



Universidade Estadual Paulista “Júlio De Mesquita Filho”  
INSTITUTO DE QUÍMICA – ARARAQUARA  
ENGENHARIA QUÍMICA



**TAMIRES MESSIAS BERTO**

**DESENVOLVIMENTO DE MODELOS NEURAIIS PARA  
CLASSIFICAÇÃO DE AMOSTRAS DE CABELO DE ACORDO COM A  
PIGMENTAÇÃO, GÊNERO, FAIXA ETÁRIA E TRATAMENTO  
QUÍMICO**

**Araraquara, SP**

**2020**

**TAMIRES MESSIAS BERTO**

**DESENVOLVIMENTO DE MODELOS NEURAIIS PARA  
CLASSIFICAÇÃO DE AMOSTRAS DE CABELO DE ACORDO COM A  
PIGMENTAÇÃO, GÊNERO, FAIXA ETÁRIA E TRATAMENTO  
QUÍMICO**

Trabalho de Conclusão de Curso apresentado ao Instituto de Química da Universidade Estadual Paulista “Júlio de Mesquita Filho” como requisito parcial para a obtenção do título de Bacharel em Engenharia Química.

**Orientadora:** Profa. Dra. Érica Regina Filletti Nascimento

**Araraquara, SP**

**2020**

TAMIRES MESSIAS BERTO

**DESENVOLVIMENTO DE MODELOS NEURAIIS PARA  
CLASSIFICAÇÃO DE AMOSTRAS DE CABELO DE ACORDO COM A  
PIGMENTAÇÃO, GÊNERO, FAIXA ETÁRIA E TRATAMENTO  
QUÍMICO**

Trabalho de Conclusão de Curso apresentado  
ao Instituto de Química da Universidade  
Estadual Paulista “Júlio de Mesquita Filho”  
como requisito parcial para a obtenção do título  
de Bacharel em Engenharia Química.

Aprovado em: \_\_\_/\_\_\_/\_\_\_

BANCA EXAMINADORA

---

Profa. Dra. Érica Regina Filletti Nascimento  
Instituto de Química – UNESP, Araraquara.

---

Profa. Dra. Fabíola Manhas Verbi Pereira  
Instituto de Química – UNESP, Araraquara.

---

Prof. Dr. Gustavo Nakamura Alves Vieira  
Instituto de Química – UNESP, Araraquara.

## **DEDICATÓRIA**

Primeiramente a Deus, razão da minha existência.

À minha família, por todo amor, carinho e incentivo em todos os momentos da minha vida.

À minha professora orientadora Érica, pelo ensinamento, paciência, confiança e dedicação durante a supervisão das minhas atividades.

## **AGRADECIMENTOS**

Agradeço primeiramente a Deus por todas as bênçãos para que meus objetivos fossem alcançados.

Aos meus pais, pelo incentivo, amor, compreensão, investimento e confiança, e a minha irmã por todo apoio e amizade.

A todos os professores que me acompanharam durante a graduação que foram cruciais para a minha constituição profissional e pessoal. Em especial à minha orientadora e professora Érica Regina Filletti Nascimento, pela convivência agradável durante a execução do trabalho, por todo ensinamento, amizade e dedicação. Também aos professores Fabíola Manhas Verbi Pereira e Gustavo Nakamura Alves Vieira pela compreensão, carinho e conselhos.

Aos meus colegas e amigos, que direta ou indiretamente, contribuíram para a realização desse trabalho, pelo companheirismo e troca de experiências.

## RESUMO

As Redes Neurais Artificiais (RNAs) são funções matemáticas em que são utilizadas técnicas computacionais para obter informações a partir destas funções. Devido à sua facilidade de implementação e de aprendizado, utilizaram RNAs para classificar amostras de cabelo humano, pois possuem grande importância para diagnóstico clínico, análise nutricional e perícia, além de vantagens para estocar, transportar e manusear tais amostras. As classificações dos fios de cabelo foram feitas em função da pigmentação (cabelo original e descolorido), faixa etária (infantil, adulto e idoso), gênero (feminino e masculino) e tratamento químico (virgem, descolorido, alisamento, tintura e tintura e alisamento), a partir de dados numéricos de 127 amostras obtidos por técnicas analíticas não invasivas, sendo elas a fluorescência de raios-x por dispersão de comprimento de onda (WDXRF – *Wavelength Dispersive x-ray Fluorescence*) e por dados de espectroscopia de quebra induzida por laser (LIBS). As redes neurais foram implementadas utilizando a ferramenta *nstart* do MATLAB R2018a, com o algoritmo de retropropagação de gradiente conjugado escalonado (*traincsg*), sendo definido o número de camadas e os seus respectivos neurônios por tentativa e erro para se obter o melhor resultado, além de parada precoce do treinamento para evitar mau desempenho de generalização. Para todas as classificações foi definido que 70% das amostras totais seriam para o conjunto de treinamento, 15% para o conjunto de validação e 15% para o conjunto de teste. Os resultados se mostraram promissores, evidenciando que as RNAs desenvolvidas são capazes de classificar as amostras de cabelo pelos parâmetros supracitados, principalmente a de pigmentação com a base de dados pelo WDXRF com 20 neurônios na camada intermediária e pelos dados do LIBS com 10 neurônios na camada intermediária, ambas com 100% de acerto em todos os conjuntos. A não classificação de algumas classes, como faixa etária e tratamento químico, pode ser devido a pequena quantidade de amostras em relação às demais e à complexidade dos dados de fio de cabelo.

**Palavras-chave:** Redes Neurais Artificiais. Amostras de cabelo. Classificação. Algoritmo de aprendizagem.

## **ABSTRACT**

Artificial Neural Networks (ANNs) are mathematical functions in which computational techniques are used to obtain information from these functions. Due to the easiness of implementation and learning, ANNs were used for the classification of human hair samples. Human hair, as they have great importance for clinical diagnosis, nutritional analysis and forensic analysis, as well as advantages for stocking, transporting and handling such material. Classifications of hair strands were made according to pigmentation (original and bleached hair), age group (infant, adult and elderly), gender (female and male) and chemical treatment (virgin, bleached, straightening, dyeing and dyeing and straightening), based on 127 samples, chosen by non-invasive analytical techniques, namely wavelength dispersion x-ray fluorescence (WDXRF) and laser-induced breakdown spectroscopy (LIBS). The neural networks were implemented using a MATLAB R2018a nnstart tool, with the staggered conjugated gradient retropropagation algorithm (traincsg), being defined the number of layers and their respective neurons per attempt and to obtain the best result, in addition to early stop training to avoid poor generalization performance. For all classifications defined for 70% of the complete definitions for the training set, 15% for the validation set and 15% for the test set. The results are promising, showing that the developed RNAs are capable of classifying hair samples by the aforementioned parameters, mainly pigmentation based on data by WDXRF with 20 neurons in the middle layer and by LIBS data with 10 neurons in the middle layer, both with 100% accuracy in all sets. The failure to classify some classes, such as age group and chemical treatment, may be due to the small number of samples compared to the others and the complexity of the hair data.

**Key-words:** Artificial Neural Networks. Hair samples. Classification. Learning algorithm.

## LISTA DE FIGURAS

Figura 1 - Esquema dos constituintes da célula neural.....	12
Figura 2 - Representação da sinapse entre os neurônios. ....	13
Figura 3 - Neurônio artificial genérico .....	13
Figura 4 - Estrutura da ferramenta <i>nnstart</i> para selecionar variáveis de entrada (Inputs) e saída (Targets). ....	20
Figura 5 - Exemplo de matriz de confusão.....	20
Figura 6 - (a) Matrizes de confusão e (b) Histograma de erros para pigmentação por WDXRF .....	22
Figura 7 - (a) Matrizes de confusão e (b) Histograma de erros para gênero por WDXRF .....	23
Figura 8 - (a) Matrizes de confusão e (b) - Histograma de erros para faixa etária por WDXRF .....	26
Figura 9 - (a) Matrizes de confusão e (b) Histograma de erros para tratamento químico por WDXRF .....	29
Figura 10 - (a) Matrizes de confusão e (b) Histograma de erros para pigmentação por LIBS.	32
Figura 11 - (a) Matrizes de confusão e (b) Histograma de erros para classificação de gênero por LIBS .....	33
Figura 12 - (a) Matrizes de confusão e (b) Histograma de erros para faixa etária por LIBS ...	36
Figura 13 - (a) Matrizes de confusão e (b) Histograma de erros para tratamento químico por LIBS .....	39

## LISTA DE TABELAS

Tabela 1 - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de teste do WDXRF .....	24
Tabela 2 - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de validação do WDXRF .....	25
Tabela 3 - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de validação do WDXRF .....	27
Tabela 4 - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de teste do WDXRF .....	28
Tabela 5 - Respostas obtidas e esperadas pela RNA para o tratamento químico no conjunto de validação do WDXRF .....	30
Tabela 6 - Respostas obtidas e esperadas pela RNA para o tratamento químico no conjunto de teste do WDXRF .....	31
Tabela 7 - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de teste do LIBS .....	34
Tabela 8 - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de validação do LIBS .....	35
Tabela 9 - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de teste do LIBS .....	37
Tabela 10 - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de validação do LIBS .....	38
Tabela 11 - Resultados obtidos e esperados pela RNA para tratamento químico no conjunto de teste do LIBS .....	40
Tabela 12 - Resultados obtidos e esperados pela RNA para tratamento químico no conjunto de validação no LIBS .....	41

## SUMÁRIO

<b>1. INTRODUÇÃO</b> .....	<b>11</b>
1.1 Neurônios Biológicos e as Redes Neurais Artificiais .....	11
1.2 Histórico das Redes Neurais Artificiais .....	14
1.3 Pesquisas recentes envolvendo Redes Neurais Artificiais.....	15
<b>2. OBJETIVOS</b> .....	<b>17</b>
2.1 Objetivo Geral.....	17
2.2 Objetivos Específicos .....	17
<b>3. MATERIAIS E MÉTODOS</b> .....	<b>18</b>
3.1 Amostras .....	18
3.2 Modelo neural com a ferramenta <i>nnstart</i> .....	18
<b>4. RESULTADOS E DISCUSSÃO</b> .....	<b>21</b>
4.1 Classificação das amostras de cabelo com dados numéricos por WDXRF, via RNAs 21	
4.1.1 Pigmentação .....	21
4.1.2 Gênero .....	22
4.1.3 Faixa etária .....	25
4.1.4 Tratamento químico.....	28
4.2 Classificação das amostras de cabelo com dados numéricos por LIBS, via RNAs...31	
4.2.1 Pigmentação .....	31
4.2.2 Gênero .....	32
4.2.3 Faixa etária .....	35
4.2.4 Tratamento químico.....	38
<b>5. CONCLUSÃO</b> .....	<b>42</b>
<b>6. REFERÊNCIAS</b> .....	<b>43</b>

## 1. INTRODUÇÃO

RNAs são funções matemáticas. Para fins práticos, utilizam-se técnicas computacionais para obter informações a partir destas funções. Elas são compostas por unidades de processamento simples, os neurônios artificiais, que estão dispostos em uma ou mais camadas interligadas por conexões associadas a pesos, que armazenam o conhecimento representado no modelo e servem para ponderar a entrada recebida por cada neurônio [1]. Inicialmente, a RNA passa por uma fase de aprendizagem em que um conjunto de exemplos é apresentado a ela, na qual se extraem as características necessárias para representar a informação fornecida. Em seguida, essas características são utilizadas para gerar respostas para o problema para dados não vistos no treinamento, essa é a fase de generalização.

O cabelo é uma matriz biológica que tem atraído a atenção de vários pesquisadores, visando desenvolver novas metodologias analíticas para adquirir conhecimentos aprofundados dos dados para diferentes propostas, tais como perícia, diagnóstico clínico e nível nutricional [7-9]. Sua alta estabilidade contra degradação natural, assim como o armazenamento, transporte e manuseio mais fáceis, permitem uma longa janela de detectabilidade para analitos de interesse do que outras amostras biológicas, principalmente em análises de cunho forense [8]. Recentemente, Santos et al. [2] aplicaram a técnica de fluorescência de raios-X dispersiva de comprimento de onda (WDXRF) combinada com ferramentas quimiométricas para avaliação de amostras de cabelo humano e sintético, antes e depois do procedimento cosmético, e elementos químicos, como Ca, Cu, Fe, K e S, foram decisivos para agrupar as amostras na Análise de Componentes Principais (PCA).

Portanto, devido a facilidade de implementação das RNAs, às respostas precisas em alta velocidade e à sua capacidade de generalização, optou-se por utilizá-las com base no algoritmo de retropropagação de gradiente conjugado escalonado [10] para classificar amostras de cabelo humano em função do tratamento químico (virgem, descolorido, alisamento, tintura e alisamento e tintura), da pigmentação (cabelo original e descolorido), da faixa etária (infantil, adulto e idoso) e do gênero (feminino e masculino). Os dados numéricos referentes à essas amostras foram obtidos por técnicas analíticas não invasivas, a WDXRF e dados de espectroscopia de quebra induzida por laser (LIBS), cujo detalhamento podem ser encontrados em Santos et al. [18].

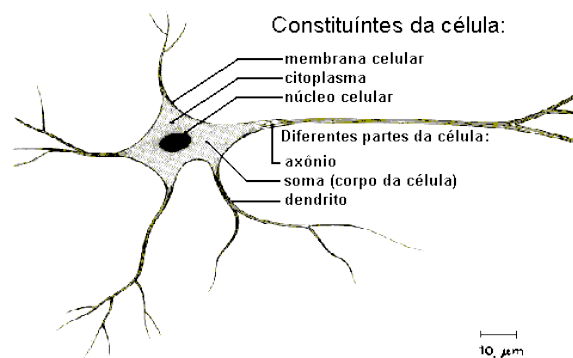
### 1.1 Neurônios Biológicos e as Redes Neurais Artificiais

O cérebro humano possui  $10^{11}$  neurônios, e cada um processa e se comunica com

milhares de outros continuamente e em paralelo, e é responsável pela emoção, percepção, pensamento, e funções sensoriomotoras e autônomas. O funcionamento das redes biológicas ainda não foi totalmente desvendado pelo homem, o que se tem são modelos, atualizados a cada nova descoberta. Entretanto, a estrutura fisiológica básica destas redes de nodos naturais é conhecida, e é nela que se baseiam as RNAs [1].

Os neurônios biológicos têm três seções: corpo da célula, dendritos e axônio. Os dendritos têm função de receber as informações, impulsos nervosos, oriundas de outros neurônios e conduzi-las até o corpo celular. O corpo da célula é responsável por coletar e combinar informações vindas de outros neurônios. E o axônio, constituído de uma fibra, é responsável por transmitir os estímulos para outras células [11]. A estrutura de uma célula neural está representada na Figura 1.

**Figura 1** - Esquema dos constituintes da célula neural.

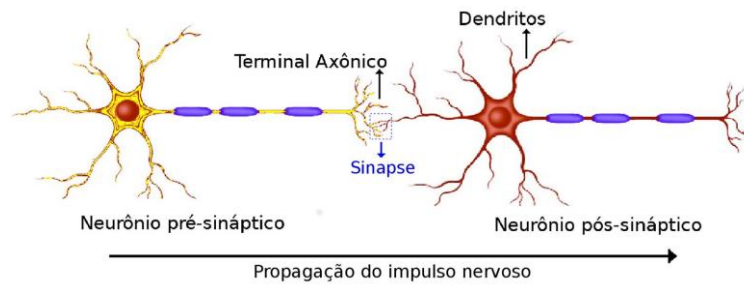


**FONTE:** CORRÊA [11].

A comunicação entre os neurônios é realizada através de impulso. Quando um impulso é recebido, o neurônio faz o processamento e, se atingir um limite de ação, dispara um segundo impulso que produz uma substância neurotransmissora que flui do corpo celular para o axônio. O neurônio que transmite o pulso pode controlar a frequência de pulsos aumentando ou diminuindo a polaridade na membrana pós-sináptica. Sendo assim, as redes neurais naturais não transmitem sinais negativos, sua ativação é medida pela frequência com que emite pulsos, que são contínuos e positivos [11].

Quando ocorre o contato entre a terminação axônica de um neurônio e o dendrito do outro se tem uma sinapse, como ilustra a Figura 2. É a partir dela que os nodos se unem funcionalmente formando redes neurais. Elas funcionam como válvulas e são capazes de controlar a transmissão de impulsos e o fluxo de informação entre os nodos na rede neural. Os efeitos das sinapses são variáveis, e é esta variação que dá ao neurônio a capacidade de adaptação [1].

**Figura 2** - Representação da sinapse entre os neurônios.



**FONTE:** BORGES [16].

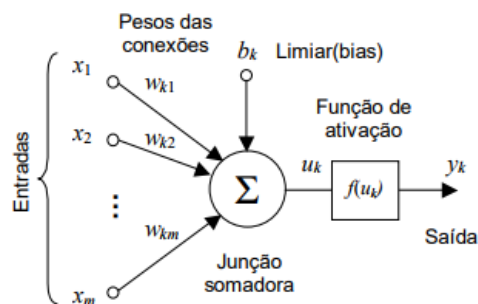
As RNAs tentam reproduzir as funções das redes biológicas, buscando programar seu comportamento básico e sua dinâmica. Dentre as características comuns têm-se: os dois sistemas são baseados em unidade de computação paralela e distribuídos que se comunicam por conexões sinápticas, possuem detectores de características, redundância e modularização das conexões. Isso permite que as RNAs reproduzam com fidelidade várias funções somente encontradas nos seres humanos [1].

Uma rede neural artificial pode ser projetada através da definição de [3]:

1. Um conjunto de neurônios artificiais;
2. Um padrão de conectividade entre os neurônios, ou seja, uma arquitetura para a rede;
3. Um algoritmo de aprendizagem ou treinamento para determinação dos parâmetros livres da rede (pesos e *bias*).

O elemento computacional básico empregado nas RNAs é o integrador, um elemento processador de informações fundamental para as operações. As principais partes do neurônio artificial genérico são: sinapses, caracterizadas pelos seus pesos associados; a junção somadora e a função de ativação [3]. O peso sináptico  $w_{kj}$  tem o primeiro subscrito correspondente ao neurônio pós-sináptico, e o segundo subscrito à sinapse ligada a ele, como ilustra a Figura 3.

**Figura 3** - Neurônio artificial genérico



**FONTE:** CASTRO e VON ZUBEN [3].

A junção somadora soma todos os sinais de entrada ponderados pelos pesos das conexões. Assume vetores de entrada e de pesos como sendo vetores coluna, que corresponde

ao produto interno do vetor de entradas  $x$  pelo vetor de pesos  $w_k$ , mais o *bias*  $b_k$ . Trata-se de uma combinação linear das entradas pelos pesos associados, mais o *bias*  $b_k$ . A função de ativação é geralmente utilizada com dois propósitos: limitar a saída do neurônio e introduzir não-linearidade no modelo [3].

Existem basicamente três camadas em uma rede neural artificial: camada de entrada, camada intermediária e camada de saída, porém nem todas redes neurais possuem camadas intermediárias. A forma pela qual os neurônios estão interconectados está intrinsecamente relacionada ao algoritmo a ser utilizado no treinamento [3].

## 1.2 Histórico das Redes Neurais Artificiais

O cientista Santiago Ramón y Cajal em 1911 identificou inicialmente que o neurônio constitui a base do sistema nervoso [17].

Warren McCulloch e Walter Pitts, em 1943, sugeriram a construção de uma máquina baseada no cérebro humano e em 1947 conseguiram demonstrar que era possível conectar os neurônios artificiais e formar uma rede capaz de executar funções complexas. O neurônio ficou conhecido como nó do Modelo de McCulloch e Pitts (MCP), porém estava restrito aos neurônios biológicos que representavam e não ao processo de aprendizagem que ocorria no cérebro [12].

Donald Hebb foi o primeiro a propor uma lei de aprendizagem específica para as sinapses dos neurônios, com o seu livro *The Organization of Behavior* de 1949 [1].

Em 1958, Rosenblatt propôs o primeiro neuro computador a obter sucesso, conhecido como *perceptron*, consistindo em neurônios organizados em camada de entrada e saída, onde os pesos das conexões eram adaptados a fim de se atingir a eficiência sináptica. O grande interesse gerado por esse neuro computador é a habilidade de generalizar seus dados do treinamento e de aprender a partir de conexões inicialmente aleatórias, apropriados para problemas simples, de classificação de padrões, consideradas redes rápidas e de confiança [12].

Em paralelo a Rosenblatt, em 1960, Widrow e Hoff criaram a primeira rede capaz de imitar o cérebro humano utilizando processadores paralelos, ao invés de um único processador, conhecida como a Rede ADALINE (*Adaptive Linear Element*), com diferença entre o *perceptron* devido ao procedimento de aprendizagem. Em seguida, foi desenvolvida a rede MADALINE (Multiple-Adaline) que é um modelo multidimensional da rede ADALINE [12].

Após a publicação em 1969 do livro “*Perceptrons*”, de Minsky e Papert, as pesquisas na área de redes neurais sofreram uma retração significativa devido a limitação apresentada pelo modelo. Assim, as pesquisas ficaram estagnadas por toda década de 70 e início da década

de 80, restando poucos pesquisadores ativos [12].

Em 1982, Teuvo Kohonen apresentou um modelo de rede neural denominado Mapas Auto-Organizáveis que se baseou na neurofisiologia do córtex cerebral.

No final da década de 80, os pesquisadores Rumelhart, Hinton e Williams, desenvolveram o algoritmo de retropropagação do erro (*error back-propagation*). Nessas redes, cada camada tem uma função específica, operando em dois passos, onde no primeiro passo as informações são apresentadas à camada de entrada da rede e a atividade flui camada por camada até a resposta ser gerada na camada de saída. A saída obtida é comparada com a desejada para esse padrão particular, se não estiver correta, o erro é calculado, a partir da camada de saída até a entrada de saída, e os pesos das conexões das unidades das camadas internas vão sendo modificados conforme o erro é retro propagado. Esse algoritmo é visto como uma resposta às críticas feitas por Minsky e Papert nos anos 60 [1].

### 1.3 Pesquisas recentes envolvendo Redes Neurais Artificiais

Apesar da capacidade das redes neurais de aprender com um conjunto de dados e ainda fornecer respostas coerentes a dados desconhecidos, um pequeno número de relatórios é encontrado na literatura combinando RNA a técnicas analíticas para prever uma tendência ou classificar os tecidos capilares humanos [13-14].

Noorizadeh et al. [13] usaram espectrometria de massa de tempo de voo por cromatografia líquida de ultra-desempenho (UPLC-TOF/MS) para analisar 52 fármacos e drogas de abuso no cabelo. Neste estudo, os estudos de mínimos quadrados parciais (PLS) e modelos de RNAs foram aplicados para a previsão dos valores do tempo de retenção (RT) dos analitos em função do tipo de fio de cabelo. Como resultado, o modelo de Redes Neurais Artificiais mostrou treinamento de alto coeficiente de correlação ( $R^2$ ), bem como baixo erro de predição do que o PLS. Em relação ao modelo PLS:  $R^2$ , erro relativo médio (RE) e os conjuntos de treinamento e teste foram (0,897, 0,780), (9,61, 17,36) e (0,56, 0,80), respectivamente. Para a RNA, o  $R^2$ , RE médio e raiz quadrada do erro médio (RMSE) para calibração, predição e conjuntos de teste foram (0,970, 0,939, 0,922), (4,74, 6,81, 8,01) e (0,29, 0,41, 0,55), respectivamente.

No relatório de Esfandiari et al. [14] foi feita a comparação entre seis algoritmos de otimização de duas camadas (*Levenberg-Marquardt*, *Resilient Back Propagation*, *Powell-Beale Conjugate Gradient*, *Fletcher-Reeves Conjugate Gradient*, *Polak-Ribiere Conjugate Gradient*, *Scaled Conjugate Gradient*) na previsão de perda de cabelo, usando RNAs. Oito

atributos (sexo, idade, fatores genéticos, cirurgia, gravidez, deficiência de zinco e ferro, anemia e uso de cosméticos) de 384 indivíduos foram tratados como parâmetros de entrada. Como resultado, usando o erro quadrático médio (MSE) e o erro absoluto médio (MAE) como critérios, o algoritmo de Levenberg-Marquardt exibiu os melhores resultados, isto é, erros de predição mínimos entre todos os algoritmos de RNA testados (0,0348 e 0,1393 para MSE e MAE, respectivamente).

Diversos artigos são publicados evidenciando resultados das redes neurais nas mais diferentes linhas de pesquisa, como na área agrícola de Ramos [4], onde as redes neurais artificiais foram utilizadas para classificar os frutos por meio de vetores de padrões, utilizando-se de uma rede perceptron de múltiplas camadas com o algoritmo de treinamento de retro-propagação para armazenar os vetores de padrões (diâmetro e peso) e classificar os padrões de entrada. Obteve-se resultado satisfatório com 15 épocas de treinamento, com 4 camadas de neurônios e que gerou apenas duas amostras com diferença entre o resultado obtido e desejado [4].

Na área medicinal foram utilizadas redes neurais artificiais para classificar o câncer a fim de obter diagnóstico específico com base nas assinaturas de expressão gênica. O treinamento foi feito usando os pequenos tumores redondos de células azuis (SRBCTs) como modelo. As RNAs classificaram corretamente todas as amostras e identificaram os genes mais relevantes para a classificação. Depois, foram utilizadas outras amostras que não haviam sido treinadas anteriormente e houve 100% de acerto na classificação [5].

## **2. OBJETIVOS**

### **2.1 Objetivo Geral**

O objetivo principal desse trabalho foi classificar amostras de cabelo em função do tratamento químico (virgem, descolorido, alisamento, tintura e alisamento e tintura), da faixa etária (infantil, adulto e idoso), do gênero (masculino e feminino) e da pigmentação (original e descolorido), utilizando ferramenta computacional alternativa, as Redes Neurais Artificiais.

### **2.2 Objetivos Específicos**

Os objetivos específicos desse trabalho foram:

- Implementar as redes neurais artificiais para todas classificações utilizando o MATLAB R2018a;
- Realizar treinamento de cada RNA a partir da base de dados de cabelo humano por WDXRF e LIBS e por tentativa e erro definir a quantidade de neurônios da camada intermediária, visando a menor porcentagem de erro para o conjunto de treinamento, validação e teste;
- Analisar o desempenho das redes neurais artificiais após treinamento a partir dos histogramas e matrizes de confusão.

### 3. MATERIAIS E MÉTODOS

#### 3.1 Amostras

As amostras de fio de cabelo foram coletados em salões de cabeleireiros da cidade de Araraquara/SP e também fornecidos por doação voluntária. Todas as amostras foram estocadas à temperatura ambiente em frascos plásticos inertes e descontaminados. As amostras de cabelo consideradas descoloridas tiveram modificação estética de descoloração semelhante ao de salões de cabeleireiro [2].

Ao todo haviam 127 amostras de cabelo humano que foram medidas em pastilhas utilizando um espectrômetro ARL Perform'X WDXRF da Thermo Scientific (Madison, WI, EUA), sendo essa uma técnica em que a radiação fluorescente de cada elemento químico ocorre em um determinado comprimento de onda e são empregados cristais para separar os vários comprimentos de onda por difração. As amostras também foram analisadas pelo sistema LIBS da Applied Spectra (Fremont, CA, USA), este que se baseia na espectrometria de emissão óptica e utiliza a microamostragem por ablação por laser e formação de um plasma para determinar os elementos químicos presentes na amostra. As medições foram realizadas diretamente nas pastilhas. As especificações das condições para as medições podem ser encontradas em Santos et al. [19]. Assim, elas puderam ser classificadas conforme gênero, faixa etária, pigmentação e procedimento estético.

#### 3.2 Modelo neural com a ferramenta *nnstart*

As redes neurais artificiais foram implementadas na interface *nnstart*, do MATLAB R2018a (MathWorks, Natick, MA, USA), com o algoritmo de retropropagação de gradiente conjugado em escala (*traincsg*). O número de camadas e de neurônios por camada foram ajustados por tentativa e erro para se obter o melhor resultado. A parada precoce do treinamento foi realizada para evitar sobreajuste pela RNA, o que poderia ocasionar mau desempenho de generalização.

A ferramenta *nnstart* do MATLAB fornece uma interface gráfica para implementação de redes neurais artificiais com funções para modelagem de sistemas não-lineares complexos, suportando a aprendizagem supervisionada com redes *feedforward*, base radial e dinâmicas. Assim como também suporta o aprendizado não supervisionado, com mapas auto organizáveis e camadas competitivas [15].

Dentre as ferramentas da *nnstart* tem-se a *Fitting Tool* (Ferramenta de Ajuste), a *PatternRecognition Tool* (Ferramenta de Reconhecimento de padrões), a *Clustering Tool*

(Ferramenta de Agrupamento) e a *Time Series Tool* (Ferramenta de Série Temporal), sendo que a segunda foi utilizada neste trabalho, já que classifica as entradas em um conjunto de categorias-alvo.

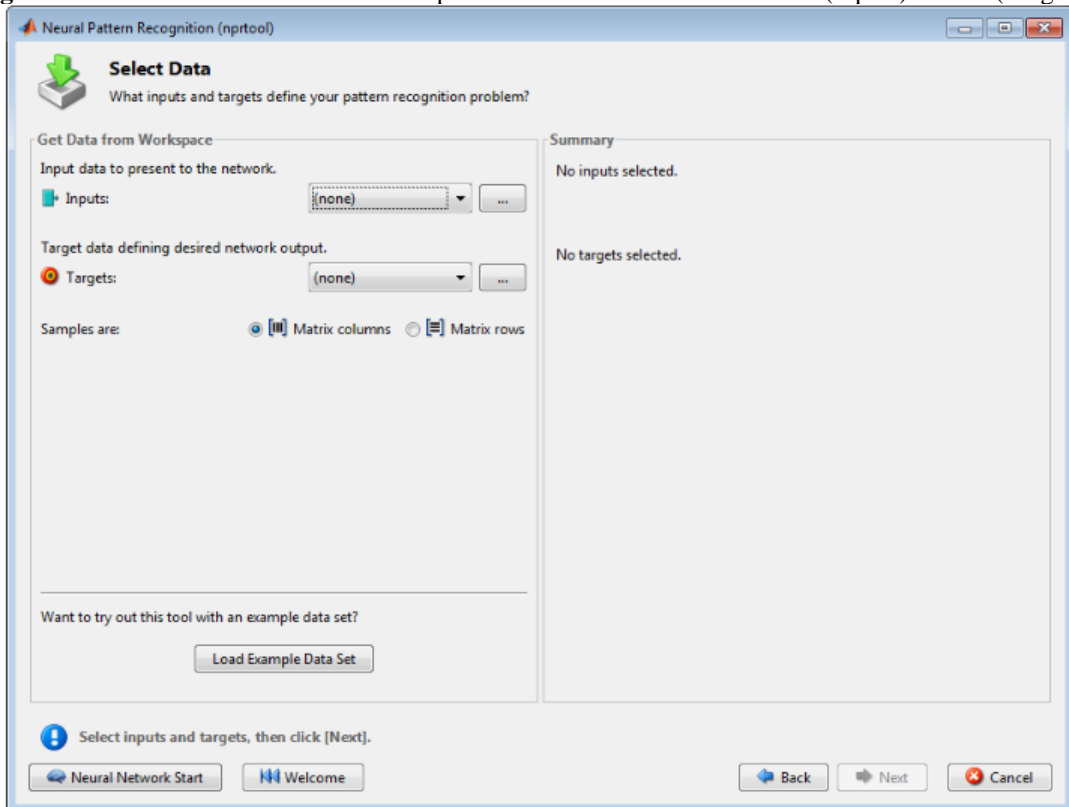
Ao selecionar a função *Pattern Recognition Tool*, foram definidas as variáveis de entrada e de saída, como ilustra a Figura 4. As de entrada foram obtidas do banco de dados das amostras de cabelo analisadas pela técnica de WDXRF, definida por uma matriz de dimensão 805x127 (805 são as variáveis referentes para todos os comprimentos de onda medidos para cada amostra e 127 o número de amostras total de cabelo), e pela técnica de LIBS, com matriz de entrada de todas as classes de dimensão 12.288x127 (12.288 são os comprimentos de onda medidos para cada amostra e 127 o número de amostras de cabelo). Já a matriz de saída foi definida de acordo com a classificação, sendo:

- Pigmentação: matriz 2x127, isto é, as 2 respostas possíveis, cabelo descolorido e original, e as 127 amostras
- Gênero: matriz 2x127, devido as 2 respostas possíveis, feminino e masculino, e as 127 amostras;
- Faixa etária: matriz 3x127, devido as 3 respostas possíveis, infantil (de 0 a 18 anos), adulto (de 19 a 59 anos) e idoso (acima de 60 anos), e as 127 amostras;
- Tratamento químico: matriz 5x127, isto é, as 5 respostas possíveis, virgem, descolorido, alisamento, tintura e tintura e alisamento, e as 127 amostras, sendo desconsiderado o tempo que a amostra teve o procedimento estético.

Em seguida, determinaram-se a quantidade de neurônios da camada intermediária, por tentativa e erro, e a porcentagem de amostras para treinamento, validação e teste, estas que foram respectivamente 70%, 15% e 15%. Por fim, executou-se o treinamento das RNAs, importou-se o código para o MATLAB e realizaram-se os testes alterando o número de neurônios da camada intermediária.

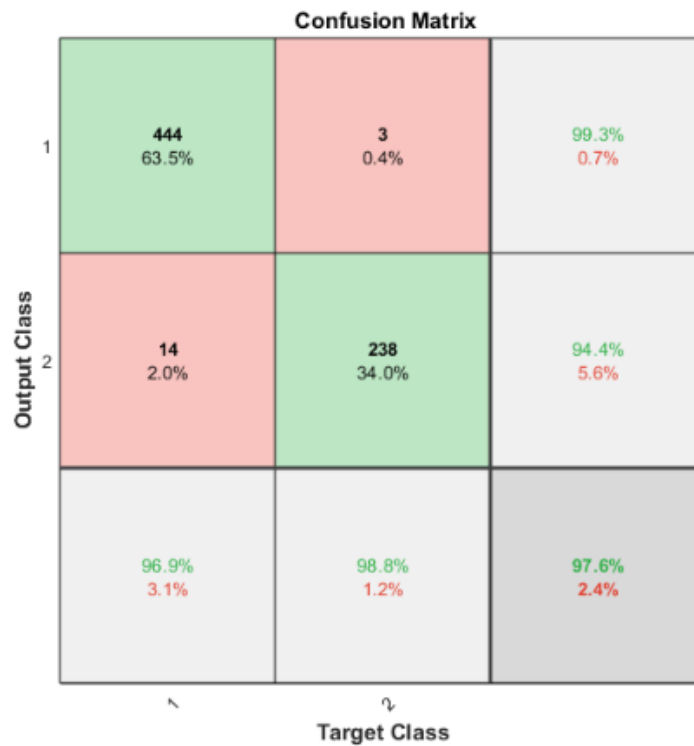
Dessa forma, com o código do MATLAB importado e realizado o treinamento, pôde-se validar o modelo que teve melhor eficácia a partir das matrizes de confusão, como exemplificado na Figura 5, em que as células das diagonais mostram o número de casos classificados corretamente e as de fora da diagonal identificam os casos classificados errados. A célula cinza no canto inferior direito evidencia a porcentagem total de casos classificados corretamente, em verde, e os classificados incorretamente, em vermelho [6].

**Figura 4** - Estrutura da ferramenta *nnstart* para selecionar variáveis de entrada (Inputs) e saída (Targets).



FONTE: MATHWORKS [6].

**Figura 5** - Exemplo de matriz de confusão.



FONTE: MATHWORKS [6].

## **4. RESULTADOS E DISCUSSÃO**

Neste capítulo serão apresentados os resultados obtidos pelas RNAs para classificar as amostras de cabelo por pigmentação, faixa etária, gênero e tratamento químico por cada uma das duas técnicas das quais foram obtidos os dados numéricos, WDXRF e LIBS.

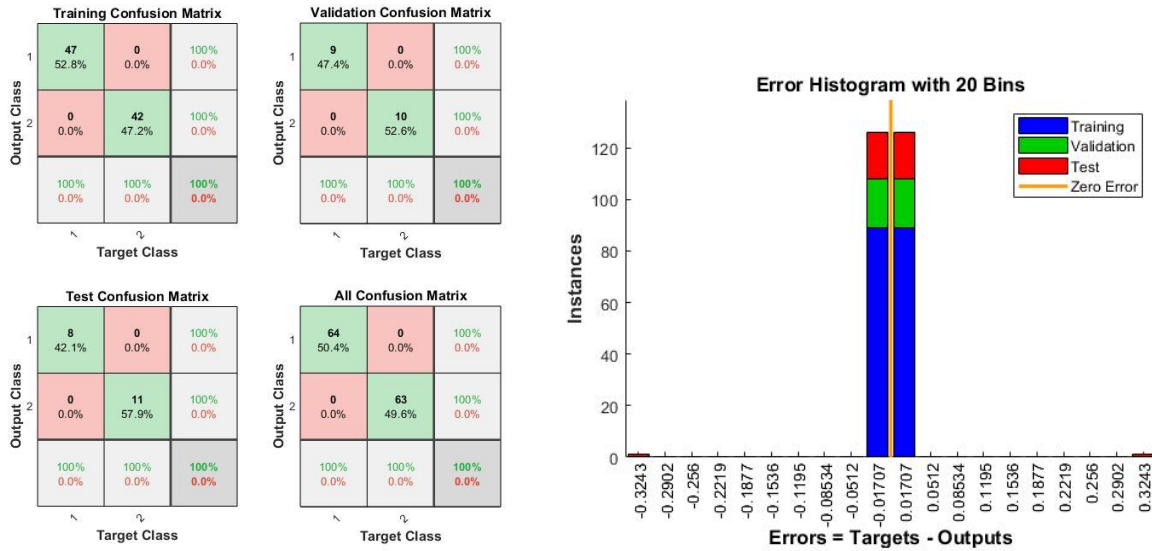
### **4.1 Classificação das amostras de cabelo com dados numéricos por WDXRF, via RNAs**

#### **4.1.1 Pigmentação**

A RNA realizada para a pigmentação possuía 2 respostas possíveis, cabelo descolorido e original, sabendo-se que das 127 amostras, 64 eram de cabelo original e 63 de cabelo descolorido. Ao realizar treinamento com a RNA foi possível obter resultado satisfatório com 20 neurônios na camada intermediária e 49 épocas, com parada precoce pelo gradiente mínimo.

Na Figura 6b é possível notar que o histograma dos erros possui distribuição normal, tendo a frequência mais alta no centro (próximo de zero) e decréscimo gradual para as caudas, com simetria em relação ao eixo central e a média e mediana aproximadamente iguais. Na Figura 6a podem-se observar os acertos da RNA, onde cabelo original e descolorido correspondem às classes 1 e 2, com coordenadas (0,1) e (1,0), respectivamente. Nota-se que não ocorreram erros no treinamento, nem na validação e no teste, e, portanto, a RNA apresentou 100% de acerto em todas as amostras. Esses resultados foram publicados em Berto et al. [10].

**Figura 6 -** (a) Matrizes de confusão e (b) Histograma de erros para pigmentação por WDXRF



(a)

(b)

FONTE: O autor.

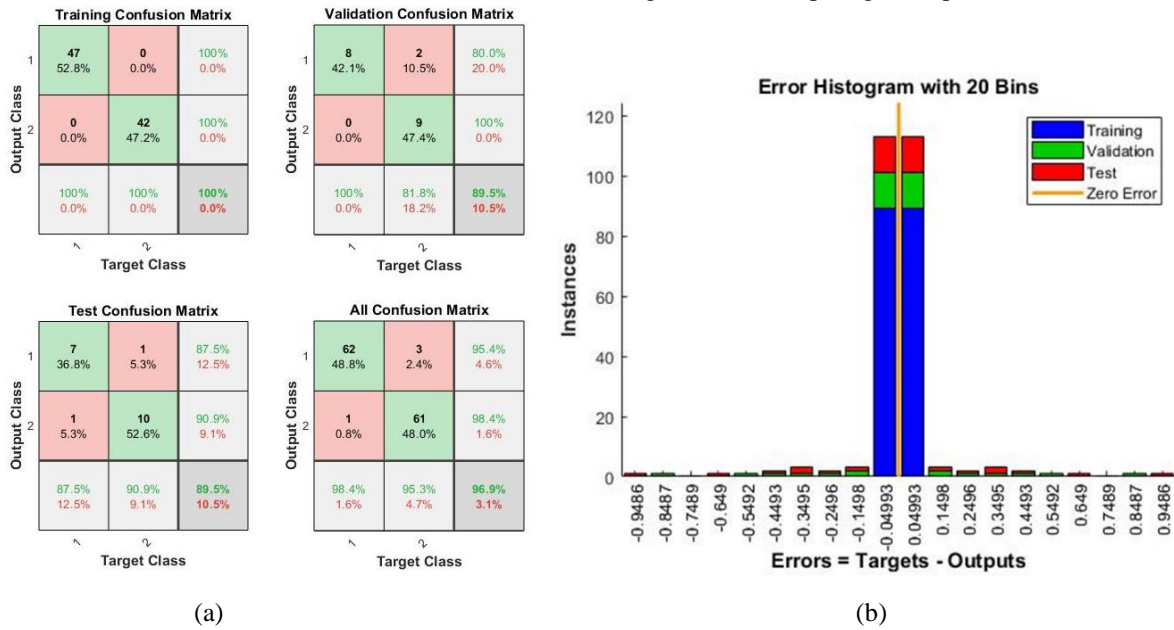
#### 4.1.2 Gênero

Nessa classificação, a matriz de saída da RNA possui dimensão 2x127, devido as 2 respostas possíveis: feminino (classe 1) e masculino (classe 2). Do total de 127 amostras, 63 eram femininos e 64 masculinos. Foi possível obter resultado eficaz com 80 neurônios na camada intermediária e 25 épocas, com parada precoce pelo conjunto de validação.

A eficiência do treinamento pode ser identificada pelas matrizes de confusão, Figura 7a, na qual se nota que não ocorreram erros no treinamento. Todavia, na validação obteve-se 89,5% de acerto, devido a 2 amostras masculinas classificadas incorretamente, e no teste também obteve-se 89,5% de acerto, pois 1 amostra feminina e 1 masculina foram classificadas erroneamente.

Por fim, na Figura 7b, tem-se o histograma de erros, apresentando distribuição normal, simétrico e com média e mediana aproximadamente iguais e localizadas no centro do histograma. Entretanto, verificam-se *outliers*, isto é, dados que estão distantes dos outros valores, porém, são mínimos devido ao total de acerto de 89,5% para o conjunto de teste. Esses resultados foram publicados em Berto et al. [10].

**Figura 7 - (a) Matrizes de confusão e (b) Histograma de erros para gênero por WDXRF**



FONTE: O autor.

Nas Tabelas 1 e 2 é possível identificar as amostras que foram classificadas incorretamente com o grifado em vermelho, tanto no conjunto de teste como no de validação, respectivamente, sendo que a rede neural classifica como 1 o maior valor obtido dentre as 2 coordenadas.

**Tabela 1** - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de teste do WDXRF

Amostra	Respostas obtidas pela RNA		Classe obtida pela RNA	Resposta esperada		Classe esperada
1	0,018	0,982	Masculino	0	1	Masculino
2	0,259	0,741	Masculino	0	1	Masculino
3	0,393	0,607	Masculino	0	1	Masculino
4	0,000	1,000	Masculino	0	1	Masculino
5	0,006	0,994	Masculino	0	1	Masculino
6	0,000	1,000	Masculino	0	1	Masculino
7	0,002	0,998	Masculino	0	1	Masculino
8	0,347	0,653	Masculino	0	1	Masculino
9	0,031	0,969	Masculino	0	1	Masculino
10	0,000	1,000	Masculino	0	1	Masculino
11	0,999	0,001	Feminino	0	1	Masculino
12	0,967	0,033	Feminino	1	0	Feminino
13	0,383	0,617	Masculino	1	0	Feminino
14	1,000	0,000	Feminino	1	0	Feminino
15	0,891	0,109	Feminino	1	0	Feminino
16	0,995	0,005	Feminino	1	0	Feminino
17	0,568	0,432	Feminino	1	0	Feminino
18	1,000	0,000	Feminino	1	0	Feminino
19	0,996	0,004	Feminino	1	0	Feminino

FONTE: O autor.

**Tabela 2** - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de validação do WDXRF

Amostra	Respostas obtidas pela RNA		Classe obtida pela RNA	Respostas esperadas		Classe esperada
1	0,0121	0,9879	Masculino	0	1	Masculino
2	0,0025	0,9975	Masculino	0	1	Masculino
3	0,8466	0,1534	Feminino	0	1	Masculino
4	0,2908	0,7092	Masculino	0	1	Masculino
5	0,0130	0,9870	Masculino	0	1	Masculino
6	0,0160	0,9840	Masculino	0	1	Masculino
7	0,0007	0,9993	Masculino	0	1	Masculino
8	0,0038	0,9962	Masculino	0	1	Masculino
9	0,0051	0,9949	Masculino	0	1	Masculino
10	0,1640	0,8360	Masculino	0	1	Masculino
11	0,5869	0,4131	Feminino	0	1	Masculino
12	0,9973	0,0027	Feminino	1	0	Feminino
13	0,9999	0,0001	Feminino	1	0	Feminino
14	0,8551	0,1449	Feminino	1	0	Feminino
15	0,9999	0,0001	Feminino	1	0	Feminino
16	0,9946	0,0054	Feminino	1	0	Feminino
17	0,6696	0,3304	Feminino	1	0	Feminino
18	0,9392	0,0608	Feminino	1	0	Feminino
19	0,5981	0,4019	Feminino	1	0	Feminino

FONTE: O autor.

#### 4.1.3 Faixa etária

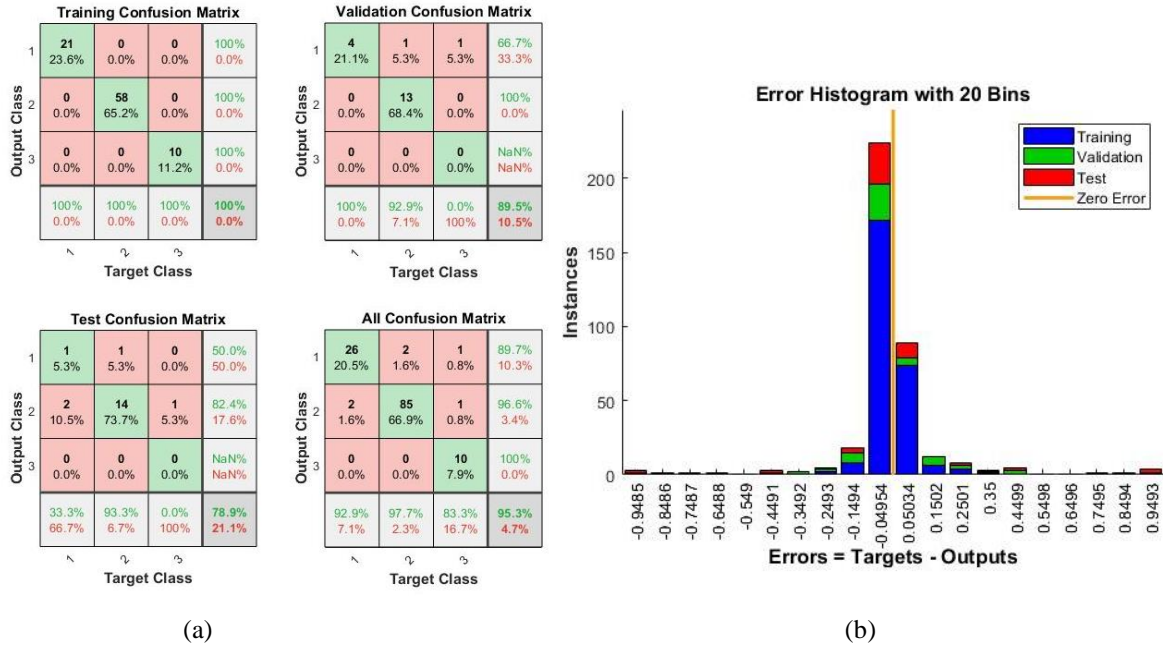
Para a classificação de faixa etária, a matriz de saída é de dimensão 3x127, isto é, as três respostas possíveis de infantil, adulto e idoso, caracterizadas pelas classes 1, 2 e 3 e as coordenadas (1,0,0), (0,1,0) e (0,0,1), respectivamente. Das 127 amostras selecionadas, 28 eram infantis, 87 adultos e 12 idosos. O resultado satisfatório ocorreu com 44 neurônios na camada intermediária e 19 épocas, com parada precoce pelo conjunto de validação.

Assim, a partir das matrizes de confusão, na Figura 8a, nota-se que não ocorreram erros no treinamento, enquanto na validação obteve-se 89,5% de acerto, devido a uma amostra de adulto classificada como infantil e uma de idoso classificada como infantil. Já no teste ocorreram quatro erros, gerando 78,9% de acerto, devido a duas amostras infantis classificadas como adulto, uma amostra de adulto classificada como infantil e uma de idoso classificada como adulto. Em ambos os casos, a RNA não classificou corretamente as amostras de idosos e, uma possibilidade para isso é o fato de ter apenas 12 amostras dessa classe.

No histograma de erros, na Figura 8b, tem-se distribuição normal, simétrico, sendo a média e mediana aproximadamente iguais e localizadas no centro do histograma. Verifica-se

outliers, porém são poucos se considerando o total de acerto de 78,9% para o conjunto de teste e 89,5% para a validação.

**Figura 8 -** (a) Matrizes de confusão e (b) - Histograma de erros para faixa etária por WDXRF



**FONTE:** O autor.

Na Tabela 3, é possível analisar tais resultados obtidos e os esperados pela RNA para o conjunto de validação, e na Tabela 4 para o conjunto de teste, destacando em vermelho os classificados incorretamente, sendo que a rede neural classifica como 1 o maior valor obtido dentre as 3 coordenadas.

**Tabela 3 - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de validação do WDXRF**

Amostra	Respostas obtidas pela RNA			Classe obtida pela RNA	Respostas esperadas			Classe esperada
1	0,6362	0,2220	0,1418	Infantil	1	0	0	Infantil
2	0,8461	0,0758	0,0781	Infantil	1	0	0	Infantil
3	0,5944	0,3592	0,0464	Infantil	1	0	0	Infantil
4	0,7952	0,0971	0,1077	Infantil	1	0	0	Infantil
5	0,0100	0,9874	0,0027	Adulto	0	1	0	Adulto
6	0,9396	0,0311	0,0293	Infantil	0	1	0	Adulto
7	0,0109	0,9793	0,0098	Adulto	0	1	0	Adulto
8	0,2112	0,7445	0,0443	Adulto	0	1	0	Adulto
9	0,4866	0,5034	0,0099	Adulto	0	1	0	Adulto
10	0,1029	0,8730	0,0242	Adulto	0	1	0	Adulto
11	0,1362	0,8433	0,0205	Adulto	0	1	0	Adulto
12	0,1243	0,8624	0,0133	Adulto	0	1	0	Adulto
13	0,1044	0,8240	0,0716	Adulto	0	1	0	Adulto
14	0,0200	0,9735	0,0065	Adulto	0	1	0	Adulto
15	0,0728	0,9090	0,0182	Adulto	0	1	0	Adulto
16	0,0519	0,8894	0,0587	Adulto	0	1	0	Adulto
17	0,0001	0,9763	0,0236	Adulto	0	1	0	Adulto
18	0,0431	0,5674	0,3896	Adulto	0	1	0	Adulto
19	0,6641	0,1281	0,2078	Infantil	0	0	1	Idoso

FONTE: O autor.

**Tabela 4** - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de teste do WDXRF

Amostra	Respostas obtidas pela RNA			Classe obtida pela RNA	Respostas esperadas			Classe esperada
1	0,0008	0,9984	0,0008	Adulto	1	0	0	Infantil
2	0,6935	0,2694	0,0371	Infantil	1	0	0	Infantil
3	0,0433	0,9541	0,0027	Adulto	1	0	0	Infantil
4	0,0004	0,9978	0,0018	Adulto	0	1	0	Adulto
5	0,0003	0,9991	0,0006	Adulto	0	1	0	Adulto
6	0,0144	0,9821	0,0035	Adulto	0	1	0	Adulto
7	0,7881	0,1905	0,0214	Infantil	0	1	0	Adulto
8	0,4059	0,5182	0,0759	Adulto	0	1	0	Adulto
9	0,0004	0,9950	0,0046	Adulto	0	1	0	Adulto
10	0,4093	0,5429	0,0478	Adulto	0	1	0	Adulto
11	0,1774	0,7704	0,0522	Adulto	0	1	0	Adulto
12	0,0251	0,9632	0,0117	Adulto	0	1	0	Adulto
13	0,0003	0,9994	0,0003	Adulto	0	1	0	Adulto
14	0,0020	0,9180	0,0800	Adulto	0	1	0	Adulto
15	0,0113	0,9697	0,0190	Adulto	0	1	0	Adulto
16	0,1202	0,7855	0,0943	Adulto	0	1	0	Adulto
17	0,0003	0,9996	0,0001	Adulto	0	1	0	Adulto
18	0,0274	0,9474	0,0252	Adulto	0	1	0	Adulto
19	0,1418	0,8367	0,0214	Adulto	0	0	1	Idoso

FONTE: O autor.

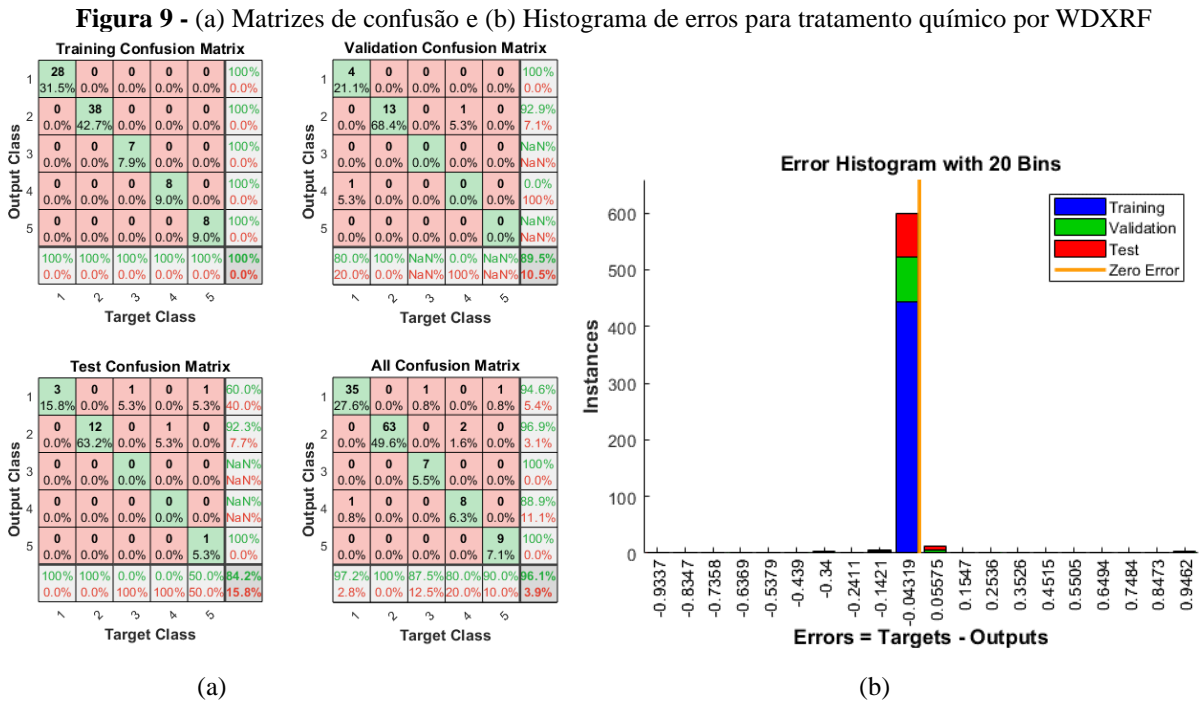
#### 4.1.4 Tratamento químico

Na classificação do cabelo humano em relação ao tratamento químico, a rede neural recebeu uma matriz de saída com dimensão 5x127, caracterizando as 5 respostas possíveis: virgem, descolorido, alisamento, tintura e alisamento e tintura, com as classes 1, 2, 3, 4 e 5, e as coordenadas (1,0,0,0,0), (0,1,0,0,0), (0,0,1,0,0), (0,0,0,1,0) e (0,0,0,0,1), respectivamente. Das 127 amostras selecionadas 36 eram virgens, 63 descoloridos, 8 com alisamento, 10 com tinturas e 10 com alisamento e tintura. Para se obter resultado satisfatório, foram utilizados 122 neurônios na camada intermediária e 34 épocas, com parada precoce pelo conjunto de validação.

A eficiência do treinamento pode ser visualizada nas matrizes de confusão, Figura 9a, onde nota-se que não ocorreram erros no treinamento, enquanto na validação obtiveram-se dois erros, devido a uma amostra com tintura classificada como descolorida e uma virgem classificada com tintura, caracterizando 89,5% de acerto. Já no teste obtiveram-se três erros, sendo um de amostra com alisamento classificada como virgem, uma de alisamento e tintura classificada como virgem e uma com tintura classificada como descolorida, gerando 84,2% de

acerto. No teste e validação nota-se que não se classificou amostras com alisamento, provavelmente devido a pequena quantidade da mesma, sendo apenas 8.

No histograma de erros, na Figura 9b, pode-se classificá-lo como tendo distribuição normal, com simetria e pequenos *outliers* proveniente dos erros da validação e teste, estes que foram poucos considerando as porcentagens de acerto de 89,5% e 84,2%, respectivamente.



FONTA: O autor.

Os resultados obtidos e esperados pela RNA para validação e teste estão dispostos nas Tabelas 5 e 6, respectivamente. Os dados classificados incorretamente estão destacados em vermelho, sendo que a rede neural classifica como 1 o maior valor obtido dentre as 5 coordenadas.

**Tabela 5** - Respostas obtidas e esperadas pela RNA para o tratamento químico no conjunto de validação do WDXRF

Amostra	Respostas obtidas pela RNA					Classe obtida pela RNA	Respostas esperadas					Classe esperada
							0	1	0	0	0	
1	5,69E-05	0,9998	0,000112	6,39E-06	5,26E-06	D	0	1	0	0	0	D
2	1,86E-10	1,0000	8,33E-08	4,23E-06	1,39E-08	D	0	1	0	0	0	D
3	0,0456	0,8448	0,00031	0,07192	0,03745	D	0	0	0	1	0	TA
4	0,0002	0,9872	0,00465	0,000226	0,007681	D	0	1	0	0	0	D
5	0,0060	0,9940	1,23E-05	2,63E-05	2,58E-06	D	0	1	0	0	0	D
6	0,0003	0,9996	5,31E-06	0,000135	9,44E-06	D	0	1	0	0	0	D
7	0,5450	0,1143	0,01641	0,295875	0,02846	V	1	0	0	0	0	V
8	0,0001	0,9996	0,00016	1,16E-05	0,00011	D	0	1	0	0	0	D
9	0,9856	1,83E-05	0,01439	2,21E-06	4,83E-06	V	1	0	0	0	0	V
10	0,7692	0,0220	0,01488	0,068648	0,1253	V	1	0	0	0	0	V
11	0,0002	0,9989	0,000543	0,000131	0,000199	D	0	1	0	0	0	D
12	0,0178	0,9820	8,38E-05	7,32E-05	7,03E-07	D	0	1	0	0	0	D
13	0,0154	0,9838	0,0006532	1,58E-05	0,00012	D	0	1	0	0	0	D
14	0,5217	0,0554	0,0029925	0,407983	0,01197	V	1	0	0	0	0	V
15	2,86E-05	1,0000	1,93E-06	4,96E-07	9,13E-07	D	0	1	0	0	0	D
16	0,00011	0,9998	3,29E-05	5,44E-06	6,28E-05	D	0	1	0	0	0	D
17	0,1676	0,0142	0,1153	0,6426	0,06019	TA	1	0	0	0	0	V
18	6,51E-09	1,0000	2,95E-09	5,09E-07	1,52E-08	D	0	1	0	0	0	D
19	0,002944	0,9963	0,0002497	0,000488	3,57E-05	D	0	1	0	0	0	D

Legenda: Virgem (V), D (descolorido), A(alisamento), TA (tintura e alisamento) e T (tintura)

**FONTE:** O autor.

**Tabela 6** - Respostas obtidas e esperadas pela RNA para o tratamento químico no conjunto de teste do WDXRF

Amostra	Respostas obtidas pela RNA					Classes obtidas pela RNA	Respostas esperadas					Classes esperadas
							0	0	0	0	1	
1	0,9832	0,0064	0,00578	0,0004	0,00428	V	0	0	0	0	1	T
2	6,09E-08	1,000	3,47E-05	1,65E-06	2,40E-06	D	0	1	0	0	0	D
3	0,0612	0,931	2,09E-05	0,00717	0,00065	D	0	1	0	0	0	D
4	6,12E-06	0,9998	5,40E-05	9,38E-06	0,00011	D	0	1	0	0	0	D
5	0,9711	0,0009	0,00217	0,003	0,02286	V	1	0	0	0	0	V
6	2,12E-10	1,000	5,79E-10	4,10E-07	1,47E-10	D	0	1	0	0	0	D
7	0,0457	0,7971	0,01013	0,13843	0,0086	D	0	0	0	1	0	T
8	0,0871	0,9123	0,00053	9,11E-06	0,0001	D	0	1	0	0	0	D
9	0,0097	0,9902	1,91E-05	9,78E-05	2,47E-05	D	0	1	0	0	0	D
10	0,0001	0,9999	7,80E-07	2,77E-06	9,75E-07	D	0	1	0	0	0	D
11	0,0471	0,932	0,01133	0,00556	0,00404	D	0	1	0	0	0	D
12	0,9784	0,0093	0,00236	0,00719	0,0027	V	1	0	0	0	0	V
13	1,29E-05	1,000	1,24E-07	2,33E-06	2,01E-07	D	0	1	0	0	0	D
14	0,9885	2,17E-06	0,00681	0,00435	0,00033	V	1	0	0	0	0	V
15	0,0006	0,999	4,30E-05	0,00028	4,29E-05	D	0	1	0	0	0	D
16	0,00428	0,9789	4,77E-05	0,00112	0,0157	D	0	1	0	0	0	D
17	0,38165	0,0027	0,1618	0,0106	0,4432	T	0	0	0	0	1	T
18	6,01E-06	0,9997	0,00012	0,00018	3,31E-07	D	0	1	0	0	0	D
19	0,3926	0,0117	0,0329	0,3801	0,1827	V	0	0	1	0	0	TA

Legenda: Virgem (V), D (descolorido), A(alisamento), TA (tintura e alisamento) e T (tintura)

FONTE: O autor.

## 4.2 Classificação das amostras de cabelo com dados numéricos por LIBS, via RNAs

As classificações com os dados numéricos pelo LIBS utilizaram as mesmas matrizes de saída que do WDXRF, pois as classes de cabelo eram semelhantes em ambos os casos e a quantidade de amostras de cada classe também permaneceu a mesma.

### 4.2.1 Pigmentação

A rede neural artificial para classificação das amostras de cabelo em relação à pigmentação (cabelo original e descolorido) possuía 10 neurônios na camada intermediária e 74 épocas, com parada precoce pelo gradiente mínimo. Obtiveram-se 100% de acerto em todas as amostras, tanto no treinamento quanto na validação e teste, como ilustra as matrizes de confusão da Figura 10a. Já o histograma de erros da Figura 10b, identifica que o desempenho da RNA foi muito eficaz. Esses resultados foram publicados em Berto et al. [10].

**Figura 10** - (a) Matrizes de confusão e (b) Histograma de erros para pigmentação por LIBS



(a)

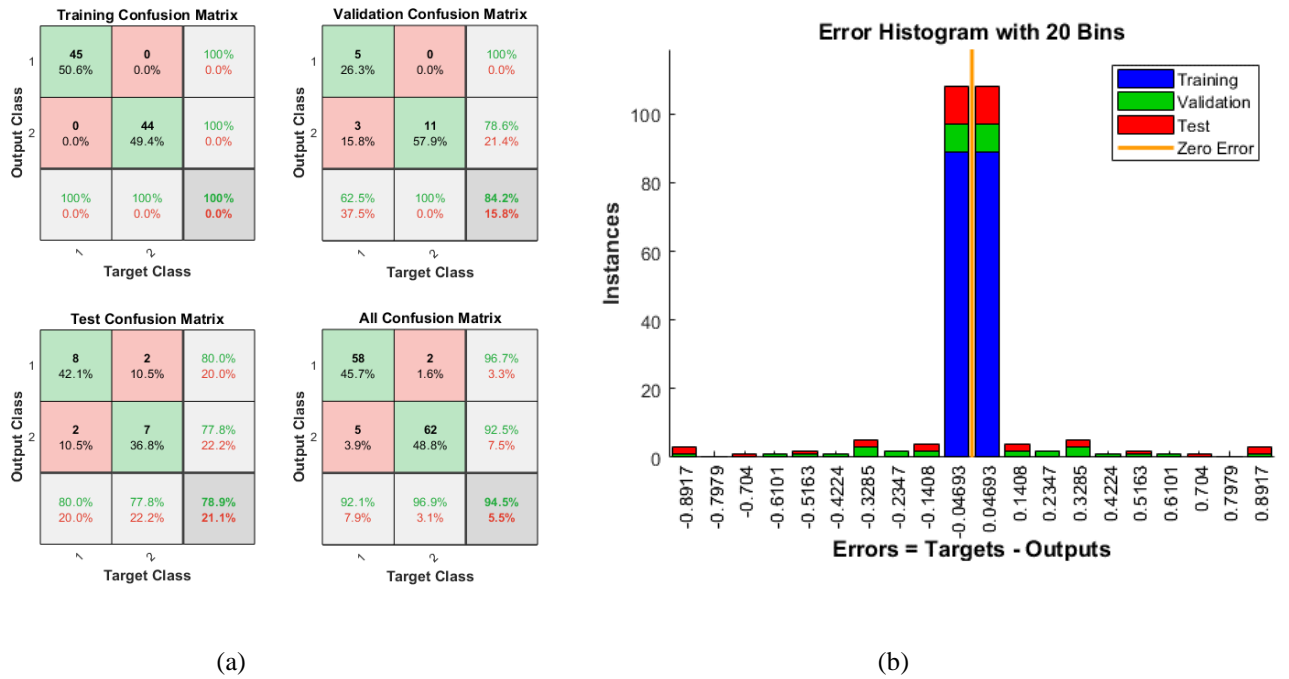
(b)

FONTE: O autor.

#### 4.2.2 Gênero

Na classificação de gênero, a rede neural artificial possuía 115 neurônios na camada intermediária e 25 épocas, com parada precoce pelo conjunto de validação. Os resultados evidenciaram 100% de acerto no conjunto onde treinamento, 84,2% no conjunto de validação, onde 3 amostras femininas foram classificadas erroneamente como masculinas, e 78,9% de acerto no conjunto de teste em que 2 amostras de cada classe foram classificadas incorretamente, como ilustra as matrizes de confusão da Figura 11a. No histograma de erros da Figura 11b, nota-se que o mesmo é normalmente distribuído, comprovando o bom desempenho da RNA. Esses resultados foram publicados em Berto et al. [10].

**Figura 11 - (a) Matrizes de confusão e (b) Histograma de erros para classificação de gênero por LIBS**



FONTE: O autor.

Além disso, nas Tabelas 7 e 8, é possível identificar, grifado em vermelho, as amostras que foram classificadas incorretamente nos conjuntos de teste e validação, respectivamente.

**Tabela 7** - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de teste do LIBS

Amostra	Resposta obtida pela RNA		Classe obtida pela RNA	Respostas esperadas		Classe esperada
1	0,905	0,095	Feminino	0	1	Masculino
2	0,018	0,982	Masculino	0	1	Masculino
3	0,006	0,994	Masculino	0	1	Masculino
4	0,010	0,991	Masculino	0	1	Masculino
5	0,184	0,816	Masculino	0	1	Masculino
6	0,081	0,919	Masculino	0	1	Masculino
7	0,059	0,941	Masculino	0	1	Masculino
8	0,001	0,999	Masculino	0	1	Masculino
9	0,515	0,485	Feminino	0	1	Masculino
10	0,999	0,001	Masculino	1	0	Masculino
11	0,641	0,359	Masculino	1	0	Masculino
12	0,963	0,038	Feminino	1	0	Feminino
13	0,708	0,292	Feminino	1	0	Feminino
14	0,823	0,177	Feminino	1	0	Feminino
15	0,128	0,872	Masculino	1	0	Feminino
16	0,341	0,659	Masculino	1	0	Feminino
17	0,992	0,008	Feminino	1	0	Feminino
18	0,974	0,027	Feminino	1	0	Feminino
19	1,000	0,001	Feminino	1	0	Feminino

FONTE: O autor.

**Tabela 8** - Respostas obtidas e esperadas pela RNA para o gênero no conjunto de validação do LIBS

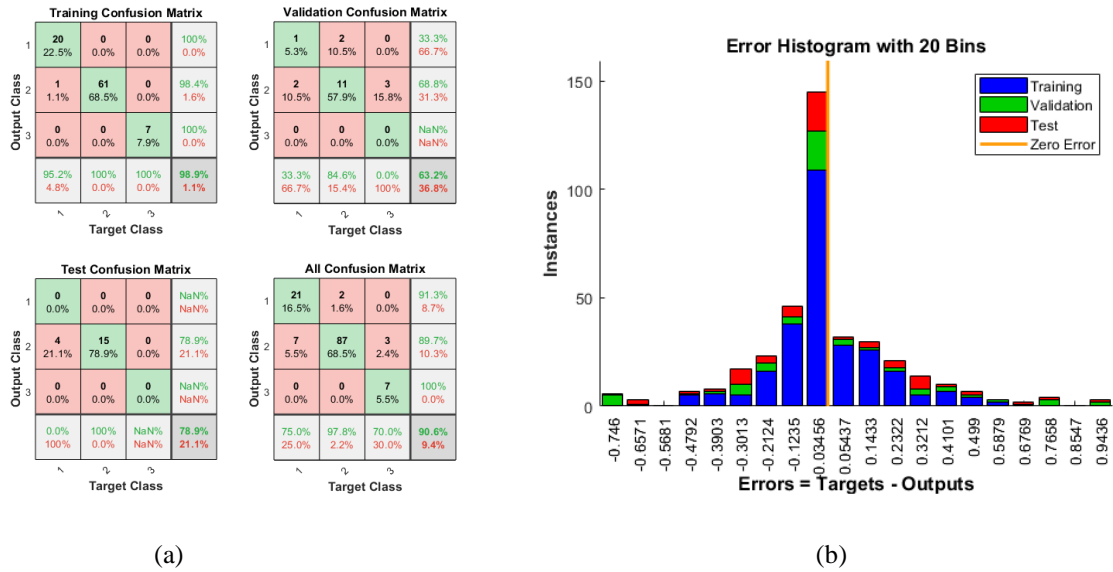
Amostra	Respostas obtidas pela RNA		Classe obtida pela RNA	Respostas esperadas		Classe esperada
1	0,4500	0,5500	Masculino	1	0	Feminino
2	0,9175	0,0825	Feminino	1	0	Feminino
3	0,8838	0,1162	Feminino	1	0	Feminino
4	0,8002	0,1998	Feminino	1	0	Feminino
5	0,9178	0,0822	Feminino	1	0	Feminino
6	0,0613	0,9387	Feminino	1	0	Feminino
7	0,4343	0,5657	Masculino	1	0	Feminino
8	0,9474	0,0526	Feminino	1	0	Feminino
9	0,1150	0,8850	Masculino	0	1	Masculino
10	0,0269	0,9731	Masculino	0	1	Masculino
11	0,3184	0,6816	Masculino	0	1	Masculino
12	0,2400	0,7600	Masculino	0	1	Masculino
13	0,3475	0,6525	Masculino	0	1	Masculino
14	0,0320	0,9680	Masculino	0	1	Masculino
15	0,0170	0,9830	Masculino	0	1	Masculino
16	0,3507	0,6493	Masculino	0	1	Masculino
17	0,0819	0,9181	Masculino	0	1	Masculino
18	0,4016	0,5984	Masculino	0	1	Masculino
19	0,0044	0,9956	Masculino	0	1	Masculino

FONTE: O autor.

#### 4.2.3 Faixa etária

Na classificação das amostras de cabelo por faixa etária, obteve-se resultado satisfatório com 56 neurônios na camada intermediária e 19 épocas, com parada precoce pelo conjunto de validação. Os resultados mostraram 98,9% de acerto no conjunto de treinamento, já que uma amostra adulta foi classificada como infantil. No conjunto de validação foi obtido 63,2% de acerto, pois duas amostras infantis foram classificadas como adulto, duas adultas classificadas como infantis e três amostras idosas classificadas como adulta. No conjunto de teste foi obtido 78,9% de acerto, onde quatro amostras de adulto foram classificadas como infantis. Na Figura 12a é ilustrado as matrizes de confusão, onde é identificado que nenhuma amostra de idoso foi classificada corretamente, decorrente da pequena quantidade de amostras dessa classe. Na Figura 12b o histograma de erros evidencia os erros descritos anteriormente, com a presença de *outliers*.

Figura 12 - (a) Matrizes de confusão e (b) Histograma de erros para faixa etária por LIBS



FONTE: O autor.

Nas Tabelas 9 e 10 tem-se destaque em vermelho as amostras que foram classificadas incorretamente em cada conjunto, de teste e validação, respectivamente.

**Tabela 9** - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de teste do LIBS

Amostra	Respostas obtidas pela RNA			Classe obtida pela RNA	Respostas esperadas			Classe esperada
1	0,2766	0,7219	0,0015	Adulto	0	1	0	Adulto
2	0,0203	0,7484	0,2313	Adulto	0	1	0	Adulto
3	0,0162	0,6695	0,3143	Adulto	0	1	0	Adulto
4	0,0041	0,9119	0,0839	Adulto	0	1	0	Adulto
5	0,0802	0,6740	0,2458	Adulto	0	1	0	Adulto
6	0,1199	0,5242	0,3559	Adulto	0	1	0	Adulto
7	0,0068	0,6693	0,3239	Adulto	0	1	0	Adulto
8	0,0058	0,8881	0,1061	Adulto	0	1	0	Adulto
9	0,3294	0,6195	0,0511	Adulto	0	1	0	Adulto
10	0,0239	0,8990	0,0771	Adulto	0	1	0	Adulto
11	0,0119	0,6451	0,3430	Adulto	1	0	0	Infantil
12	0,2218	0,7745	0,0037	Adulto	1	0	0	Infantil
13	0,3371	0,6472	0,0157	Adulto	1	0	0	Infantil
14	0,4926	0,5003	0,0070	Adulto	1	0	0	Infantil
15	0,2615	0,7120	0,0265	Adulto	0	1	0	Adulto
16	0,0287	0,8092	0,1622	Adulto	0	1	0	Adulto
17	0,0740	0,8619	0,0641	Adulto	0	1	0	Adulto
18	0,0096	0,6974	0,2930	Adulto	0	1	0	Adulto
19	0,0472	0,7590	0,1937	Adulto	0	1	0	Adulto

FONTE: O autor.

**Tabela 10** - Respostas obtidas e esperadas pela RNA para a faixa etária no conjunto de validação do LIBS

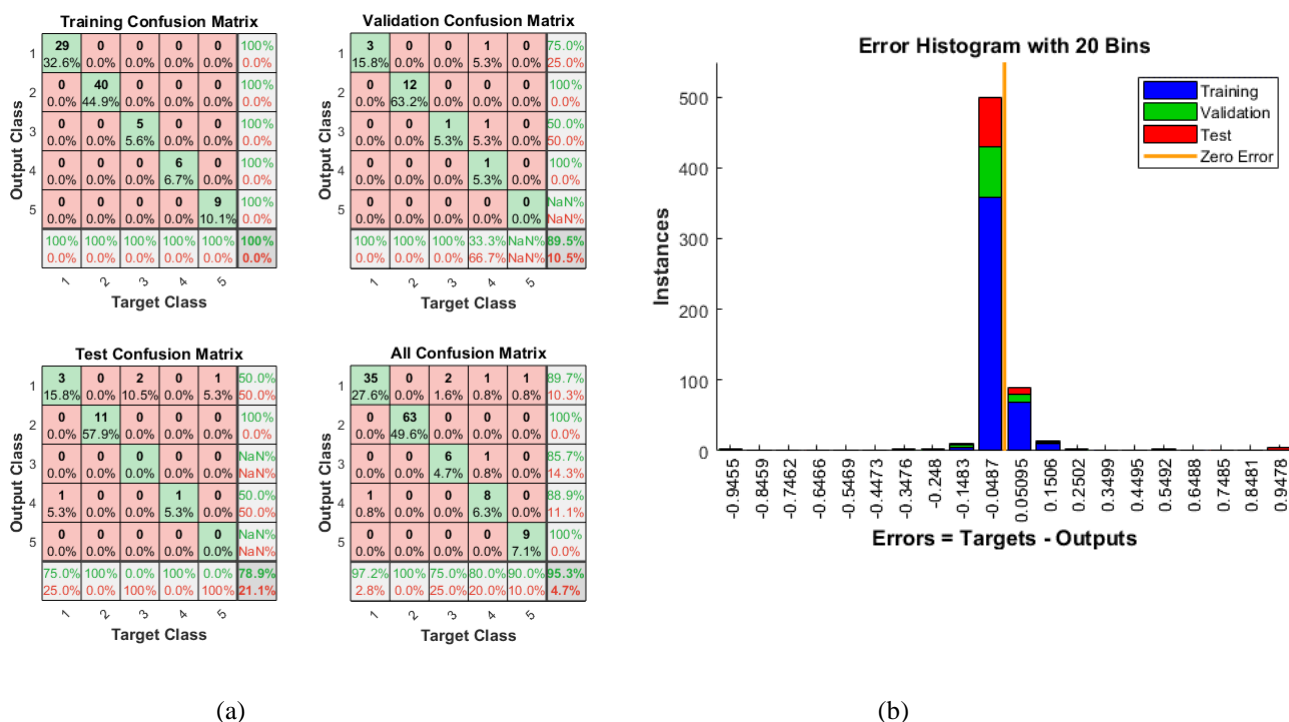
Amostra	Respostas obtidas pela RNA			Classe obtida pela RNA	Respostas esperadas			Classe esperada
1	0,6911	0,2680	0,0409	Infantil	1	0	0	Infantil
2	0,2349	0,7079	0,0572	Adulto	1	0	0	Infantil
3	0,0787	0,6197	0,3015	Adulto	0	1	0	Adulto
4	0,3394	0,6375	0,0231	Adulto	0	1	0	Adulto
5	0,2815	0,6173	0,1013	Adulto	0	1	0	Adulto
6	0,6245	0,3671	0,0084	Infantil	0	1	0	Adulto
7	0,3591	0,5097	0,1312	Adulto	0	1	0	Adulto
8	0,0672	0,8912	0,0416	Adulto	0	1	0	Adulto
9	0,0080	0,7682	0,2237	Adulto	0	0	1	Idoso
10	0,0336	0,7209	0,2455	Adulto	0	0	1	Idoso
11	0,0630	0,7347	0,2022	Adulto	1	0	0	Infantil
12	0,0386	0,9278	0,0336	Adulto	0	1	0	Adulto
13	0,0067	0,9748	0,0185	Adulto	0	1	0	Adulto
14	0,2122	0,7739	0,0139	Adulto	0	1	0	Adulto
15	0,4743	0,4185	0,1072	Infantil	0	1	0	Adulto
16	0,1700	0,7625	0,0675	Adulto	0	1	0	Adulto
17	0,0128	0,9865	0,0007	Adulto	0	1	0	Adulto
18	0,0598	0,6442	0,2961	Adulto	0	1	0	Adulto
19	0,1944	0,7904	0,0151	Adulto	0	0	1	Idoso

FONTE: O autor.

#### 4.2.4 Tratamento químico

Para a classificação das amostras de cabelo por tratamento químico, utilizou-se uma rede neural artificial com 18 neurônios na camada intermediária e 34 épocas, com parada precoce pelo conjunto de validação. Os resultados mostraram 100% de acerto no conjunto de treinamento, 89,5% de acerto no conjunto de validação, onde uma amostra de cabelo tingido foi classificada incorretamente como cabelo original e outra amostra de cabelo tingido foi classificada como alisado. No conjunto de teste houve 78,9% de acerto pois uma amostra de cabelo original foi classificada como tingida, 2 amostras alisadas foram classificadas como virgem, uma amostra de cabelo virgem classificada como tintura e uma amostra de tintura e alisamento classificada como virgem. Na Figura 13a é ilustrada as matrizes de confusão com as porcentagens de acertos em cada conjunto e na Figura 13b o histograma de erros com distribuição normal.

Figura 13 - (a) Matrizes de confusão e (b) Histograma de erros para tratamento químico por LIBS



FONTE: O autor.

Nas Tabelas 11 e 12 é possível identificar as amostras que foram classificadas incorretamente nos conjuntos de teste e validação, respectivamente.

**Tabela 11** - Resultados obtidos e esperados pela RNA para tratamento químico no conjunto de teste do LIBS

Amostra	Respostas obtidas pela RNA					Classe obtida pela RNA	Respostas esperadas					Classe esperada
							1	0	0	0	0	
1	0,782	0,006	0,129	0,063	0,021	V	1	0	0	0	0	V
2	0,856	0,016	0,073	0,029	0,026	V	1	0	0	0	0	V
3	0,213	0,122	0,037	0,623	0,005	TA	1	0	0	0	0	V
4	0,887	0,003	0,062	0,027	0,021	V	1	0	0	0	0	V
5	0,000	0,999	0,000	0,000	0,000	D	0	1	0	0	0	D
6	0,000	1,000	0,000	0,000	0,000	D	0	1	0	0	0	D
7	0,001	0,998	0,000	0,001	0,000	D	0	1	0	0	0	D
8	0,004	0,992	0,001	0,003	0,000	D	0	1	0	0	0	D
9	0,004	0,984	0,001	0,010	0,002	D	0	1	0	0	0	D
10	0,003	0,994	0,000	0,003	0,000	D	0	1	0	0	0	D
11	0,006	0,984	0,002	0,007	0,001	D	0	1	0	0	0	D
12	0,001	0,998	0,000	0,001	0,000	D	0	1	0	0	0	D
13	0,013	0,920	0,012	0,048	0,008	D	0	1	0	0	0	D
14	0,004	0,994	0,000	0,002	0,000	D	0	1	0	0	0	D
15	0,019	0,936	0,002	0,043	0,001	D	0	1	0	0	0	D
16	0,995	0,000	0,002	0,001	0,001	V	0	0	1	0	0	A
17	0,895	0,003	0,067	0,013	0,023	V	0	0	1	0	0	A
18	0,114	0,403	0,065	0,413	0,005	TA	0	0	0	1	0	TA
19	0,953	0,008	0,017	0,001	0,021	V	0	0	0	0	1	T

Legenda: Virgem (V), D (descolorido), A(alisamento), TA (tintura e alisamento) e T (tintura)

**FONTE:** O autor.

**Tabela 12 - Resultados obtidos e esperados pela RNA para tratamento químico no conjunto de validação no LIBS**

Amostra	Respostas obtidas pela RNA					Classe obtida pela RNA	Respostas esperadas					Classe esperada
							0	1	0	0	0	
1	0,0007	0,9987	0,0002	0,0003	0,0002	D	0	1	0	0	0	D
2	0,2107	0,1461	0,1429	0,4264	0,0739	TA	0	0	0	1	0	TA
3	0,0010	0,9937	0,0014	0,0022	0,0017	D	0	1	0	0	0	D
4	0,0004	0,9992	0,0001	0,0003	0,0001	D	0	1	0	0	0	D
5	0,0005	0,9977	0,0005	0,0006	0,0007	D	0	1	0	0	0	D
6	0,9862	0,0010	0,0064	0,0037	0,0027	V	0	0	0	1	0	TA
7	0,2767	0,0587	0,3824	0,1881	0,0941	A	0	0	0	1	0	TA
8	0,9933	0,0003	0,0029	0,0032	0,0002	V	1	0	0	0	0	V
9	0,0005	0,9991	0,0001	0,0003	0,0000	D	0	1	0	0	0	D
10	0,0006	0,9972	0,0003	0,0015	0,0004	D	0	1	0	0	0	D
11	0,7894	0,0390	0,0782	0,0662	0,0272	V	1	0	0	0	0	V
12	0,0032	0,9942	0,0002	0,0024	0,0001	D	0	1	0	0	0	D
13	0,0006	0,9988	0,0002	0,0003	0,0001	D	0	1	0	0	0	D
14	0,0009	0,9971	0,0002	0,0018	0,0001	D	0	1	0	0	0	D
15	0,0035	0,9842	0,0025	0,0089	0,0009	D	0	1	0	0	0	D
16	0,8302	0,0231	0,0358	0,0069	0,1039	V	1	0	0	0	0	V
17	0,0027	0,9907	0,0007	0,0055	0,0003	D	0	1	0	0	0	D
18	0,3073	0,0814	0,4505	0,0941	0,0667	A	0	0	1	0	0	A
19	0,0053	0,9862	0,0008	0,0076	0,0002	D	0	1	0	0	0	D

Legenda: Virgem (V), D (descolorido), A(alisamento), TA (tintura e alisamento) e T (tintura)

**FONTE:** O autor, 2020.

## 5. CONCLUSÃO

Os resultados obtidos pelas redes neurais artificiais para classificação de cabelos humanos usando dados da técnica WDXRF em relação ao gênero, faixa etária e tratamento químico foram mais promissores do que usando os dados do LIBS, considerando a precisão alta e mais classes classificadas corretamente. Em ambos os casos os gráficos de distribuição de erro estiveram na faixa esperada, caracterizada pela distribuição normal, com maior frequência no centro e diminuição gradativa para as caudas, simetricamente, com média e mediana aproximadamente iguais.

Nas classificações de faixa etária e tratamento químico ocorreu a não classificação de algumas classes, devido ao fato de existirem poucas amostras em relação as demais, além dos dados dos cabelos serem complexos.

Portanto, a tecnologia de Redes Neurais Artificiais pode ser utilizada para classificar amostras de cabelo em relação à pigmentação, gênero, faixa etária e tratamento químico. E um estudo futuro com mais amostras das outras classes poderá gerar resultados melhores, indicando que RNAs poderiam ser utilizadas para as outras classificações também.

## 6. REFERÊNCIAS

- [1] BRAGA, A.P.; LUDERMIR, T.B; CARVALHO, A.C.P.L.F. **Redes Neurais Artificiais: Teoria e Aplicações**. Rio de Janeiro: LTC, 2000.
- [2] SANTOS, M. C.; SPERANÇA, M. A.; PEREIRA, F. M. V. **Wavelength dispersive X-ray fluorescence (WDXRF) and chemometric investigation of human hair after cosmetic treatment**. X-Ray Spectrometry, v.47, p. 252-257, Mar. 2018. doi.org/10.1002/xrs.2836.
- [3] CASTRO, L. N. ; VON ZUBEN, F. J. **Redes Neurais Artificiais**. UNICAMP, 2015. Disponível em: [ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia006\\_03/topico5\\_03.pdf](ftp://ftp.dca.fee.unicamp.br/pub/docs/vonzuben/ia006_03/topico5_03.pdf). Acesso em: 01 set. 2020.
- [4] RAMOS, S. P. J. **Redes Neurais Artificiais na classificação de frutos: cenário bidimensional**. 2003. Disponível em: <https://www.scielo.br/pdf/cagro/v27n2/a15v27n2.pdf>. Acesso em: 31 ago. 2020.
- [5] KHAN, J. et al. **Classification and diagnostic prediction of cancers using gene expression profiling and artificial neural networks**. 2001.
- [6] MATHWORKS. **Classifique padrões com uma Rede Neural Artificial**. Disponível em: <https://www.mathworks.com/help/deeplearning/gs/classify-patterns-with-a-neural-network.html>. Acesso em: 31 ago. 2020.
- [7] BARBOSA, J. et al. **Hair as an alternative matrix in bioanalysis**. Bioanal. 5 (2013) 895-914.
- [8] SAMANIDOU, V. F. **The challenges in hair analysis from the perspective of an analytical chemist**. J. Appl. Bioanal. 2 (2016) 103-107.
- [9] POZEBON D.; SCHEFFLER G. L.; DRESSLER V. L. **Elemental hair analysis: A review of procedures and applications**. Anal. Chim. Acta 992 (2017) 1-23.

- [10] BERTO, T. M.; DOS SANTOS, M. C.; PEREIRA, F. M. V; FILLETTI, É. R. **Artificial neural networks applied to the classification of hair samples according to pigment and sex using noninvasive analytical techniques.** X-RAY SPECTROMETRY. 2020; 1-10. <https://doi.org/10.1002/xrs.3163>
- [11] CORRÊA, F. **Aplicação de redes neurais artificiais no setor de transportes no Brasil /** Disponível em: <https://repositorio.ufscar.br/bitstream/handle/ufscar/4257/2582.pdf?sequence=1>. Acesso em: 02 set. 2020. São Carlos: UFSCar, 2008.
- [12] NASCIMENTO, E. R. F. **Desenvolvimento de Modelos Neurais para o Processamento de Sinais Acústicos Visando a Medição de Propriedades Topológicas em Escoamentos Multifásicos.** 2007. Tese (Doutorado em Engenharia Mecânica) - Escola de Engenharia de São Carlos, Universidade de São Paulo, São Carlos, 2007.
- [13] NOORIZADEH, H.; FARMANY, A.; NARIMANI, H.; NOORIZADEH, M. **QSRR using evolved artificial neural network for 52 common pharmaceuticals and drugs of abuse in hair from UPLC–TOF-MS.** Drug Test. Analysis (2011).
- [14] ESFANDIARI A.; KALANTARI K. R.; BABAEI A. **Hair Loss Diagnosis using Artificial Neural Networks.** Int. J. Comp. Sci. 9 (2012) 174-180.
- [15] NASCIMENTO, D. S. C. **Obtenção de zona de conforto utilizando Redes Neurais Artificiais.** 2016. 153f. Tese (Doutorado). Escola de Minas, Universidade Federal de Ouro Preto, Ouro Preto, 2016.
- [16] BORGES, R. R. et al. **Sincronização de disparos em redes neuronais com plasticidade sináptica.** *Rev. Bras. Ensino Fís.* [online]. 2015, vol.37, n.2, pp.2310-1-2310-9. ISSN 1806-1117. <http://dx.doi.org/10.1590/S1806-11173721787>.
- [17] HAYKIN, S. **Neural Networks: a comprehensive foudation.** 1999. 2nd ed. New Jersey: Prentice Hall.

[18] SANTOS, M. C.; DAI, C.; PEREIRA, F. M. V. **Chemical Element Profiles in Commercial Woven Fabric Combining Laser-Induced Breakdown Spectroscopy and Chemometrics.** JOURNAL OF APPLIED SPECTROSCOPY, v. 85, p. 543-551, 2018.

[19] SANTOS, M. C.; PEREIRA, F. M. V. **Direct Analysis of Human Hair Before and After Cosmetic Modification Using a Recent Data Fusion Method.** JOURNAL OF THE BRAZILIAN CHEMICAL SOCIETY, v. 31, p. 33-39, 2020.