



## Uso de redes neurais para classificação da resposta cerebral medida com espectroscopia no infravermelho próximo.

Nicolas Gabriel S. R. de Souza\*, Sérgio L. Novi, Rickson C. Mesquita  
Instituto de Física "Gleb Wataghin", Universidade Estadual de Campinas (UNICAMP)

### Introdução

As técnicas de neuroimagem são ferramentas importantes para o desenvolvimento da neurociência, possuindo aplicações médicas tanto no diagnóstico de doenças quanto nos processos de tratamento delas. Nas últimas décadas, a técnica de espectroscopia no infravermelho próximo (NIRS, do inglês, *near-infrared spectroscopy*) tem se mostrado promissora por conta da sua capacidade de medir alterações na hemodinâmica do córtex cerebral de forma não invasiva [1].

A técnica de NIRS utiliza luz na região do infravermelho próximo (700-900 nm) de baixa potência (~10 mW). Como nessas condições os fótons conseguem atravessar o crânio e penetrar até à superfície do córtex, pode-se usar a NIRS para medir a ativação neural resultante de uma tarefa externa a partir das alterações nas concentrações de oxí-hemoglobina (HbO) e deoxi-hemoglobina (HbR) decorrentes da tarefa.

No entanto, as variações de HbO e HbR decorrentes da ativação funcional do cérebro são muito pequenas, da ordem de 3% em relação ao estado basal, o que torna sua detecção muito difícil. Alguns trabalhos recentes têm mostrado o potencial de redes neurais artificiais para a análise de séries temporais dos dados coletados com NIRS [2], especialmente para a detecção da atividade neural associada a uma tarefa externa. Desse modo, o objetivo deste projeto foi avaliar o uso de redes neurais artificiais para classificação da atividade cerebral a partir do sinal da NIRS em seres humanos.

### Materiais e Métodos

Para análise dos algoritmos, foram utilizados dados de NIRS previamente coletados de 18 voluntários sadios durante o repouso (*resting state*). Os dados foram coletados com um sistema comercial de NIRS de ondas contínuas (NIRScout, NIRx Medical Systems) contendo 32 fontes (cada fonte tem 2 comprimentos de onda, 760 e 840 nm) e 32 detectores de luz. O sensor óptico usava todas as fontes e detectores posicionados em 64 combinações (pares fonte-detector) diferentes, de forma a cobrir os lobos frontal e parietal dos dois hemisférios. Após as coletas, as intensidades de luz detectadas em cada par fonte-detector (canal) foram convertidas em concentrações de HbO e HbR através da lei de Beer-Lambert modificada.

Para imitar um experimento funcional, uma resposta hemodinâmica simulada foi adicionada às séries temporais de cada canal de HbO e HbR dos voluntários em 9 instantes de tempo pré-determinados (Figura 1). A amplitude da resposta simulada foi variada para estudar a performance das redes neurais em detectar ativações funcionais em diferentes razões sinal-ruído.

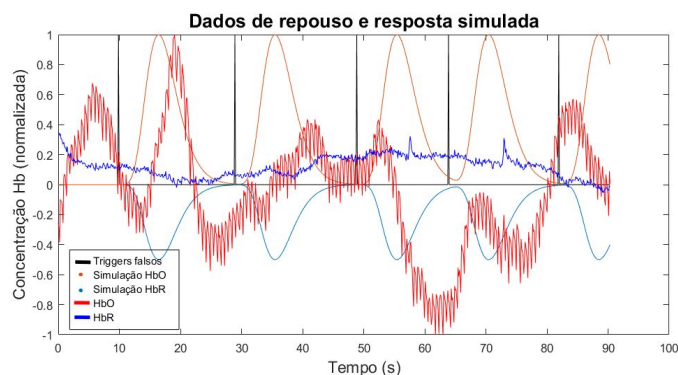


Figura 1: Exemplo de uma série temporal de HbO (vermelho) e HbR (azul) coletada durante o estado basal, e da resposta de atividade neural simulada. As barras verticais em preto representam os instantes em que a simulação foi incluída.

Após esse processo, os dados em torno da atividade simulada foram segmentados em blocos de tarefa e repouso (Figura 2). A duração de cada bloco variou entre 2 e 6 s, com intervalos de 2 s, a fim de estudar como a quantidade de informação influencia na classificação. Ao todo, cada voluntário contribuiu com até 1.152 amostras para análise (equivalente a 64 canais, tendo cada canal 9 simulações, e cada simulação tendo dois blocos). Considerando os 18 voluntários e descartando canais com baixa razão sinal-ruído, as análises posteriores utilizaram um total de 20.304 amostras. Cada amostra foi representada por uma matriz de 2 dimensões, contendo informação tanto de HbO quanto de HbR numa dimensão, e a duração do bloco na segunda dimensão. Parte destas amostras foram utilizadas como *input* da rede neural artificial.

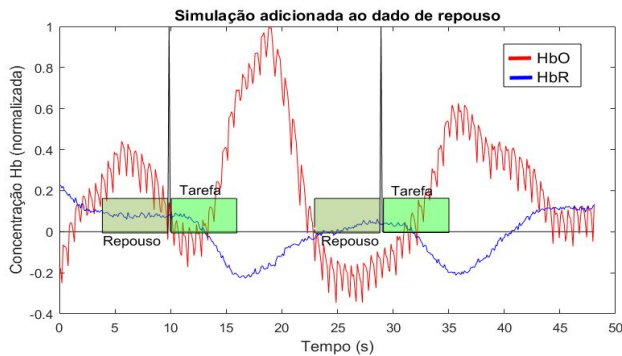


Figura 2: Série temporal de HbO (vermelho) e HbR (azul) após a inserção da atividade neural simulada. A figura mostra um recorte dos primeiros dois triggers da figura 6.

Foram escolhidas duas redes neurais, sendo uma do tipo "perceptron multicamadas" (*multilayer perceptron - MLP*) e outra do tipo rede neural convolucional (CNN). A *MLP* é formada por estruturas denominadas neurônios, cuja função é armazenar um número. Eles são organizados em camadas e conectados entre si como um grafo. Cada camada possui uma quantidade livre de neurônios, que é o que vai caracterizar a arquitetura da rede. Para cada ligação entre neurônios é feita operações matemáticas e o resultado é o argumento da função de ativação que define se o neurônio será ativado ou não. Existem vários exemplos de funções de ativação, como *sigmóide*, *ReLU*, *Leaky-ReLu* entre outras. Neste trabalho, a rede MLP foi construída com uma camada *Flatten*, que recebe os dados de input na forma de uma matriz e transforma em um vetor. Em seguida, tem-se uma camada *Dense* composta por 80 neurônios utilizando a função de ativação *ReLU*. Por fim, a última camada de classificação foi composta por 2 neurônios com ativação *Softmax*, que devolve a probabilidade do dado ser ativado ou não ativado. Um esquema da rede pode ser visto na Figura 3.

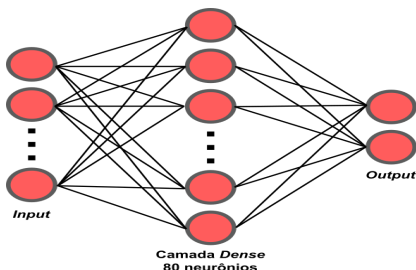


Figura 3: Arquitetura da rede neural MLP.

A rede CNN segue a mesma lógica da rede MLP. Contudo, possuem um maior número de camadas e outras propriedades. *A priori*, temos uma matriz de input em que é feita uma convolução desse dado com uma submatriz, chamada de *kernel*. Desse modo, cada kernel utilizado gera uma outra matriz. O conjunto das matrizes geradas é conhecido como *feature maps*. Além disso, pode ser aplicado uma camada de condensamento (*pooling layer*) em que é feita uma subamostragem da matriz original com objetivo de simplificar a informação, reduzindo o tamanho do dado. Neste trabalho, a primeira camada da CNN recebeu um *input 2D*, no qual foi feita uma convolução em uma dimensão (*Conv1D*). Em seguida, tem-se mais duas camadas de *Conv1D* e uma camada de *Maxpooling1D*. Ademais, para vetorizar esses dados é colocada uma camada *Flatten*, seguida de duas camadas *Dense* com ativação *ReLU*. Finalmente, uma camada de *Dropout* (0.5) é aplicada seguida da última camada *Dense* com *Softmax* e dois neurônios para classificação. Uma representação dessa rede se encontra na Figura 4.

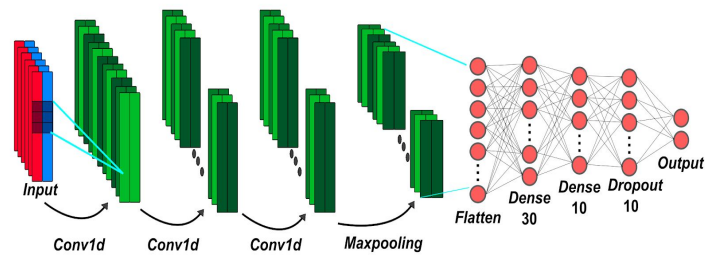


Figura 4: Arquitetura da rede neural CNN.

## Resultados e Discussão

Inicialmente, todas as amostras foram embaralhadas de forma igual, sem distinção entre os voluntários. Separamos 20% dos dados para testar a rede após esta ser treinada. Os outros 80% dos dados foram utilizados para treinar a rede., e parte destes dados foram utilizados para validação da rede antes do teste.

Obtivemos uma acurácia muito alta (~99%) com as duas redes, e em todas as condições de duração dos blocos e amplitudes da resposta simulada, o que nos fez levantar a hipótese de que havia *overfitting*. E mesmo utilizando apenas 20% dos dados para treino e 80% dos dados para teste, chegamos em um resultado parecido ao mostrado acima. Logo, a rede, de fato, parece ter aprendido algum padrão inerente aos dados.

Com objetivo de verificar se esse padrão era, de fato, proveniente da função de ativação que usamos para simular a resposta neural, passamos a utilizar apenas os dados brutos de repouso. Nessa situação, não incluímos nenhuma função de ativação nos dados, embora continuássemos a marcar os blocos como "ativação". Após os mesmos procedimentos de treino e teste supracitados, a rede continuou tendo uma acurácia superior a 99%.

Num teste seguinte, reduzimos a quantidade de canais utilizados, selecionando apenas um canal por voluntário. A ideia por trás deste teste é que canais que estão medindo regiões próximas do cérebro possuem uma correlação (espacial) maior, e a rede poderia estar aprendendo esta informação. Contudo, mesmo nestas condições, a acurácia ainda se mostrou alta, em torno de 99%.

Isso sugeriu que a rede aprendeu algum padrão de dependência temporal inerente à estrutura dos dados provenientes de NIRS. De fato, a Figura 5 mostra a autocorrelação de um canal aleatório de um dos voluntários, e cujo padrão se repete em outros canais. Podemos ver que a correlação é bem alta até um certo ponto da série temporal, ou seja, com uma quantidade de pontos iniciais é possível prever o comportamento dos pontos futuros.

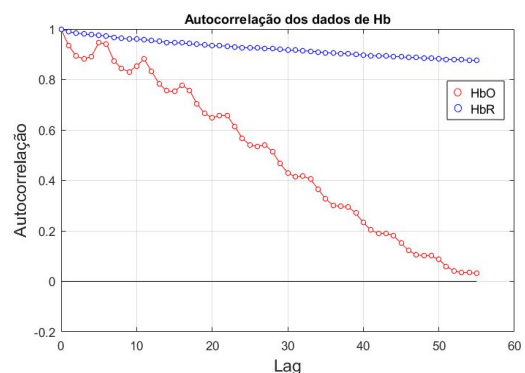


Figura 5: Exemplo de gráfico da autocorrelação da série temporal para HbO e HbR.

Com objetivo de testar nossa hipótese e remover a correlação temporal com aleatoriedade, mudamos a forma como as amostras de tarefa e repouso são categorizadas. Antes, para cada amostra de repouso tinha uma amostra de ativação imediatamente após a de repouso. Passamos então a cortar, os dados de cada voluntário em blocos de 6s, e para cada bloco foi escolhido um *label* aleatório (50%) para ativação ou repouso. Seguindo esse princípio foram feitas duas abordagens.

Na primeira abordagem, quando uma escolha de *label* é feita, ela é aplicada em todos os canais e para todos voluntários. É perceptível que não temos mais uma correlação temporal, dado que cada bloco é classificado de forma aleatória. Contudo, como os labels são colocados em todos os canais, mantivemos a correlação espacial dos canais. Primeiramente inserimos os dados de todos os canais, usando 80% dos dados para treino e o restante para teste. Com isso, diminuimos a acurácia da classificação para 70% no caso da rede MLP e 81% para a CNN.

Podemos ver que utilizando uma classificação de blocos aleatórias conseguimos reduzir um pouco da correlação temporal, o que dificulta o aprendizado da rede. Ademais, utilizando apenas 5 canais de cada voluntário, chegamos na acurácia de 61% com a rede MLP e 57% com a CNN. Por fim, utilizando apenas um canal de cada voluntário, obtivemos 52% para as duas redes.

Dessa forma, podemos ver que, dependendo do caso, pode ser possível treinar uma rede utilizando informação de apenas um canal. Além disso, vimos que a correlação temporal inerente ao dado NIRS tem um peso elevado, porque mesmo supondo um caso irreal, em que o repouso não é seguido da tarefa, a rede consegue aprender um padrão em dados de repouso classificados de forma aleatória.

Visando aleatorizar mais os *labels* e reduzir a correlação espacial, a classificação do bloco foi feita de forma aleatória, individualizada para todos os canais e voluntários. Testando a rede usando dados de todos os canais chegamos na acurácia de 50% para ambas as redes. E mesmo reduzindo a quantidade de canais, mantivemos esse mesmo valor, que é o que se espera de um chute. Logo, com essa abordagem apenas confirmamos que a correlação espacial pode ter um peso relevante nos dados de NIRS, pelo fato de que conseguimos reduzir de 20% a 30% a acurácia da rede. Isso pode nos indicar que fornecer dados de muitos canais pode apenas servir como ruído para a classificação dos dados NIRS pelas redes neurais.

## Conclusões

Os testes realizados ao longo do projeto permitiram ter acesso a informações que não são muito bem documentadas na literatura e mostraram novos desafios. O desafio principal mostra claramente que, a fim de se classificar a atividade neural em dados de neuroimagem, é necessário encontrar uma forma de reduzir a correlação temporal nos dados. Esta conclusão impacta diretamente técnicas com alta resolução temporal, como NIRS ou EEG. Somente após este passo podemos retornar às simulações realizadas e estudar novamente o comportamento das redes. Outro ponto importante observado nos experimentos foi analisar como a informação espacial (refletida na quantidade de canais) interfere na acurácia da rede. Isso pode ser um indicativo de que nem sempre fornecer dados de muitos canais será

melhor, considerando que o objetivo é classificar a resposta de uma única tentativa (*single trial*) de pelo menos um canal. Os resultados deste estudo têm impactos significativos no uso de NIRS para aplicações em interfaces cérebro-máquina.

## Agradecimentos

Agradeço ao CNPq pela concessão da bolsa de iniciação científica. Agradeço ao Professor Rickson Mesquita pela orientação, suporte e paciência. Agradecimento especial ao Sérgio Novi pela motivação e por ter me ajudado ao longo desse projeto.

1. Huppert TJ, Diamond SG, Franceschini MA, Boas DA. HomER: a review of time-series analysis methods for near-infrared spectroscopy of the brain. *Applied Optics*. 2009. p. D280. doi:10.1364/ao.48.00d280
2. Trakoolwilaiwan T, Behboodi B, Lee J, Kim K, Choi J-W. Convolutional neural network for high-accuracy functional near-infrared spectroscopy in a brain-computer interface: three-class classification of rest, right-, and left-hand motor execution. *Neurophotonics*. 2017. p. 1. doi:10.1117/1.nph.5.1.011008
3. Barker JW, Aarabi A, Huppert TJ. Autoregressive model based algorithm for correcting motion and serially correlated errors in fNIRS. *Biomed Opt Express*. 2013;4: 1366–1379.