



1 Introdução

Esclerose lateral amiotrófica (ELA), síndrome do encarceramento e lesões na coluna vertebral são condições que emprisionam pessoas dentro de seus próprios corpos. Nos casos mais graves, os olhos são a última janela que permite a comunicação com o mundo. É pensando nessas pessoas que o projeto foi desenvolvido. O projeto tem o objetivo de estudar uma Interface Cérebro-Computador (BCI do inglês Brain Computer Interface) para permitir o controle de uma cadeira de rodas a partir dos sinais cerebrais de encefalografia, devolvendo assim, algum grau de liberdade para os pacientes. Parafraseando o ponto de vista de Hugh Herr, chefe do grupo de biomecatrônica do MIT MediaLab: um corpo nunca pode ser “quebrado” (‘broken’ no original). O ambiente por nós construído e a tecnologia estão “quebradas” e têm limitações. Nós, como pessoas, não podemos aceitar essas limitações como definitivas, mas precisamos buscar avanços tecnológicos.[Herr, 2014].

2 Objetivos e Metodologia

2.1 SSVEP

O método utilizado no projeto é a identificação de potenciais estacionários evocados por estimulação visual repetitiva (SSVEP do inglês Steady State Visually Evoked Potentials). Um paciente é apresentado a um monitor que contém 4 imagens de tabuleiros de xadrez. Cada uma dessas imagens alterna suas casas brancas com as pretas nas frequências de 11, 12, 13 e 14Hz. Quando o paciente se concentra em apenas um tabuleiro de xadrez, os neurônios da região occipital do cérebro, responsável pelo processamento de imagens, são ativados na mesma frequência em que o tabuleiro inverte a cor de suas casas. Esses neurônios produzem campos elétricos pequenos, mas que podem ser medidos por eletrodos posicionados na região occipital (parte de trás da cabeça). Após a etapa de processamento desses sinais, uma rede neural é capaz de identificar em qual tabuleiro de xadrez o paciente estava olhando. Como cada tabuleiro de xadrez está associado com um comando para a cadeira de rodas (frente, esquerda, direita e freio), um paciente, com apenas o controle dos olhos, pode se movimentar sem a interferência de terceiros.

2.2 Coleta de dados

A base de dados coletada para o projeto contém sinais de EEG de dois adultos voluntários e saudáveis (a coleta foi realizada no contexto de projeto aprovado pelo comitê de ética da UNICAMP - CAAE 0617.0.146.000-10). Foi utilizada uma touca no padrão 10-20 estendido, com a referência nos mastóides e os eletrodos em F9, F10 e FPZ foram utilizados para o cancelamento de ruído. Os testes foram realizados com 5 classes balanceadas: 4 imagens de tabuleiros de xadrez apareciam na tela de um computador durante os testes e alternavam suas casas brancas e pretas em uma frequência de 10, 11, 12 e 13Hz sendo cada frequência atribuída a uma classe. A última classe foi atribuída ao caso em que o voluntário não se concentrava em nenhum dos tabuleiros de xadrez. Cada voluntário passa por uma bateria de 60 testes, 12 para cada classe, com cada teste tendo 6 segundos. Os testes foram feitos com uma taxa de amostragem de 256Hz. Os dados coletados de um dos voluntários, no entanto, se mostraram extremamente ruidosos, e tiveram de ser desconsiderados para o treinamento da rede. Com a pandemia, o laboratório ficou indisponível, impossibilitando a coleta de mais dados. Este problema limitou os resultados e ajustes tiveram de ser feitos para trabalhar com um conjunto de dados reduzido. O número de classes foi diminuído de 5 para 2 (10Hz e 13Hz): por esse motivo, portanto, foram analisados um total de 24 testes neste projeto.

2.3 Processamento de dados

Primeiro é necessário discretizar e amplificar o sinal vindo do eletrodo, que é originalmente da ordem de microvolts, para uma amplitude mais fácil de ser lida pelo computador. Depois o sinal passa por um filtro para eliminar a interferência proveniente da rede elétrica em 60 Hz. Após esta etapa, os dados são discretizados e passados para o computador. Embora tenham sido coletados informação de 16 eletrodos, nem todos possuem a mesma porcentagem de informação relevante para a classificação. Eletrodos mais próximos da região occipital (Oz, O1 e O2), por exemplo, são os mais relevantes. Com base neste estudo [Sarah N.Carvalho, 2015] e em testes feitos com a base de dados coletada nesse projeto, foi selecionado um subconjunto de 6 canais (Oz,O1,O2,POz,PO3,PO4) para a análise. A nova matriz de dados tem o formato: $24 \times 6 \times 1536$. Movimentos nos músculos faciais como o franzir da testa e o piscar do olhos não podem ser evitados durante o uso prolongado, mas podem poluir os campos elétricos medidos pelos eletrodos da região occipital, o que dificulta a classificação. Um filtro LMS (do inglês Least Mean Square) (figura 6) foi eleito para o cancelamento desses ruídos. O filtro LMS tem o objetivo de aproximar o sinal mensurado nos músculos da face (figuras 1 e 2) dos artefatos presentes no sinal de EEG (figura 3) para que se possa subtrair a saída do filtro (figura 4) do sinal de EEG, cancelando assim os artefatos (figura 5). Estão ilustrados abaixo (figuras 1 a 5) os sinais envolvidos na filtragem. O eixo vertical apresenta o potencial elétrico, já normalizado, medido pelo eletrodo indicado no título da respectiva imagem. O eixo horizontal representa a passagem do tempo. Cada 256 pontos equivale a 1 segundo.

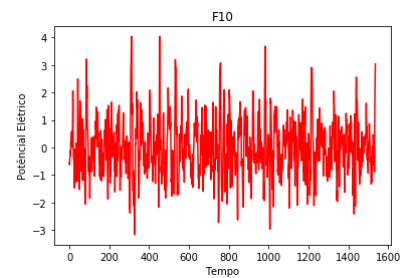
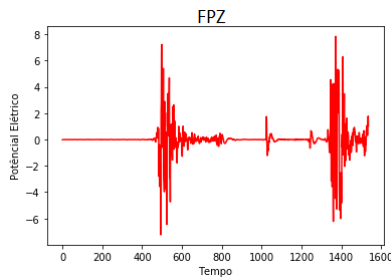


Figura 1: Eletrodo FPZ, posicionado na musculatura da testa
Figura 2: Eletrodo F10 posicionado na musculatura responsável pelo olho direito

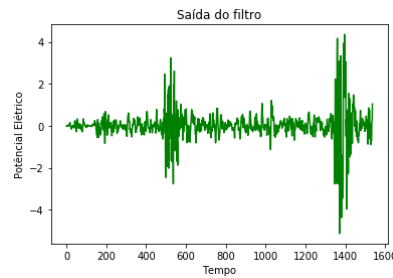
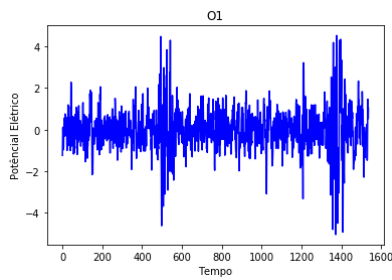


Figura 3: Eletrodo posicionado na região occipital responsável pelo processamento dos estímulos visuais.
Figura 4: Na saída do filtro LMS foi capaz de devolver os mesmos padrões do cérebro, responsável pelo processamento dos estímulos visuais em FPZ e F10 (principalmente em FPZ nesse caso) que estão presentes em O1, porém na escala que são observados em O1

Rede Convolucional

Uma arquitetura de rede neural eficiente para esse problema é a convolucional. Com esse método temos vetores de pesos adaptáveis, chamados de kernels, que realizam um produto escalar com uma parte da sequência de dados para obter uma saída. Em seguida, se utiliza o mesmo kernel em uma parte vizinha dos dados para se obter uma outra saída. Espera-se que cada kernel de tamanho K , após o treinamento da rede, seja responsável por reconhecer algum padrão, independente de onde se encontra. A arquitetura da rede convolucional inspirada por [Kwak et al., 2017] e adaptada para este projeto funciona da seguinte forma (figura 9):

Como já foi discutido na seção de processamento, os dados que chegam na rede estão no domínio da frequência, com 6 canais. Cada janela é uma sequência de N pontos dados que englobam as frequências nas faixas de 9-14Hz para $1 \leq K \leq 3$, 8-15Hz para $4 \leq K \leq 7$ ou 7-16Hz para $8 \leq K \leq 10$. A primeira camada é convolucional e

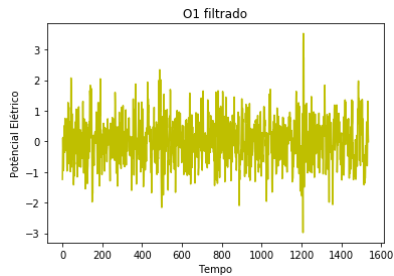


Figura 5: A saída do filtro é subtraída do sinal de O1 para se obter uma versão do sinal de O1 sem os artefatos

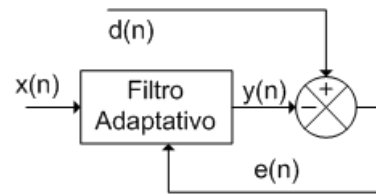


Figura 6: Esquemático geral de um filtro adaptativo. $d(n)$ se refere ao sinal de EEG (figura 3), $x(n)$ se refere aos sinais ruidosos (figuras 1 e 2), $y(n)$ se refere à saída do filtro (figura 5) e $e(n)$ se refere ao erro do filtro.

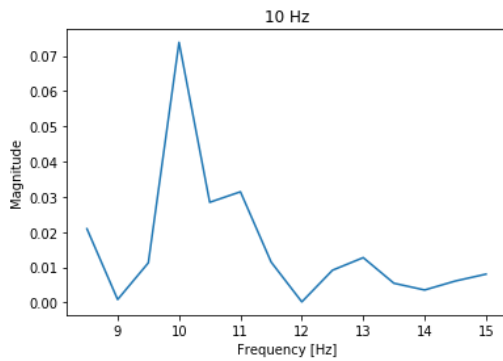


Figura 7: Exemplo de VEP (Visually Evoked Potencial) no canal OZ em 10Hz

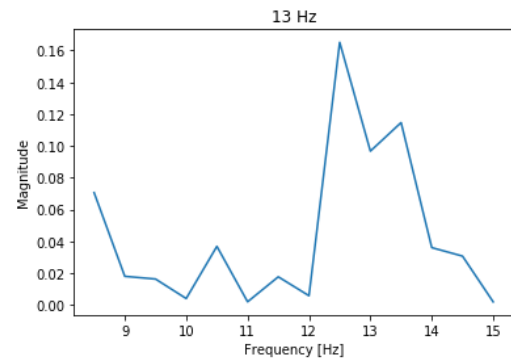


Figura 8: Exemplo de VEP no canal OZ em 13Hz

tem 6 kernels com dimensão $K \times 1$. observa-se na saída uma matriz com $6 \times (N-K)$. Espera-se que os kernels sejam responsáveis por identificar VEPs, ou seja, picos nas frequências de interesse. A Segunda e última camada é formada por dois perceptrons que tem seus pesos conectados a cada um dos pontos produzidos na etapa anterior. Ela é responsável por gerar duas saídas, cada uma representa a confiança que a rede tem em associar aquela janela com sua respectiva classe. O parâmetro mais importante da rede é o K . Ele define o tamanho do kernel responsável por identificar os VEPs. Diferentes tamanhos de K levam a rede a diferentes interpretações de um VEP. 5000 loops de treinamento foram realizados para valores de K entre 1 e 10, então, as redes com acurácia maior que 75 % foram selecionadas para integrar um comitê composto X redes com diferentes interpretações de um VEP. Cada rede tem direito em um voto, a classe com mais votos é a escolhida. Como as redes tem diferentes interpretações do que é um VEP o comitê se mostrou bem diverso e com boa generalização, mesmo que, individualmente, cada um dos integrantes não obteve um resultado satisfatório.

3 Resultados

A rede convolucional CNN1 por [Kwak et al., 2017] foi uma primeira versão de arquitetura proposta pelo projeto. A rede CNN1, obteve acurácia de 99,28% em um ambiente estático sem nenhum método de regularização. O volume de dados em que esta rede foi treinada ($N \leq 67,500$) é bem superior a base de dados disponível para o treinamento ($N = 20$), portanto, o projeto teve de aplicar uma série de estratégias de regularização para obter resultados. Foi utilizado uma regularização L2 com $\lambda = 0.35$. O tamanho dos lotes fui ajustado para 1024. A taxa de aprendizado segue o cronograma proposto em [Smith and Topin, 2017] e o otimizador utilizado foi o ADAM [Diederik P. Kingma, 2015]. A média das acurácias obtidas no conjunto de validação (25% dos dados) não se mostrou muito promissora, porem, 432 redes apresentaram resultados com acurácia maior que 75%. Um conjunto com as 10 melhores redes foram escolhidas para compor um comitê. Na tabela abaixo, temos as acurácias obtidas.

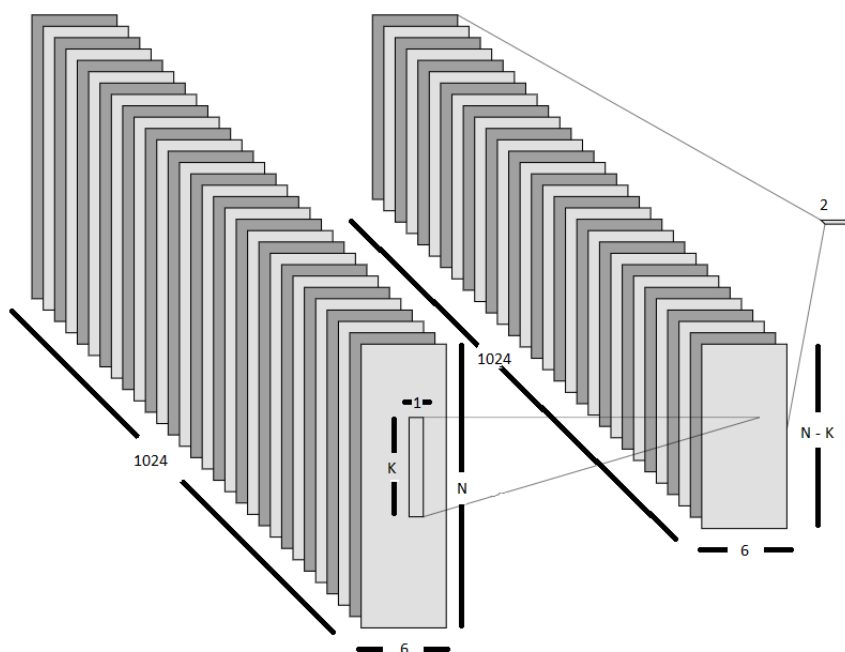


Figura 9: Arquitetura da rede neural convolucional inspirada por [Kwak et al., 2017] e adaptada para este projeto

Base de Dados	Acurácia
Subconjunto A (Os 20 testes com menos ruído do voluntário 1)	95,04
Todos os dados do voluntário 1	65,35
Todos os dados do voluntário 2	47,17
Todos os dados	56,26
Subconjunto A sem qualquer cancelamento de ruído	87,57

4 Conclusão

Neste projeto, foi realizado um estudo sobre um conjunto de temas ligados às temáticas de processamento de sinais e aprendizado de máquina em interfaces cérebro-computador. O aluno realizou um estudo que abrangeu, além de interfaces baseadas em SSVEP, temas como filtragem adaptativa, redes neurais e cancelamento de interferência. Foi proposto um esquema de filtragem do sinal de EEG para classificação por redes neurais profundas, o qual levou a resultados relevantes. Houve, todavia, limitações que, provavelmente têm a ver com a qualidade dos dados empregados, que corresponderam a um ensaio experimental de dimensões modestas. Pretendia-se coletar mais dados, mas a pandemia e o isolamento social impossibilitaram o desenvolvimento esperado para a parte experimental. Buscou-se, assim, extrair o máximo dos dados disponíveis.

Consideramos que a iniciação foi bem-sucedida, pois permitiu que o aluno adquirisse formação em vários temas de vanguarda no contexto do tratamento da informação, bem como em interfaces cérebro-computador. O esquema de cancelamento e classificação proposto requererá futuros estudos experimentais mais completos para ser validado, mas os resultados aqui relatados podem ser considerados promissores.

5 apoio e agradecimentos

Gostaria de agradecer ao meu orientador Romis Attux pela sua dedicação e sabedoria. Aos meus pais por terem acreditado no meu potencial e me incentivado durante todo o processo. Também sou grato ao aluno de Doutorando Thiago Bulhões pela sua ajuda na etapa de coleta de dados. Ao CNPQ pelo financiamento, e à UNICAMP por toda a estrutura e laboratórios que possibilitaram a minha pesquisa. Agradeço a todos os professores que formaram o meu conhecimento, em especial Jackson Neo Padilha cuja didática e paixão pela física me inspirou a ser o cientista que sou hoje. A todos os outros e outras cientistas que citei nas referências, gostaria de deixar aqui também minha gratidão.

Referências

- [Diederik P. Kingma, 2015] Diederik P. Kingma, J. B. (2015). Adam: A method for stochastic optimization.
- [Herr, 2014] Herr, H. (2014). Hugh herr: A nova biônica que nos permite correr, escalar e dançar. TED Talk.
- [Kwak et al., 2017] Kwak, N.-S., Müller, K.-R., and Lee, S.-W. (2017). A convolutional neural network for steady state visual evoked potential classification under ambulatory environment. *PLOS ONE*, 12:e0172578.
- [Sarah N.Carvalho, 2015] Sarah N.Carvalho, Thiago B.S.Costa, L. F. D. C. G. F. L. C. R. A. (2015). Comparative analysis of strategies for feature extraction and classification in ssvep bcis. *Elsevier*, pages 40–42.
- [Smith and Topin, 2017] Smith, L. N. and Topin, N. (2017). Super-convergence: Very fast training of residual networks using large learning rates. *CoRR*, abs/1708.07120.