-01	Cr-035
Gaussiana 1	0,074	0,177	2,035	2,876	7,103	2,311
Gaussiana 2	0,196	0,356	15,521	15,614	7,176	11,124
Gaussiana 3	0,445	0,590	14,092	14,649	6,849	11,819
Gaussiana 4	0,916	0,980	14,479	3,606	2,258	10,546
Gaussiana 5	0,408	0,365	0,413	0,827	18,812	16,236

Fonte: O autor.

Os resultados da predição da espectroscopia Raman para a amostra TLZCr01 apresentaram erro máximo de 15,839% na largura da Gaussiana 4 e de 14,489% na área da Gaussiana 5. Para a amostra TLZCr035, o maior erro observado foi de 14,861% na largura da Gaussiana 4 e de 14,489% na área da Gaussiana 5.

Na rede 1C-1Ag utilizando a base 6, os valores obtidos para validação (0,0000404482) e treinamento (0,0000417963) apresentaram tendência de diminuição, indicando que a

rede está aprendendo de forma adequada, sem ocorrência de *overfitting*. Assim, para as amostras de vidros teluretos, verificou-se que a rede conseguiu realizar a predição para a TLZCr01, embora tenha ocorrido um aumento nos erros em comparação com a base 5, atingindo 15,415% na largura da Gaussiana 2 e 18,822% na área da Gaussiana 5. Para a amostra TLZCr035, os resultados foram semelhantes, com erro de 16,033% na largura da Gaussiana 3 e de 15,996% na área da Gaussiana 5.

- 1C-2Ag base 5 – 1C-2Ag base 6 com um 1% de ruído

A rede 1C-2Ag, com adição de 1% de ruído, apresentou valores de validação (0,000047275) e de treinamento (0,0000473252) bastante próximos, ambos com tendência de diminuição. Esse comportamento indica ausência de *overfitting*. Em comparação, a mesma rede sem a adição de 1% de ruído apresentou valores mais elevados. Na Tabela 39, são apresentados os erros obtidos para as amostras TLZCr050 e TLZCr2, considerando os parâmetros de centro, largura e área.

Tabela 39 – Parâmetros das gaussianas ajustadas para as redes 1C-2Ag, das amostras TLZCr-0.50 e TLZCr-2 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-050	Cr-2	Cr-050	Cr-2	Cr-050	Cr-2
Gaussiana 1	0,248	0,566	1,335	3,445	9,562	0,829
Gaussiana 2	0,132	0,204	15,137	14,957	8,364	17,052
Gaussiana 3	0,583	0,534	15,016	16,102	7,949	21,122
Gaussiana 4	0,930	1,114	0,220	1,426	0,517	8,170
Gaussiana 5	0,504	0,316	0,213	1,456	22,396	14,015

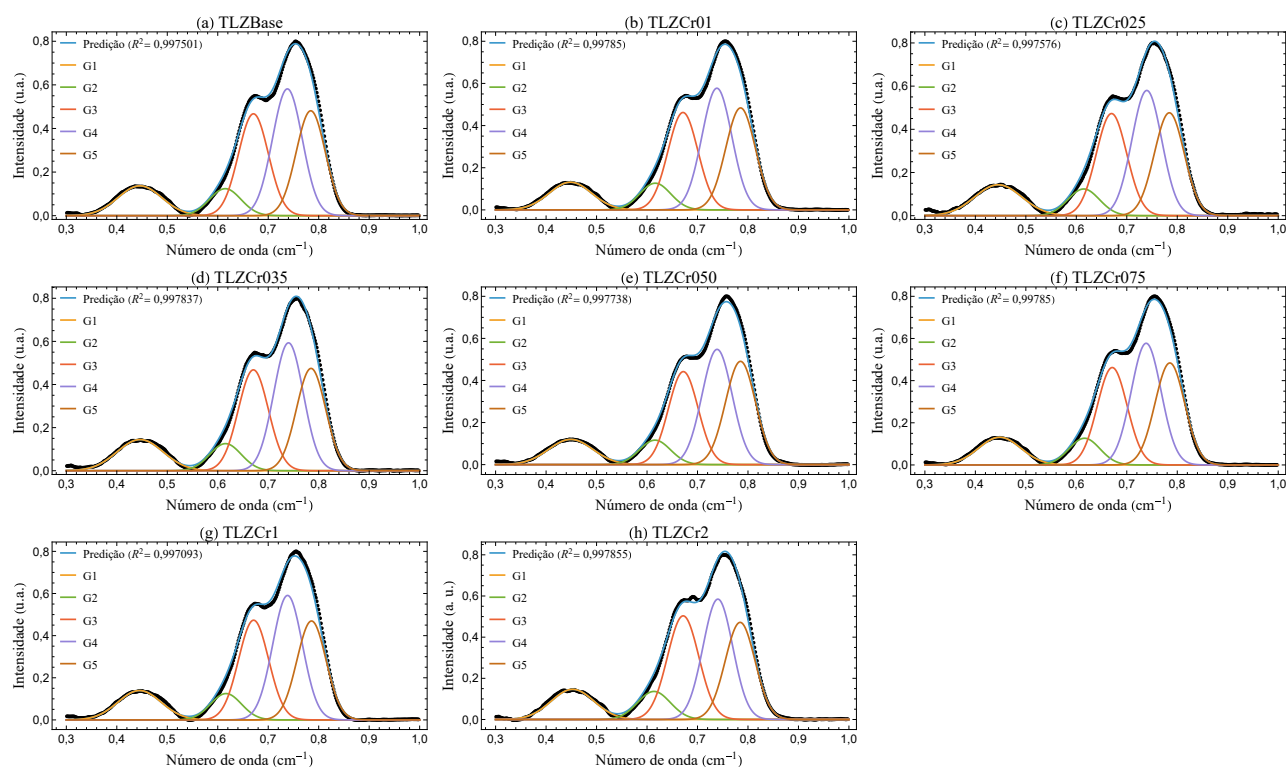
Fonte: O autor.

Assim, os resultados obtidos com a rede 1C-2Ag na base 5 para a predição da espectroscopia Raman das amostras mostraram-se consistentes. Ainda que tenham sido observados erros, o desempenho da rede em relação aos dados experimentais pode ser considerado satisfatório, uma vez que os centros das bandas permaneceram dentro de margens de erro aceitáveis, apresentando, além disso, um bom ajuste global do perfil espectral.

Na amostra TLZCr050, a rede 1C-2Ag apresentou seu maior erro médio na largura da Gaussiana 2, atingindo 15,137%, enquanto para a área o valor máximo foi de 22,396% na Gaussiana 5. Por sua vez, para a amostra TLZCr2, o maior erro identificado correspondeu a 16,102% na largura da Gaussiana 2 e a 21,122% na área da Gaussiana 3.

A rede 1C-2Ag, na base 6, apresentou valores de validação (0,0000414263) e de treinamento (0,0000427483) inferiores aos obtidos na base 5, indicando uma melhora no desempenho, já que os erros diminuíram. Esse resultado evidencia que a rede na base 6 apresentou maior capacidade de aprendizado em comparação à base anterior, sem ocorrência de *overfitting*.

Figura 52 – Resultados da predição 1C-2Ag na base 5.



Fonte: O autor.

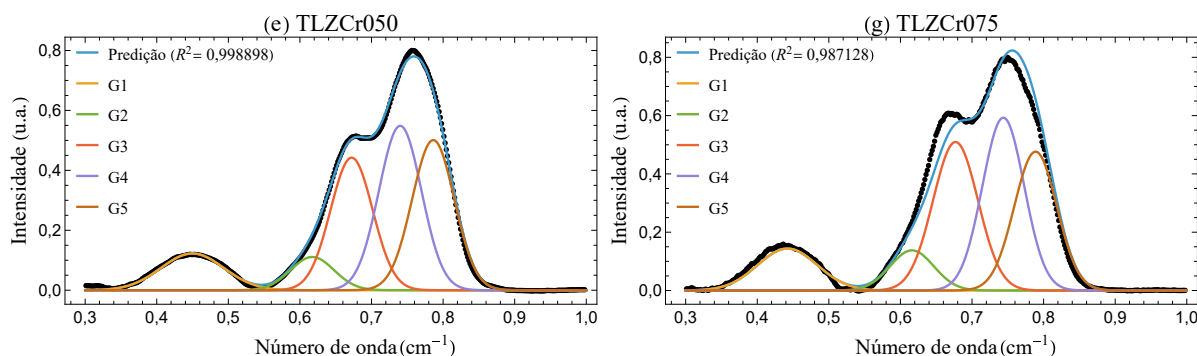
Para a amostra TLZCr050, os erros alcançaram 17,097% na largura da Gaussiana 3 e 22,430% na área da Gaussiana 5, valores que se mantiveram muito próximos aos obtidos na base 5, sugerindo que a mudança de base não resultou em melhorias significativas para essa amostra. No caso da TLZCr2, observaram-se erros de 18,937% na largura da Gaussiana 3 e de 23,366% na área, também na Gaussiana 3. Esses resultados evidenciam que, embora a rede tenha mantido consistência no processo de predição, os erros permanecem relativamente elevados, indicando possíveis limitações na capacidade de ajuste fino do modelo para essas amostras específicas.

- 1C-3Ag base 5 – 1C-3Ag base 6 com um 1% de ruído

Para a rede 1C-3Ag, com a inserção de 1% de ruído na base de dados, os valores de erro de validação e treinamento apresentaram um ligeiro aumento em comparação com a mesma rede sem ruído, atingindo respectivamente 0,0000492224 e 0,0000455176. Esses resultados podem ser considerados promissores, uma vez que não foram observados indícios de overfitting, indicando que a rede manteve sua capacidade de generalização e conseguiu realizar de forma adequada a predição dos espectros Raman. No caso dos resultados obtidos por espectroscopia Raman dos vidros teluretos, foram selecionadas, para fins de comparação, a amostra que apresentou o melhor desempenho TLZCr050 e

aquela que apresentou o maior erro TLZCr075. Esses dados estão organizados na Tabela 29, a fim de evidenciar de forma clara as diferenças de ajuste entre os extremos do conjunto analisado.

Figura 53 – Resultados da predição 1C-3Ag na base 5.



Fonte: O autor.

Tabela 40 – Parâmetros das gaussianas ajustadas para as redes 1C-3Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-050	Cr-75	Cr-050	Cr-75	Cr-050	Cr-75
Gaussiana 1	0,045	0,613	1,383	1,816	7,482	5,378
Gaussiana 2	0,361	0,320	13,835	12,478	2,131	9,484
Gaussiana 3	0,630	1,319	15,116	13,276	8,106	2,856
Gaussiana 4	1,102	1,593	0,347	10,963	0,093	18,795
Gaussiana 5	0,637	0,967	0,596	1,454	21,430	16,688

Fonte: O autor.

Para a rede 1C-3Ag, aplicada às amostras de vidros teluretos, observou-se que a amostra TLZCr050 apresentou o melhor desempenho, ainda que com erros residuais de 15,116% na largura da gaussiana 3 e de 21,430% na área da gaussiana 5. Em contraste, a amostra TLZCr075 obteve o pior ajuste entre as avaliadas, com erros de 13,276% e 12,478% na largura das gaussianas 3 e 2, respectivamente, e de 18,795% na área da gaussiana 4. Esses resultados evidenciam que, mesmo nas condições de melhor predição, ainda persistem discrepâncias significativas em parâmetros específicos, o que sugere limitações da rede em capturar integralmente as variações estruturais presentes nos espectros Raman das diferentes composições vítreas.

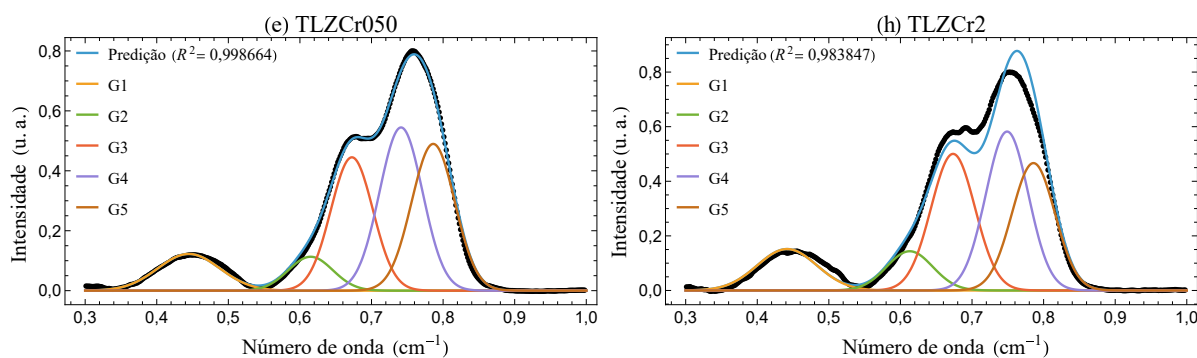
Na configuração da rede 1C-3Ag utilizando a base 6 com a inserção de 1% de ruído, os valores de erro de validação (0,0000406174) e de treinamento (0,0000418765) apresentaram redução em comparação à mesma rede treinada na base 5. Esse comportamento indica que a rede foi capaz de aprender de forma mais eficiente, apresentando menor discrepância entre os conjuntos de treinamento e validação. No que se refere ao desempenho sobre as amostras, a TLZCr050 apresentou erros semelhantes aos observados na base 5,

com 15,967% na largura da gaussiana 3 e 20,705% na área da gaussiana 5. A amostra TLZCr075 manteve também um comportamento próximo ao obtido anteriormente, com erros de 13,9% e 14,539% na largura das gaussianas 2 e 3, respectivamente, e de 18,933% na área da gaussiana 4. Esses resultados sugerem que, embora a mudança da base tenha favorecido o aprendizado global da rede, o impacto sobre a predição individual das amostras foi limitado, permanecendo erros significativos em parâmetros específicos.

- 1C-4Ag base 5 – 1C-4Ag base 6 com um 1% de ruído

Para a rede 1C-4Ag, utilizando a base 5, observou-se um aumento nos valores de erro de predição em comparação à rede anterior, sugerindo que essa configuração apresentou menor capacidade de aprendizado. Os valores obtidos foram de 0,0000538729 para a validação e 0,000049733 para o treinamento. Ainda assim, tais resultados podem ser considerados promissores, uma vez que não foram identificados indícios de *overfitting*, indicando que a rede manteve capacidade de generalização adequada. Na Figura 54, são apresentados os resultados das amostras de vidros teluretos que obtiveram, respectivamente, a melhor predição TLZCr050 e a menor predição TLZCr2, de forma a evidenciar os extremos de desempenho da rede no ajuste dos espectros Raman.

Figura 54 – Resultados da predição 1C-4Ag na base 5.



Fonte: O autor.

Tabela 41 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0.50 e TLZCr-075 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-050	Cr-2	Cr-050	Cr-2	Cr-050	Cr-2
Gaussiana 1	1,167	1,854	0,320	5,406	8,451	0,481
Gaussiana 2	0,014	0,339	16,201	14,135	6,363	22,483
Gaussiana 3	0,687	0,822	14,909	14,436	8,575	18,631
Gaussiana 4	1,266	2,322	0,710	0,627	0,435	8,566
Gaussiana 5	0,635	0,600	1,568	0,267	21,517	143,471

Fonte: O autor.

A amostra TLZCr050 apresentou uma predição satisfatória, com erro máximo de 16,201% na largura da gaussiana 2 e de 21,517% na área da gaussiana 5. Em contrapartida, a amostra TLZCr2 não obteve desempenho adequado, indicando que a rede não conseguiu aprender de forma consistente seus padrões espectrais. Além disso, observam-se indícios de overfitting para essa amostra, o que sugere uma limitação da rede em generalizar frente às características estruturais específicas dessa composição vítrea.

Na configuração da rede 1C-4Ag utilizando a base 6, observou-se uma discreta melhora nos valores de erro em comparação à base 5, alcançando 0,0000497945 para a validação e 0,0000460827 para o treinamento. Assim como no caso anterior, não foram identificados indícios de overfitting, o que indica que a rede manteve sua capacidade de generalização.

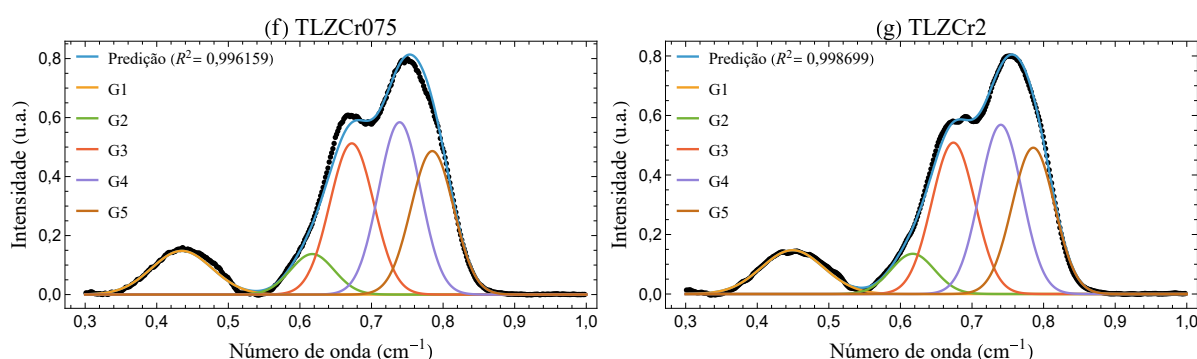
Para a amostra TLZCr050, não foram observadas mudanças significativas em relação à base 5, apresentando erro máximo de 15,582% na largura da gaussiana 3 e de 22,674% na área da gaussiana 5. Para a amostra TLZCr2, os resultados permaneceram praticamente inalterados, evidenciando novamente a incapacidade da rede em realizar a predição adequada. Nesse caso, observa-se a ocorrência de overfitting, indicando que a rede na base 6 também não conseguiu generalizar os padrões espectrais dessa composição vítrea.

APÊNDICE D – RESULTADOS DA PREDIÇÃO EXPERIMENTAL PELAS REDES COM NORMALIZAÇÃO E 2% DE RUÍDO

- 1C-1Ag base 5 – 1C-1Ag base 6 com um 2% de ruído

Para a rede 1C-1Ag com 2% de ruído, os valores de erro obtidos foram 0,0000645501 para validação e 0,0000506268 para treinamento. Não houve ocorrência de *overfitting*, porém observa-se que esses valores são superiores aos obtidos para a mesma rede com 1% de ruído. Esse aumento indica que a rede aprendeu um pouco menos, o que é esperado, considerando que o acréscimo de ruído aos resultados teóricos torna o aprendizado mais real. Na Figura 55, apresentam-se os erros das amostras TLZCr075 e TLZCr2, considerando os parâmetros de centro, largura e área das gaussianas.

Figura 55 – Resultados da predição 1C-1Ag na base 5.



Fonte: O autor.

Tabela 42 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0,75 e TLZCr-2 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-075	Cr-2	Cr-075	Cr-2	Cr-075	Cr-2
Gaussiana 1	2,234	0,173	2,821	1,300	1,713	1,732
Gaussiana 2	0,514	0,356	12,346	7,681	9,069	8,698
Gaussiana 3	0,656	0,891	10,557	12,984	1,012	19,098
Gaussiana 4	0,988	1,125	11,846	1,347	17,982	5,309
Gaussiana 5	0,474	0,458	1,452	0,930	14,906	9,919

Fonte: O autor.

Os resultados da predição da espectroscopia Raman para a amostra TLZCr075 apresentaram erro máximo de 12,346% na largura da Gaussiana 4 e de 14,906% na área da Gaussiana 5. Esses valores indicam que, para alguns picos, a rede não conseguiu realizar uma predição precisa da amostra TLZCr075. Para a amostra TLZCr2, o maior erro observado foi de 12,984% na largura da Gaussiana 3 e de 19,098% na área da Gaussiana 5. Apesar do erro mais elevado em determinados parâmetros, a predição para essa

amostra foi satisfatória, mostrando que a rede conseguiu representar de forma adequada os resultados espectroscópicos da TLZCr2.

Na rede 1C-1Ag utilizando a base 6, os valores de erro obtidos foram 0,0000555533 para validação e 0,0000505852 para treinamento. Observa-se uma tendência contínua de diminuição dos erros ao longo do treinamento, indicando que a rede está aprendendo de maneira adequada e que não houve ocorrência de *overfitting*. Em comparação com a mesma rede treinada com menor ruído, os valores de erro aumentaram devido à adição de 2% de ruído à base de dados. Esse aumento é esperado e indica que a rede apresentou um pouco mais de dificuldade para aprender os resultados, refletindo o efeito do ruído adicional nos espectros teóricos.

Para as amostras de vidros teluretos, verificou-se que a rede conseguiu realizar a predição da TLZCr075, embora tenha ocorrido um aumento nos erros em comparação com a base 5, atingindo 11,885% na largura da Gaussiana 2 e 19,851% na área da Gaussiana 4. A amostra TLZCr2, os resultados foram semelhantes, apresentando erro de 13,931% na largura da Gaussiana 3 e 18,907% na área da Gaussiana 5.

De maneira geral, os resultados obtidos pela rede na base 6 com 2% de ruído mostraram-se mais promissores em termos de capacidade de predição do que os observados para a mesma rede na base 5, evidenciando que a rede manteve bom desempenho mesmo com o aumento do ruído nos espectros teóricos.

- 1C-2Ag base 5 – 1C-2Ag base 6 com um 2% de ruído

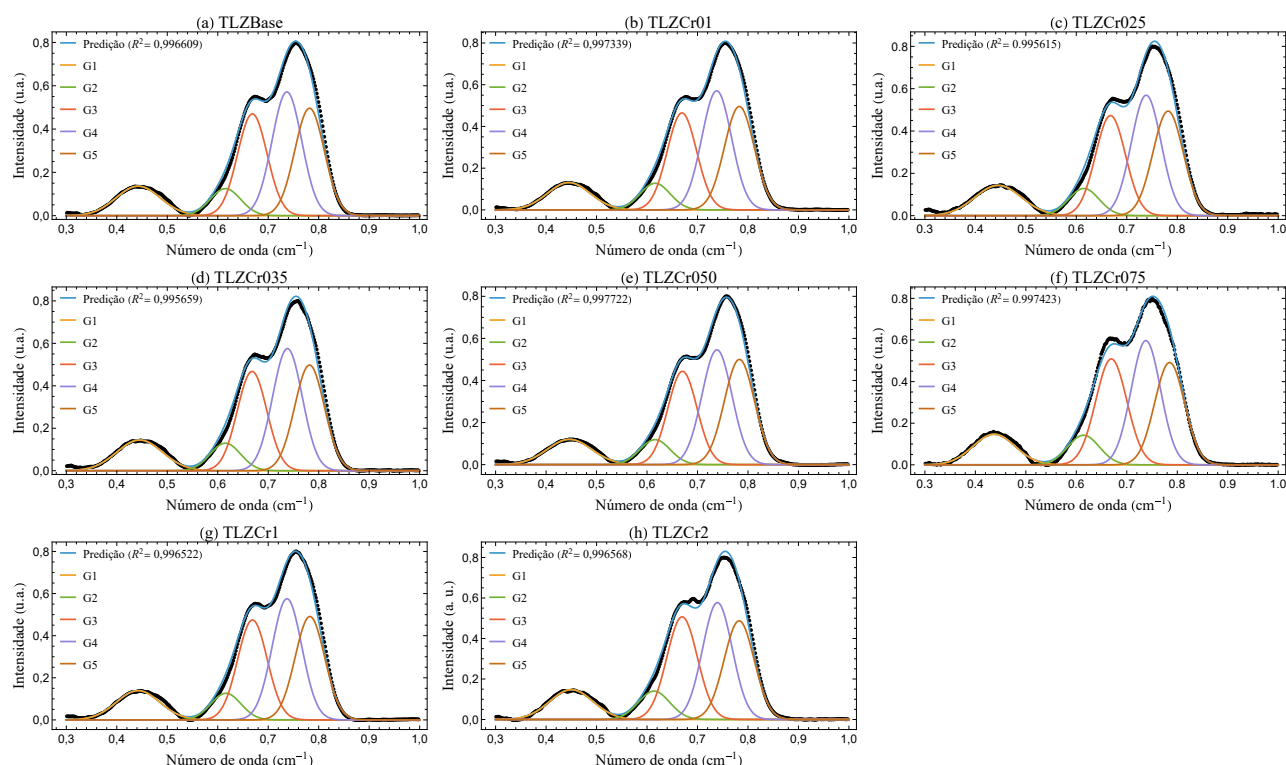
Na rede 1C-2Ag com adição de 2% de ruído, os erros registrados foram de 0,0000604551 para validação e 0,0000542242 para treinamento. Não foi identificada a ocorrência de *overfitting*, porém nota-se que esses valores são mais elevados em comparação aos obtidos para a mesma rede com 1% de ruído. Esse acréscimo sugere uma ligeira redução na capacidade de aprendizado, o que é esperado, já que a introdução de ruído nos resultados teóricos aproxima a rede de condições mais realistas de análise. Na sequência, são apresentados os erros das amostras TLZCr050, que apresentou o melhor ajuste, e TLZCr2, cuja predição pela rede apresentou um erro um pouco maior, considerando os parâmetros de centro, largura e área das funções gaussianas.

A Tabela 43, são apresentados os resultados da predição de todas as amostras de espectroscopia Raman obtidos pela rede 1C-3Ag utilizando a base 5.

Assim, observa-se que a rede 1C-2Ag na base 5 se mostra uma das mais promissoras, uma vez que foi capaz de realizar a predição da espectroscopia Raman para todas as amostras de vidros teluretos. Para a amostra TLZCr050, os maiores desvios encontrados foram de 13,129% na largura da Gaussiana 3 e de 20,698% na área da Gaussiana 5.

No caso da amostra TLZCr2, o erro máximo identificado foi de 13,113% na largura da Gaussiana 3 e de 18,0809% na área da Gaussiana 3. Apesar do R^2 ligeiramente inferior,

Figura 56 – Resultados da predição 1C-2Ag na base 5.



Fonte: O autor.

Tabela 43 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0,50 e TLZCr-2 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		Área	
	Cr-050	Cr-2	Cr-050	Cr-2	Cr-050	Cr-2
Gaussiana 1	0,283	0,020	1,939	2,402	0,270	0,668
Gaussiana 2	0,461	0,180	9,880	10,333	0,172	15,434
Gaussiana 3	0,436	0,228	12,657	13,113	0,640	18,809
Gaussiana 4	0,894	1,023	0,938	1,571	0,815	6,681
Gaussiana 5	0,472	0,059	0,193	0,529	20,240	10,369

Fonte: O autor.

conforme apresentado na Figura 56, e dos erros mais elevados em alguns parâmetros, a predição para essas amostras pode ser considerada satisfatória, evidenciando que a rede conseguiu representar de forma consistente os resultados espectroscópicos das amostras TLZCr050 e TLZCr2.

Para a mesma rede, mas considerando agora a base 6, os valores de erro na validação 0,0000555118 e no treinamento 0,0000505822 apresentaram redução em relação à base 5. Esses resultados indicam que não houve ocorrência de *overfitting*, demonstrando que a rede foi capaz de realizar de forma adequada a predição da espectroscopia Raman.

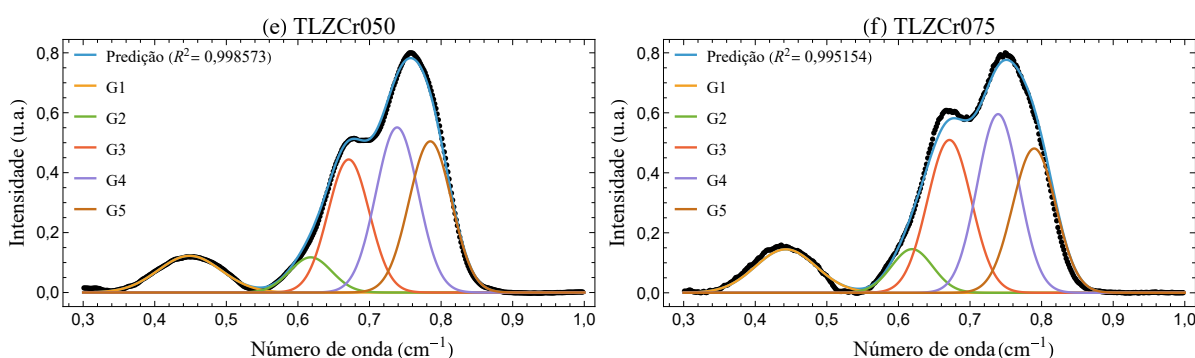
Assim, as amostras que apresentaram, respectivamente, os menores e maiores erros foram

TLZCr035 e TLZCr075. A amostra TLZCr035 obteve o melhor desempenho, com erro máximo de 13,065% na largura da Gaussiana 3 e de 15,464% na área da Gaussiana 5. Por outro lado, a amostra TLZCr075 apresentou a pior predição, com erro de 15,519% na largura da Gaussiana 2 e de 18,988% na área da Gaussiana 4.

- 1C-3Ag base 5 – 1C-3Ag base 6 com um 2% de ruído

A rede 1C-3Ag também apresentou comportamento semelhante ao observado nas demais, com valores de erro de validação de 0,0000562354 e de treinamento de 0,0000504321. Observa-se que esses valores aumentaram em função da adição de 2% de ruído à base de dados. Ainda assim, os resultados obtidos por essa rede, assim como pelas anteriores, mantêm-se bastante satisfatórios, sem indícios de *overfitting*. O melhor desempenho dessa rede foi observado para a amostra TLZCr050, enquanto a predição menos satisfatória ocorreu para a amostra TLZCr075, conforme apresentado na Tabela 44.

Figura 57 – Resultados da predição 1C-3Ag na base 5.



Fonte: O autor.

Então temos na Figura 57 os resultados das amostras da espectroscopia raman da rede 1C-3Ag na base 5.

Tabela 44 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZCr-0,50 e TLZCr-0,75 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		área	
	Cr-050	Cr-0,75	Cr-050	Cr-0,75	Cr-050	Cr-0,75
Gaussiana 1	0,804	0,538	1,939	3,543	8,973	3,072
Gaussiana 2	0,156	0,795	9,880	6,466	7,737	9,492
Gaussiana 3	0,329	0,467	12,657	10,660	6,516	0,492
Gaussiana 4	0,887	0,976	0,938	11,221	0,304	19,620
Gaussiana 5	0,227	1,084	0,193	0,370	20,698	14,935

Fonte: O autor.

Dessa forma, verifica-se que a rede 1C-3Ag, na base 5 foi capaz de realizar a predição dos espectros Raman de todas as amostras de vidros teluretos. Para a amostra TLZCr050,

os maiores desvios observados corresponderam a 12,657% na largura da Gaussiana 3 e 20,240% na área da Gaussiana 5.

Para a amostra TLZCr075, identificou-se um erro máximo de 11,221% na largura da Gaussiana 4 e 19,620% na área da mesma gaussiana. Embora o valor de R^2 tenha se mostrado ligeiramente inferior, conforme ilustrado na Figura 58, e alguns parâmetros tenham apresentado desvios mais expressivos, a predição para essas amostras ainda pode ser considerada satisfatória. Isso demonstra que a rede foi capaz de representar de maneira consistente os resultados espectroscópicos tanto da amostra TLZCr050 quanto da TLZCr075.

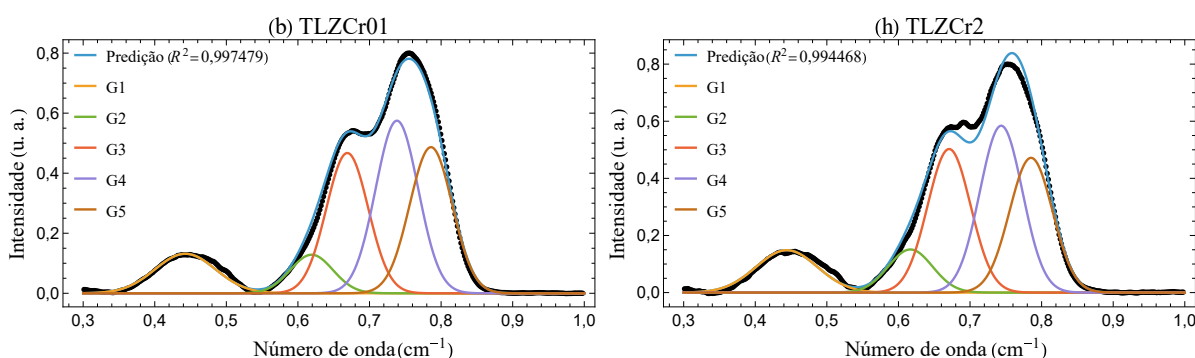
Considerando a mesma rede, mas agora aplicada à base 6, foram obtidos valores de erro de 0,0000556294 para a validação e 0,0000498928 para o treinamento, indicando uma redução em comparação aos resultados obtidos na base 5. Esses valores confirmam que não houve ocorrência de *overfitting*, demonstrando que a rede manteve um bom desempenho na predição da espectroscopia Raman.

No conjunto das amostras, a que apresentou o melhor resultado foi a TLZCr025, com erros máximos de 13,709% na largura da Gaussiana 3 e 15,436% na área da Gaussiana 5. A amostra TLZCr1 apresentou a pior predição, registrando desvios de 14,002% na largura da Gaussiana 3 e 14,566% na área da Gaussiana 2.

- 1C-4Ag base 5 – 1C-4Ag base 6 com um 2% de ruído

A rede apresenta o menor número de neurônios na camada de entrada e obteve os maiores valores de erro de validação 0,0000641706 e de treinamento 0,0000565719. Ainda assim, os resultados podem ser considerados promissores, uma vez que não foram identificados indícios de *overfitting*. O melhor desempenho dessa rede foi observado para a amostra TLZCr01, enquanto a predição menos satisfatória ocorreu para a amostra TLZCr2, conforme apresentado na Tabela 45.

Figura 58 – Resultados da predição 1C-4Ag na base 5.



Fonte: O autor.

Então temos os resultados das amostras na Figura 58 da espectroscopia raman da rede 1C-4Ag na base 5.

Tabela 45 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0,1 e TLZCr-2 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		área	
	Cr-01	Cr-2	Cr-01	Cr-2	Cr-01	Cr-2
Gaussiana 1	0,754	1,346	1,599	2,644	7,739	1,193
Gaussiana 2	0,670	0,276	13,624	9,479	9,944	23,432
Gaussiana 3	0,148	0,420	13,921	12,219	6,293	16,887
Gaussiana 4	0,863	1,575	16,273	0,504	4,205	9,109
Gaussiana 5	0,495	0,415	1,869	0,121	19,911	12,629

Tabela 46 – Parâmetros das gaussianas ajustadas para as redes 1C-4Ag, das amostras TLZCr-0,1 e TLZCr-2 incluindo centro, largura e área de cada banda.

Fonte: O autor.

Assim, observa-se que a rede 1C-4Ag, aplicada à base 5, conseguiu prever os espectros Raman de todas as amostras de vidros teluretos. Para a amostra TLZCr01, os maiores desvios encontrados foram de 16,273% na largura da Gaussiana 4 e de 19,911% na área da Gaussiana 5.

Para a amostra TLZCr2, o maior desvio encontrado foi de 12,219% na largura da Gaussiana 3 e 23,432% na área da Gaussiana 2. Ainda assim, o resultado pode ser considerado satisfatório, uma vez que os valores de erro estão dentro do esperado.

A rede aplicada na base 6, teve os erros de 0,0000556294 na validação e 0,0000498928 no treinamento, valores inferiores aos observados para a base 5. Esses resultados confirmam a ausência de *overfitting* e indicam que a rede manteve um desempenho consistente na predição dos espectros Raman.

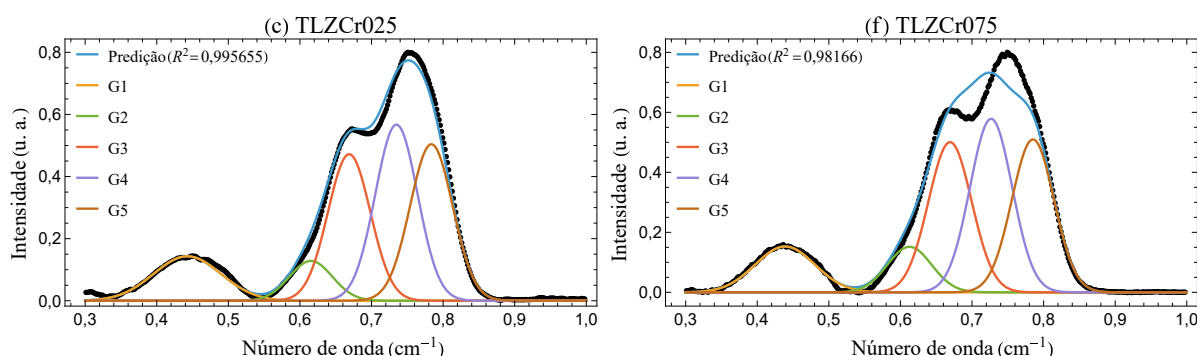
Entre as amostras analisadas, a TLZCr025 apresentou o melhor ajuste, com desvios máximos de 13,709% na largura da Gaussiana 3 e 15,436% na área da Gaussiana 5. Em contrapartida, a TLZCr1 exibiu a predição menos satisfatória, registrando erros de 14,002% na largura da Gaussiana 3 e 14,566% na área da Gaussiana 2.

- 2C-3Ag base 5 – 2C-3Ag base 6 com um 2% de ruído

A rede 2C-3Ag, única composta por duas camadas de entrada, apresentou valores de erro de 0,0000556313 na validação e 0,0000517037 no treinamento. Esses resultados são muito próximos aos obtidos pelas redes com apenas uma camada, podendo ser considerados promissores, uma vez que não foram identificados indícios de *overfitting* nos valores globais de erro. No entanto, ao analisar os resultados de predição individuais, observa-se que para algumas amostras de vidros teluretos dopados com Cr a rede não conseguiu reproduzir adequadamente os espectros, sugerindo comportamento de *overfitting* em casos específicos, conforme apresentado na Tabela 47.

Então temos os na Figura 59 resultados da predição de todas as amostras da espectroscopia raman da rede 2C-3Ag na base 5.

Figura 59 – Resultados da predição 1C-11Ag na base 5.



Fonte: O autor.

Tabela 47 – Parâmetros das gaussianas ajustadas para as redes 1C-1Ag, das amostras TLZBase e TLZCr-2 incluindo centro, largura e área de cada banda.

Gaussiana	Centro		Largura		área	
	Cr-025	Cr-075	Cr-025	Cr-075	Cr-025	Cr-075
Gaussiana 1	0,756	1,373	2,165	1,464	1,216	0,117
Gaussiana 2	0,060	0,228	12,548	13,984	16,506	22,232
Gaussiana 3	0,120	0,179	14,667	8,809	15,563	2,892
Gaussiana 4	0,387	0,735	4,035	12,319	12,273	17,296
Gaussiana 5	0,211	0,501	1,333	2,376	8,899	11,698

Fonte: O autor.

Portanto, a rede 2C-3Ag com 2% de ruído apresenta comportamento semelhante ao observado na mesma rede com 1% de ruído, indicando que o modelo assimilou de forma excessiva as particularidades do conjunto de treinamento, comprometendo sua capacidade de generalização. Essa limitação pode estar associada à elevada complexidade da rede em relação ao volume e à variabilidade dos dados disponíveis, favorecendo o sobreajuste. Além disso, a baixa diversidade estrutural das amostras também pode ter contribuído para esse efeito, restringindo a capacidade do modelo em representar adequadamente novas instâncias do sistema vítreo analisado.