



APLICAÇÃO DE REDES NEURAIS ARTIFICIAIS COM ESPECTROSCOPIA NO INFRAVERMELHO PARA DISCRIMINAÇÃO DE MULHERES COM CÂNCER DE MAMA: UM PROCESSO DE TRIAGEM RÁPIDO, BARATO E EFICIENTE

4

RESUMO

INTRODUÇÃO: Detectar o surgimento de neoplasias malignas de modo precoce, eficiente e de baixo custo é a base para que uma nova tecnologia de triagem possa ser aplicada em escala e trazer benefícios aos pacientes e ao sistema de saúde. O uso da inteligência artificial com algoritmos de *deep learning*, tais como redes neurais artificiais (ANN), representam o mais alto patamar computacional da análise de dados de grande volume e padrões complexos. A espectroscopia no infravermelho com transformada de Fourier (FT-IR) é uma técnica aplicada em amostras biológicas para analisar o conjunto de informações de energia vibracional característicos das diferentes ligações químicas presentes nas moléculas da amostra. Os dados gerados pela FT-IR são tratados quimiometricamente para discriminar padrões de categorias com o uso de redes neurais artificiais, sendo que estas categorias podem ser condições de saúde ou doença. O câncer de mama é uma patologia que requer triagem que vai desde o autoexame, a mamografia e quando da presença de nódulos suspeitos a realização da biópsia, etapas que podem demorar, são invasivas e angustiantes para a paciente. Por isso, existe a necessidade de novas tecnologias eficientes, rápidas e baratas para triagem desta doença. **OBJETIVO:** Discriminar pacientes com câncer de mama através da análise do plasma por espectroscopia no infravermelho (FT-IR), seguido de tratamento quimiométrico e processamento com redes neurais artificiais. **METODOLOGIA:** 56 mulheres oriundas do Hospital Universitário da UFSM que realizaram mamografia e tiveram o diagnóstico de câncer de mama após biópsia com estágios I, II e III, independente do tipo histológico. O grupo controle foi composto de 18 mulheres normais. Para a análise do FT-IR, as alíquotas de plasma foram depositadas no cristal de leitura e desidratadas em corrente de ar (60-65°C) por 1,5 min. As leituras foram realizadas em triplicata em espectrômetro Spectrum™ 400 FT-IR/FT-NIR (PerkinElmer, Waltham, Massachusetts, EUA) no modo de refletância total atenuada (ATR-FTIR), na faixa de 650-4000 cm^{-1} , com resolução espectral de 4 cm^{-1} e 4 pulsos de varredura. Foi usado o algoritmo Kennard-Stone para dividir as amostras em 70% (controle = 13 e caso = 39) para construção do modelo de calibração e otimização por validação cruzada com o algoritmo k-fold venetion blinds, onde $k=15$. Os outros 30% (controle = 5 e caso = 13) foram utilizados para validação externa e obtenção das figuras de mérito. Para discriminação em caso ou controle, foi utilizado análise discriminante com redes neurais artificiais (ANNDA) com arquitetura de rede de perceptron multicamadas. Para a redução de dimensionalidade das variáveis de entrada optou-se pelo modelo de mínimos quadrados parciais (PLS) com 5 variáveis latentes e somente 1 neurônio na camada oculta intermediária. As configurações usadas na ANNDA foram algoritmo = backpropagation network, função de ativação = tangente hiperbólica, taxa de aprendizagem = 0.125 e ciclos

¹ Universidade de Santa Cruz do Sul (Unisc), nikolaspereira96@gmail.com

² Universidade de Santa Cruz do Sul (Unisc), valer@unisc.br

³ Hospital Universitário de Santa Maria (HUSM), drpadoim@gmail.com

⁴ Universidade de Santa Cruz do Sul (Unisc), nieger@unisc.br

de aprendizagem = 20. Os dados foram normalizados pela amplitude entre 0-1, transformados com 1ª derivada (janela de 5 pontos) e aplicado correção ortogonal de sinal. **RESULTADOS:** O modelo foi capaz de prever a classe de todas as amostras do conjunto teste, discriminando 100% em câncer ou não. Logo, a sensibilidade e especificidade foram iguais a 100%. **DISCUSSÃO:** Mesmo uma configuração de baixa complexidade adotada na ANNDA, utilizando somente 5 variáveis latentes de entrada e 1 único neurônio intermediário gerou um modelo com 100% de acerto com o uso do algoritmo backpropagation que recalcula os pesos e remolda o modelo a cada nova amostra introduzida. **CONCLUSÃO:** FT-IR com redes neurais artificiais demonstrou ser um excelente método de triagem para discriminar mulheres com e sem câncer de mama, podendo auxiliar no diagnóstico precoce proporcionando agilidade para encaminhamento de biópsia e tratamento.

PALAVRAS-CHAVE: Espectroscopia no infravermelho, redes neurais artificiais, câncer de mama, quimiometria