

Aprendizado de redes neurais profundas para diagnóstico molecular rápido e não invasivo de COVID-19 pela saliva

Anisio Pereira dos Santos Junior¹, Murillo Guimarães Carneiro¹
Robinson Sabino-Silva²

¹Faculdade de Computação (FACOM) – Universidade Federal de Uberlândia
Uberlândia – MG – Brasil

²Instituto de Ciências Biomédicas (ICBIM) – Universidade Federal de Uberlândia
Uberlândia – MG – Brasil

{anisio4322, mgcarneiro, robinsonsabino}@ufu.br

Abstract. *This work investigates a hybrid Deep Learning architecture (CNN-BiLSTM) to classify ATR-FTIR saliva spectra for COVID-19 detection. A crucial differential of this study is simulating a real-world pandemic triage scenario: it differentiates positive patients not from healthy individuals, but from flu-like symptomatic negative patients. The architecture captures subtle biomolecular alterations across infrared bands. Evaluated on 200 clinical samples, the approach achieved 80% average accuracy and F1-score, statistically outperforming traditional algorithms. The results validate the method's practical potential as a rapid, reagent-free screening tool for complex clinical settings with overlapping symptomatology.*

Resumo. *Este trabalho investiga uma arquitetura híbrida de Aprendizado Profundo (CNN-BiLSTM) para classificar espectros ATR-FTIR de saliva na detecção da COVID-19. Um diferencial crucial da pesquisa é simular um cenário real de triagem em pandemia: o modelo diferencia pacientes positivos não de pessoas saudáveis, mas de um grupo controle de sintomáticos gripais (negativos). A rede extrai alterações biomoleculares sutis no infravermelho perante forte similaridade clínica. Avaliada em 200 amostras, o modelo alcançou acurácia e F1-score médios de 80%, superando estatisticamente métodos tradicionais. Os achados validam o enorme potencial prático da abordagem como uma ferramenta de diagnóstico rápido e não invasivo.*

1. Introdução

A pandemia de COVID-19 e a emergência de novas variantes reforçam a necessidade contínua de métodos de testagem rápidos, acessíveis e sustentáveis [Frederick et al. 2024]. A espectroscopia vibracional, em particular a Espectroscopia de Infravermelho com Transformada de Fourier por Refletância Total Atenuada (ATR-FTIR), emerge como uma tecnologia promissora devido à sua rapidez, portabilidade, custo-benefício e natureza livre de reagentes [Naseer et al. 2021, Alkhuder et al. 2022]. A utilização da saliva como biofluido é particularmente atraente por permitir testes não invasivos, fáceis de coletar e propícios a testagens em larga escala [Caixeta et al. 2023].

O principal desafio na análise espectroscópica biomédica, no entanto, não é apenas diferenciar indivíduos doentes de saudáveis, mas sim de outros pacientes com sintomas

semelhantes (gripes comuns). Este estudo se depara com o desafio realista de classificar um grupo positivo contra um grupo de controle composto completamente por pacientes sintomáticos negativos. Nestes cenários, a diferenciação biomolecular capturada pelo infravermelho é muito sutil [Oliveira et al. 2023].

Para superar essas limitações, este artigo investiga uma arquitetura de aprendizado profundo (Deep Learning) híbrida que combina Redes Neurais Convolucionais (CNN) com Células de Memória Longa de Curto Prazo Bidirecionais (BiLSTM). A CNN é eficaz em extrair características espaciais (picos e vales morfológicos espectrais) [Taye 2023], enquanto a BiLSTM captura dependências sequenciais e de longo alcance dos números de onda do infravermelho [Fan et al. 2024].

Este artigo está organizado da seguinte forma: A Seção 2 descreve os dados, a arquitetura proposta e o protocolo experimental. A Seção 3 apresenta os resultados, análises estatísticas e interpretação. Por fim, a Seção 4 traz as considerações finais.

2. Materiais e Métodos

2.1. Aquisição e Pré-processamento de Dados

O conjunto de dados, coletado após aprovação ética, abrange 200 amostras de saliva de pacientes divididos em dois grupos de igual tamanho ($N=100$ por classe): pacientes diagnosticados (RT-PCR) positivamente para COVID-19, e um grupo de controle composto por pacientes negativos mas com sintomas gripais na triagem.

Espectros de ATR-FTIR das amostras de saliva operaram como as impressões digitais moleculares. Conforme as Figuras 1a e 1b, o perfil vibracional entre amostras COVID-19 e o controle sintomático sofre uma considerável sobreposição. Para mitigar o ruído experimental e variâncias de percurso óptico, todos os espectros passaram por uma filtragem de suavização via algoritmo Savitzky-Golay (SG) [Savitzky and Golay 1964] seguida de normalização vetorial pela Norma Euclidiana.

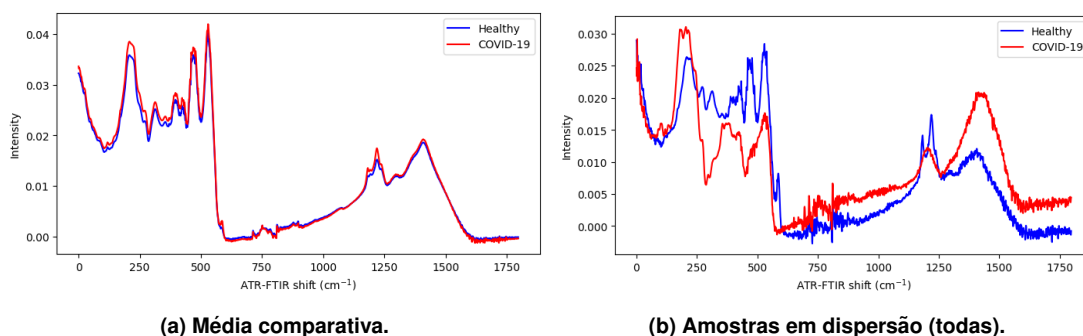


Figura 1. Espectros ATR-FTIR de saliva pré-processados, destacando a severa sobreposição espectral, dificultando a discriminação entre classes.

2.2. Arquitetura CNN-BiLSTM Proposta

O modelo de aprendizado profundo híbrido utiliza duas camadas Conv1D seguidas de MaxPooling1D. O mapeamento convolucional extrai características locais do array 1D representativo dos espectros de infravermelho (intensidades como sequências paramétricas

não temporais). Essa sequência reduzida e abstraída em componentes mais discriminativos flui para as células BiLSTM da rede. Ao processar dados espectrais bidirecionalmente, a BiLSTM estabelece dependências contextuais e detecta assinaturas macromoleculares complexas e distribuídas (ex. modos vibracionais simétricos e assimétricos coatuantes) não rastreáveis isoladamente [Fan et al. 2024, Educative.io 2024].

Camadas contíguas tipo Flatten, Dense, e Dropout mitigam overfittig [Srivastava et al. 2014] preparando os dados vetoriais para a saída sigmoidal de classificação binária. O fluxo lógico da arquitetura é esquematizado na Figura 2.

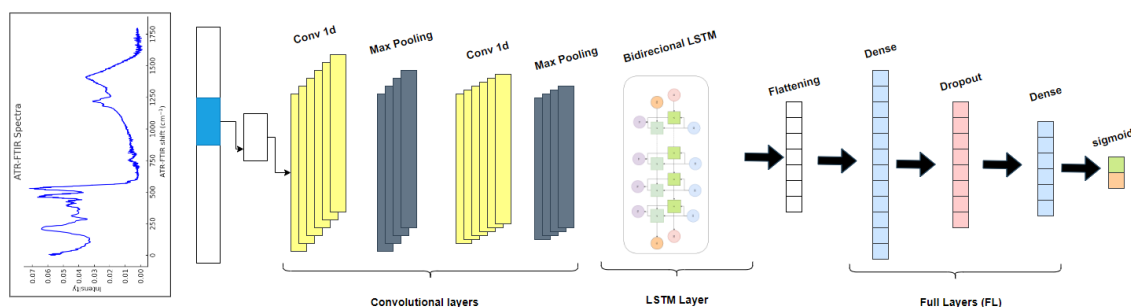


Figura 2. Estrutura do modelo híbrido de Machine Learning 1D CNN-BiLSTM empregado na análise de espectroscopia ATR-FTIR da saliva para detectar infecções de COVID-19.

2.3. Configuração Experimental e Avaliação

Para comprovar a validade clínica, implementou-se validação cruzada 10-fold estratificada (com testes distribuídos por 10 simulações independentes). O conjunto de desenvolvimentos sofreu divisão de 80%/20% para otimização de hiperparâmetros (como dropout, ativadores, unidades ocultas e filtros) pela abordagem TPE do *Optuna* [Akiba et al. 2019] por 200 iterações. Parâmetros otimizados treinaram os 90% do fold validados contra uma amostra cega de teste formadora dos 10% remanescentes.

Classificadores Baseline abrangidos pela mesma metodologia envolveram CNN unitária, Multi-Layer Perceptron (MLP), Random Forest (RF), Naive Bayes (NB), XGBoost (XGB), Support Vector Machine (SVM) e processadores sequenciais Transformer [Wen et al. 2022].

3. Resultados e Discussão

Os resultados médios ($\pm$ desvio padrão) extraídos das 100 avaliações de teste confirmam a performance diagnóstica superior do método acoplado CNN-BiLSTM frente a variados modelos comparativos. Observado na Tabela 1, a assertividade classificatória (Acurácia) estabilizou a patamares de 80%, liderando métricas críticas como Especificidade (77%) - fundamental por denotar reduzidas ocorrências de alarmes falsos vitais frente a confusões gripais - e um F1-Score generalista de 80%.

A arquitetura referendada evidenciou um crescimento de 4% contra as previsões unitárias do seu núcleo convolucional de extração espectral per si (a CNN unitária cravou 0.76 de acurácia com F1 de 0.77), sugerindo validade à hipótese da forte modelagem

Tabela 1. Resultados paramétricos preditivos do modelo em classificação ATR-FTIR, melhores índices grifados em negrito.

Modelo	Acurácia	Sensibilidade	Especificidade	Precisão	F1-Score
NB	0.50±0.107	0.72±0.162	0.27±0.164	0.50±0.076	0.58±0.162
RF	0.55±0.101	0.54±0.171	0.55±0.232	0.56±0.117	0.54±0.171
XGB	0.54±0.125	0.51±0.208	0.56±0.190	0.53±0.138	0.51±0.208
MLP	0.54±0.129	0.51±0.208	0.56±0.127	0.53±0.127	0.51±0.208
SVM	0.54±0.094	0.88±0.193	0.20±0.133	0.52±0.064	0.65±0.193
CNN	0.76±0.057	0.83±0.125	0.69±0.120	0.74±0.070	0.77±0.125
Transformer	0.67±0.111	0.65±0.143	0.68±0.215	0.70±0.163	0.66±0.104
CNN-BiLSTM	0.80±0.093	0.82±0.103	0.77±0.157	0.80±0.125	0.80±0.086

de dependências sequenciais propiciadas à amostra (através do tensor originado no layer Flatten). Pode-se observar que as redes neurais não profundas (Multi-layer Perceptron e algoritmos de Ensemble) atuaram com extrema inferioridade frente ao espaço caótico dos ruídos vibracionais por não processarem adequadamente as altas dimensionalidades subjacentes nas assinaturas. Ademais, o SVM demonstra clara instabilidade viés (Alta Sensibilidade, mas Especificidade subótima - classificando positivos com grande folga e ignorando sintomatologias da influenza).

3.1. Análises Estatísticas e Visualização t-SNE

A fim de se garantir que esta superioridade manifesta do agrupamento CNN-BiLSTM não é um efeito estocástico das partições nos K-folds, aplicou-se análises de Testes de Friedman e Nemenyi [Demšar 2006]. O Friedman indicou dissimilaridades severas por acurácia ($\chi^2 = 42.81$, $p < 0.001$) e Especificidade ($\chi^2 = 42.27$, $p < 0.001$). Nas comparações pareadas, o CNN-BiLSTM demonstrou diferenças estatísticas notórias com $p < 0.05$ (Testes *Post-Hoc* com Nemenyi) em relação a todos os baselines de *machine learning* estrito clássicos e redes não sequenciais, alicerçando a validação e sua robustez no suporte analítico biomolecular.

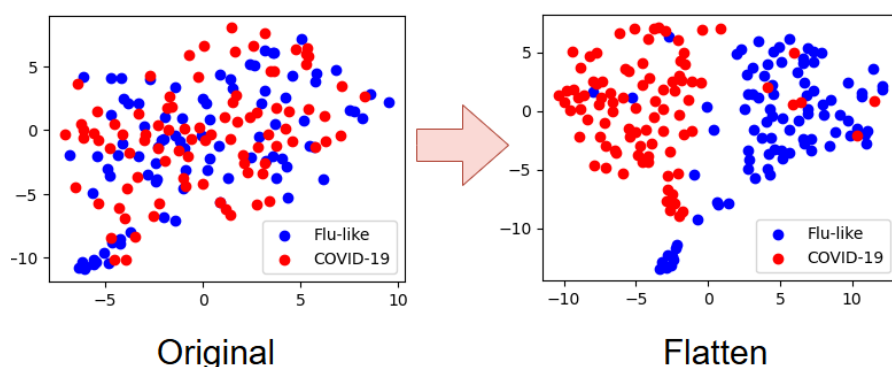


Figura 3. A clusterização por algoritmos t-SNE evidenciam na esquerda a ausência latente de diferenciais discerníveis de assinaturas nos Espectros FTIR originais entre grupos COVID_19 (vermelhos) e de sintomas Gripais genéricos (azuis). À direita, a hiperdimensionalidade processada pelo Flatten Layer das redes CNN-BiLSTM revela alta abstração e separabilidade notável dos grupos para classificação Sigmoid final.

Adicionalmente, mecanismos interpretativos atestaram a adequabilidade clínica. A visão abstracional redutora T-SNE disposta na Figura 3 corrobora de forma imagética que as sub-redes convolucionais atuaram eficientemente na codificação latente; reestruturando as dimensões espaciais para formar limites demarcatórios eficientes a fim de viabilizar o output preditivo das portas densas com muito êxito frente a representação originária confusa das assinaturas refletidas.

Para aprofundar a interpretabilidade biológica das predições, o método *SHAPley Additive exPlanations* (SHAP) foi aplicado para identificar quais bandas do infravermelho foram mais determinantes na classificação. O modelo CNN-BiLSTM demonstrou forte dependência de regiões espectrais associadas a lipídios, proteínas e ácidos nucleicos, conforme detalhado na Tabela 2. A alteração metabólica nestas macromoléculas (como nos modos vibracionais de cadeias acil e bandas Amida) na saliva está alinhada à literatura de patofisiologia viral. Isso reforça a convicção de que o modelo realmente aprendeu características biomoleculares válidas para distinguir a COVID-19 de outras síndromes gripais.

Tabela 2. Principais picos espectrais ATR-FTIR identificados pela análise SHAP e suas respectivas atribuições biomoleculares na amostra salivar.

Número de Onda (cm ⁻¹)	Atribuição Vibracional	Biomolécula
2922 e 2920	Estiramento CH ₂ assimétrico de cadeias acil	Lipídios
1744	Vibração do grupo éster de triglicerídeos	Lipídios
1696	Folha- β antiparalela (Amida I)	Proteínas
1674 e 1670	Estiramento C=O (Amida I)	Proteínas
1375 e 1374	Estiramento C-N de citosina	Ácidos nucleicos
1312	Componentes da banda Amida III	Proteínas

4. Conclusões

A pesquisa delineou uma viável alternativa de diagnóstico biomolecular rápido via exploração vibracional ATR-FTIR por aprendizado profundo contornante de desafios espectroscópicos complexos (como sobreposição latente interclasses sintomatológicas em pacientes da vida real num pronto-socorro). A associação hierárquica e bidirecional de Convoluções e Células LSTM para o trato dos tensores de biofluidos de saliva evidenciou, a nível estatístico formal, resultados animadores perante abordagens lineares para triagem clínica sustentável de alta velocidade não-invasiva.

Impacto e Reconhecimento

As contribuições desta linha de pesquisa foram validadas externamente por reconhecimentos e produções científicas que reforçam sua originalidade e potencial de aplicação. Uma versão preliminar da abordagem, voltada à análise de espectros Raman de soro sanguíneo via CNN, foi publicada nos anais da *Brazilian Conference on Intelligent Systems* [Santos Jr et al. 2023], validando a estratégia metodológica por revisão por pares. O caráter inovador do projeto também foi reconhecido em duas edições consecutivas do *Google Latin America Research Awards* (LARA), nas edições de 2020 e 2021, com repercussão na imprensa institucional¹. Adicionalmente, a expertise consolidada culminou no

¹Disponível em: <https://comunica.ufu.br/noticias/2022/02/pesquisadores-da-ufu-ganham-premio-do-google-pela-segunda-vez>

depósito da patente internacional “*Spectral profile for diagnosing covid-19, use thereof, method, system and platform for diagnosing Covid-19*”, em colaboração com a iniciativa privada, evidenciando a viabilidade de translação dos avanços aqui apresentados para aplicações diagnósticas reais.

Agradecimentos

Os autores agradecem o apoio financeiro concedido pelo Google (por meio dos prêmios Google Latin America Research Awards de 2020 e 2021), pela Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) (processo nº APQ-00410-21), pelo Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (processos nº 402196/2021-0 e 408216/2022-0 e 420212/2023-0) e pelo Instituto Nacional de Ciência e Tecnologia em Teranóstica e Nanobiotecnologia (INCT-Teranano) (processo nº CNPq-465669/2014-0). Agradecemos ainda à NVIDIA Corporation pela doação de uma GPU Titan V utilizada nesta pesquisa.

Referências

- Akiba, T. et al. (2019). Optuna: A next-generation hyperparameter optimization framework. *Proceedings of the 25th ACM SIGKDD*, pages 2623–2631.
- Alkhuder, K. et al. (2022). Attenuated total reflection Fourier transform infrared (ATR-FTIR) spectroscopy of saliva for the detection of COVID-19. *Clinical Spectroscopy*, 4:100018.
- Caixeta, D. C. et al. (2023). Salivary ATR-FTIR spectroscopy for COVID-19 diagnosis: A systematic review. *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, 286:121950.
- Demšar, J. (2006). Statistical comparisons of classifiers over multiple data sets. *Journal of Machine learning research*, 7(Jan):1–30.
- Educative.io (2024). Introduction to convolutional neural networks (CNN). Accessed: 2024-07-01.
- Fan, Y., Tang, Q., Guo, Y., and Wei, Y. (2024). BiLSTM-MLAM: A Multi-Scale Time Series Prediction Model for Sensor Data Based on Bi-LSTM and Local Attention Mechanisms. *Sensors*, 24(12):3962.
- Frederick, J. et al. (2024). Biosensor integration with smartphone applications for clinical diagnosis: A review. *Sensors*, 24(1):100.
- Naseer, K. et al. (2021). ATR-FTIR spectroscopy: A facile and robust tool for diagnosing viral diseases. *Photodiagnosis and Photodynamic Therapy*, 34:102220.
- Oliveira, L. C. et al. (2023). Salivary vibrational spectroscopy for COVID-19 screening. *Diagnostics*, 13(8):1443.
- Santos Jr, A. P., Filho, A. C. M., Sabino-Silva, R., and Carneiro, M. G. (2023). Convolutional Neural Networks for the Molecular Detection of COVID-19. In *Brazilian Conference on Intelligent Systems*, pages 51–62. Springer.
- Savitzky, A. and Golay, M. J. (1964). Smoothing and differentiation of data by simplified least squares procedures. *Analytical chemistry*, 36(8):1627–1639.

- Srivastava, N., Hinton, G., Krizhevsky, A., Sutskever, I., and Salakhutdinov, R. (2014). Dropout: a simple way to prevent neural networks from overfitting. *The journal of machine learning research*, 15(1):1929–1958.
- Taye, M. M. (2023). Theoretical understanding of convolutional neural network: Concepts, architectures, applications, future directions. *Computation*, 11(3):52.
- Wen, Q., Zhou, T., Zhang, C., Chen, W., Ma, Z., Yan, J., and Sun, L. (2022). Transformers in time series: A survey. *arXiv preprint arXiv:2202.07125*.