



Universidade Federal de Ouro Preto
Instituto de Ciências Exatas e Aplicadas
Departamento de Engenharia Elétrica



Trabalho de Conclusão de Curso

Aplicação de técnicas de *deep learning* em sinais cerebrais

Larissa Rangel de Azevedo

João Monlevade, MG

2022

Larissa Rangel de Azevedo

**Aplicação de técnicas de *deep learning* em
sinais cerebrais**

Trabalho de Conclusão de Curso apresentado à Universidade Federal de Ouro Preto como parte dos requisitos para obtenção do Título de Bacharel em Engenharia Elétrica pelo Instituto de Ciências Exatas e Aplicadas da Universidade Federal de Ouro Preto.

Orientadora: Prof.^a Dr.^a Sarah Negreiros de Carvalho Leite

Coorientador: Prof. Dr. Harlei Miguel de Arruda Leite

Universidade Federal de Ouro Preto

João Monlevade

Dezembro, 2022

A994a Azevedo, Larissa Rangel de

Aplicação de técnicas de deep learning em sinais cerebrais.
[Manuscrito]./ Larissa Rangel de Azevedo. - 2022.

38 f. : il.: color., gráf., tab.

Orientador: Prof. Dra. Sarah Negreiros de Carvalho.

Coorientador: Prof. Dr. Harlei Miguel de Arruda Leite.

Monografia (Bacharelado). Universidade Federal de Ouro Preto.
Instituto de Ciências Exatas e Aplicadas. Graduação em Engenharia
Elétrica

1. Aprendizado do computador. 2. Interface cérebro-computador.
3. Potencial evocado (Eletrofisiologia). 4. Interação homem-máquina. I.
Leite, Sarah Negreiros de Carvalho. II. Leite, Harlei Miguel de Arruda.
III. Universidade Federal de Ouro Preto. IV. Título.

CDU 004.5

Bibliotecário(a) Responsável: Flávia Reis - CRB6-2431



FOLHA DE APROVAÇÃO

Larissa Rangel de Azevedo

Aplicação de técnicas de *deep learning* em sinais cerebrais

Monografia apresentada ao Curso de Engenharia Elétrica da Universidade Federal de Ouro Preto como requisito parcial para obtenção do título de bacharela em Engenharia Elétrica

Aprovada em 16 de dezembro de 2022.

Membros da banca

Profa. Dra. Sarah Negreiros de Carvalho Leite - Orientadora - Universidade Federal de Ouro Preto

Prof. Dr. Harlei Miguel de Arruda Leite - Coorientador - Universidade Federal de Ouro Preto

O trabalho foi avaliado somente pela professora orientadora e pelo professor coorientador em virtude do aproveitamento do artigo publicado em evento científico ou periódico, conforme previsto pela resolução do colegiado do curso.

Profa. Dra. Sarah Negreiros de Carvalho Leite, orientadora do trabalho, aprovou a versão final e autorizou seu depósito na Biblioteca Digital de Trabalhos de Conclusão de Curso da UFOP em 16 de dezembro de 2022.



Documento assinado eletronicamente por **Sarah Negreiros de Carvalho Leite**, **PROFESSOR DE MAGISTERIO SUPERIOR**, em 16/12/2022, às 11:06, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site http://sei.ufop.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0, informando o código verificador **0435756** e o código CRC **5AD4331B**.

"²⁰Falou Daniel e disse: Seja bendito o nome de Deus para todo o sempre, porque dele é a sabedoria e a força;

²¹Ele muda os tempos e as horas; Ele remove os reis e estabelece reis; Ele dá sabedoria aos sábios e ciência aos inteligentes.

²²Ele revela o profundo e o escondido e conhece o que está em trevas; e com Ele mora a luz."

Agradecimentos

Ao final desta fase quero agradecer primeiramente a Deus por me sustentar até o fim. Nos momentos alegres ou tristes e mesmo longe de casa e da família nunca me senti só, pois Sua presença sempre esteve comigo em tudo e em todas as coisas.

Em segundo lugar agradeço aos meus pais, Adriana e Luís Fernando, por serem anjos em minha vida e por me darem todo amor e carinho. Por terem lutado desde o início para que eu tivesse uma educação acima da média. Por aceitarem embarcar nesta jornada junto comigo. Pela compreensão quando tive que abdicar dos nossos momentos em família para dedicar-me ao meu sonho.

Agradeço aos meus irmãos Letícia e Luís Isac, pelas alegrias que me proporcionaram ao longo desta caminhada, cada bilhete surpresa de "Te amo" me fizeram sentir presente mesmo estando distante.

A todos os professores do ICEA os quais tive o privilégio de ser aluna, deixo meu muito obrigada. Em especial agradeço a minha orientadora, Prof^a Sarah Negreiros de Carvalho Leite, pela confiança depositada para realizar este trabalho que começamos ainda na iniciação científica, pelos ensinamentos e conversas. Agradeço também ao meu coorientador Harlei Miguel de Arruda Leite.

Obrigada aos meus colegas da turma de Engenharia Elétrica 16.2 e aos amigos de república por compartilhar em nosso dia a dia altos e baixos, conquistas e fracassos e tornarem esta caminhada mais leve e feliz, cada um que passou pela minha vida me trouxe alguma lição valiosa.

Finalmente, ao caminhar pela última vez pelos corredores desta instituição como aluna, consigo refletir sobre tudo que vivi nestes últimos anos e posso afirmar que os aprendizados vão além do que se vê em sala de aula, não mudaria nada do que fiz, pois cada escolha, cada passo que me levou para trás, me fizeram chegar até aqui.

Obrigada UFOP!

Resumo

Interfaces-cérebro computador (BCI) são sistemas que permitem realizar uma comunicação direta entre o cérebro do usuário e uma máquina. Neste trabalho, foi estudada a abordagem de BCI baseada em potenciais evocados visualmente em regime estacionário (SSVEP). Todas as etapas de processamento de sinais de uma BCI-SSVEP (filtragem, extração de características e classificação) foram implementadas. O enfoque do trabalho foi avaliar o impacto de técnicas de classificação no desempenho do sistema BCI. Para isso, três abordagens distintas de discriminação foram testadas: (1) classificador linear baseado no método dos mínimos quadrados que forneceu desempenho médio de 47%, (2) rede neural convolucional usando extração de características pelo método da FFT e topologia de 2 camadas ocultas empregando $kernel = 5$ e função de ativação do tipo tangente hiperbólica, que forneceu desempenho médio de 87% e (3) rede neural convolucional com os dados brutos, eliminando a etapa de extração de características pelo método da FFT, nesta configuração, a rede tem sua topologia dependente dos dados EEG colocados na entrada, assim foi feito o tunelamento somente para o indivíduo 32 da base de dados, e obteve-se um desempenho de 81%.

Palavras-Chave: BCI-SSVEP, CNN, Pré-processamento, Aprendizado de Máquina.

Abstract

Brain-Computer Interfaces (BCI) are systems that allow direct communication between the user's brain and a machine. In this work, the BCI approach based on steady-state visually evoked potentials (SSVEP) was studied. All signal processing steps of a BCI-SSVEP (pre-processing, feature extraction and classification) were implemented. The focus of this work was to evaluate the impact of ranking techniques on the performance of the BCI system. To this end, three different discrimination approaches were tested: (1) linear classifier based on least squares method that provided an average performance of 47%, (2) convolutional neural network with 2 hidden layers employing $kernel = 5$ and hyperbolic tangent as activation function and feature extraction, which provided an average performance of 87% and (3) convolutional neural network with the raw data, eliminating the feature extraction step, in this configuration, the network has its topology dependent on the EEG data placed in the input, so it was tunneling was performed only for individual 32 of the database, and a performance of 81% was obtained.

Keywords: BCI-SSVEP, CNN, preprocessing, Machine Learning.

Lista de ilustrações

Figura 1 – Esquema em blocos de uma BCI-SSVEP.	2
Figura 2 – Subdivisão anatômica básica do sistema nervoso.	4
Figura 3 – Subconjuntos da área de inteligência artificial.	8
Figura 4 – Neurônio biológico.	11
Figura 5 – Representação do perceptron.	12
Figura 6 – Representação de uma RNA.	13
Figura 7 – Representação gráfica da operação de convolução para uma entrada 1D.	14
Figura 8 – Representação gráfica da operação de <i>pooling</i> 1D.	15
Figura 9 – Disposição dos eletrodos por região do cérebro durante o processo de aquisição.	16
Figura 10 – Sinal ideal a) no domínio do tempo e b) no domínio da frequência.	20
Figura 11 – Sinais EEG evocados em 10 Hz no domínio do tempo e frequência. a) Sinal sem filtragem, b) Sinal filtrado.	21
Figura 12 – Espectro do sinal cerebral EEG evocado em 6 Hz (a) 3 s e (b) 12 s.	22
Figura 13 – Resposta SSVEP para as 4 frequências (6, 10, 12 e 15 Hz).	23
Figura 14 – Rateio de treinamento e validação.	24
Figura 15 – Arquitetura de uma CNN composta de uma camada de entrada, 3 camadas convolucionais e uma camada de saída.	26
Figura 16 – Primeira etapa de testes com a CNN-dados lapidados.	28
Figura 17 – Segunda etapa de testes com a CNN-dados lapidados.	28
Figura 18 – <i>Overfitting</i> - Sobreajuste dos dados.	31
Figura 19 – Acurácia do Indivíduo 32.	32

Lista de tabelas

Tabela 1 – Matriz de confusão para o Indivíduo 32.	24
Tabela 2 – Taxa de acertos para as 4 classes pelo método do CL.	25
Tabela 3 – Taxa de acertos para CNN-1, dados lapidados.	27
Tabela 4 – Taxa de acertos para CNN-2, dados lapidados.	29
Tabela 5 – Placar de desempenho dos classificadores.	30

Sumário

1	INTRODUÇÃO	1
1.1	Objetivos	5
1.2	Revisão Bibliográfica	6
2	APRENDIZADO DE MÁQUINA	8
2.1	Classificador Linear baseado no Método dos Mínimos Quadrados	9
2.2	Rede Neural Convolutacional	11
2.2.1	Modelo do Perceptron	11
2.2.2	Camada de Convolução	13
2.2.3	Camada de <i>Pooling</i>	15
3	METODOLOGIAS	16
3.1	Aquisição do Sinal	16
3.2	Pré-processamento	17
3.3	Extração de Características	18
3.4	Classificação	19
4	RESULTADOS E DISCUSSÃO	20
4.1	Pré-Processamento e Extração de Características	20
4.2	Classificação	23
4.2.1	Classificador Linear baseado no Método dos Mínimos Quadrados	23
4.2.2	Rede Neural Convolutacional	25
4.2.3	Análise Comparativa: Classificador Linear vs. Rede Neural Convolutacional	30
4.2.4	CNN com os Dados Brutos	31
5	CONCLUSÃO	33
	ANEXOS	34
	REFERÊNCIAS	37

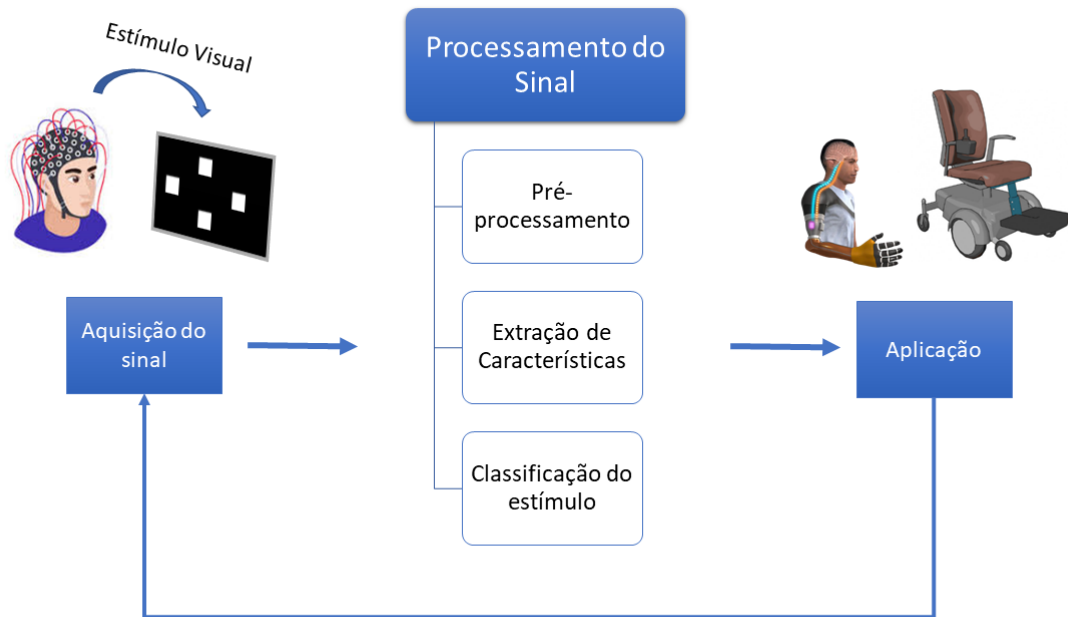
1 Introdução

O Aprendizado de Máquina (ML do inglês, *Machine Learning*) é uma área da Inteligência Artificial (IA), que tem como intuito desenvolver sistemas capazes de adquirir conhecimento através da experiência. Uma subcategoria do aprendizado de máquina é o Aprendizado Profundo de Máquina (DL do inglês, *Deep Learning*), que reúne técnicas mais sofisticadas para resolver problemas de predição mais complexos. As Redes Neurais Artificiais representam um tipo de técnica de aprendizagem, cujo modelo matemático é inspirado no cérebro humano e podem ser classificadas em rasas ou profundas, a depender do número de nós e camadas intermediárias. Sua principal funcionalidade é aprender a partir de exemplos e resolver problemas mais complexos melhorando seu desempenho. Estas estruturas têm sido amplamente usadas em sistemas de reconhecimento de voz, piloto automático de carros, detecção de fraudes, etc (GOODFELLOW, 2016).

Uma recente área de aplicação destas técnicas são as Interfaces Cérebro-Computador (BCI do inglês, *Brain-Computer Interface*), que são sistemas capazes de estabelecer uma comunicação direta entre o cérebro do usuário e uma máquina. Estes sistemas coletam os impulsos elétricos cerebrais provenientes de um estímulo e os associam ao acionamento de um elemento final de controle, que irá refletir a intenção do usuário, por exemplo ao movimentar uma cadeira de rodas ou uma prótese robótica.

O princípio de funcionamento da BCI baseada no potencial evocado visualmente em regime estacionário (SSVEP), ocorre devido a capacidade dos neurônios do córtex visual, localizados na região do lobo occipital do indivíduo, sincronizarem seus disparos com a frequência do estímulo visual ao qual esta sendo exposto. Como cada estímulo cintila em uma frequência diferente e é associado a um comando da BCI, os ritmos cerebrais são coletados por meio do EEG, processados e usados para o desenvolvimento de uma aplicação, como mostra o esquema de uma BCI-SSVEP da Figura 1.

Figura 1 – Esquema em blocos de uma BCI-SSVEP.



Fonte: Própria.

Segundo (GRAIMANN, 2009), os sinais elétricos oriundos da atividade cerebral podem ser medidos usando métodos invasivos ou não-invasivos. O primeiro método requer que o paciente seja submetido a uma cirurgia, para que os eletrodos sejam inseridos no tecido cortical, sem penetrar o cérebro, ou diretamente no cérebro. Quando dispostos apenas na superfície do córtex cerebral, o sinal é obtido da atividade de um grupo de neurônios próximo ao eletrodo, sendo chamado de eletrocorticograma (ECoG). Já o sinal registrado diretamente do cérebro é feito por meio de uma gravação intracortical. Estes procedimentos fornecem um sinal de melhor qualidade, porém são quase sempre inviabilizados pelo custo de se realizar o procedimento e risco de expor o paciente a esta cirurgia muito delicada.

Nos métodos de aquisição não-invasivos o sinal cerebral pode ser obtido através de uma ressonância magnética, magnetoencefalografia (MEG) ou de um eletroencefalograma (EEG), por exemplo, com eletrodos dispostos sobre o couro cabeludo do paciente. Esta técnica de EEG é bem estabelecida e comumente usada na medicina, para diagnósticos em exames. Além disso, é um procedimento que envolve menor custo, se comparado ao procedimento cirúrgico, mesmo levando em conta o valor do equipamento de aquisição por isso são os mais utilizados para desenvolvimento de sistemas BCI. Todavia, uma das desvantagens do EEG é ser suscetível a ruídos internos, originários das atividades involuntárias do paciente, como piscar de olhos, respiração e ruídos externos induzidos da rede elétrica, com isso é necessário tratar o sinal adquirido.

A fonte do sinal que alimenta a BCI é responsável por caracterizá-la em níveis de topologia e implementação, podendo ser oriundo de duas formas: endógena, quando o próprio usuário induz o sinal de resposta, como é o caso da imagética motora, ou exógena, quando o sinal é induzido através de um estímulo externo (GRAIMANN, 2009). Neste trabalho, foi adotada a abordagem de fonte de estímulo exógena, pois exige menor esforço do indivíduo. Diferentes abordagens podem ser implementadas como é explicado em (CARVALHO et al., 2015), no entanto as mais comuns são através do potencial evocado auditivo e potencial evocado visualmente em regime estacionário, sendo este último o modelo escolhido para objeto de estudo deste trabalho, denominado BCI-SSVEP.

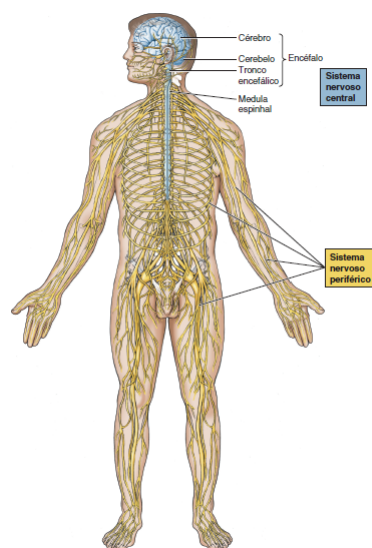
A BCI-SSVEP pode ser implementada usando diferentes técnicas e abordagens, no entanto sua construção elementar está compreendida em: aquisição do sinal cerebral, processamento do sinal e o desenvolvimento de uma aplicação, como mostra a Figura 1. O foco deste trabalho está na etapa de processamento do sinal cerebral.

Dentro da etapa de processamento existem subetapas, como o pré-processamento, realizado para atenuar os ruídos de aquisição, a extração das características e a classificação dos dados. A etapa de classificação é muito importante para o desempenho do sistema pois, é ela quem irá mapear os dados de entrada e gerar saídas que são usadas como um sinal de controle para a aplicação desejada.

A BCI-SSVEP implementada no trabalho foi composta por sinais oriundos de aquisição por meio de um EEG, em que as configurações usadas para os testes iniciais foram baseadas em resultados anteriores das pesquisas de iniciação científica (AZEVEDO et al., 2020) e (AZEVEDO; LEITE; CARVALHO, 2022) realizadas pela aluna e os orientadores.

A grande importância em estudar o comportamento do cérebro humano é que este sempre foi um mistério e motivo de muitos questionamentos sobre nosso comportamento, sentimentos e emoções. Alguns registros arqueológicos de quase 7 mil anos indicam hominídeos que morreram por golpe fatal no crânio originário de lutas entre guerreiros, evidenciando que já naquela época havia o entendimento que o cérebro era um dos órgãos mais importantes do corpo humano (BEAR, 2002). A medida que a neurociência evoluiu foi compreendido que o encéfalo, se comunica com o restante do corpo através de nervos, constituindo o sistema nervoso e que o mecanismo usado para movimentar os músculos são estímulos elétricos gerados pelo cérebro, o principal elemento que constitui o encéfalo, como mostra a Figura 2.

Figura 2 – Subdivisão anatômica básica do sistema nervoso.



Fonte: Adaptado de (BEAR, 2002).

Como se sabe, algumas lesões graves podem ocorrer em qualquer parte do sistema nervoso, sendo mais comuns na altura da coluna, que podem levar o indivíduo a um quadro de deficiência motora. Segundo o relatório mundial sobre a deficiência, cerca de um bilhão de pessoas possuem alguma forma de deficiência, das quais, 200 milhões têm dificuldades funcionais significativas (OMS, 2011).

O cenário brasileiro, de acordo com a Pesquisa Nacional de Saúde (PNS) realizada em 2019, mostra que das 17,3 milhões de pessoas de 2 anos ou mais de idade que foram investigadas e declaram ter algum tipo de deficiência, cerca de 7,8 milhões tinham deficiência física nos membros inferiores, 5,5 milhões tinham deficiência física relacionada aos membros superiores e o restante apresentava outros tipos de deficiência (visual, auditiva, mental, etc). Sendo índice de deficiência física nos membros inferiores maior entre mulheres (4,8%)

do que para homens (2,7%).

Desta forma, a principal motivação em estudar o comportamento do sistema nervoso é curar doenças e solucionar problemas. Com a ajuda dos sistemas BCI é possível reduzir ainda mais algumas limitações como dificuldade de locomoção, da fala, da audição, etc. vivida pelas pessoas com deficiência para que elas possam ter uma qualidade de vida melhor.

1.1 Objetivos

O objetivo central deste trabalho é aplicar Redes Neurais Artificiais à sinais cerebrais, afim de melhorar o desempenho de uma BCI-SSVEP. O modelo escolhido neste trabalho foi a Rede Neural Convolutacional. Os objetivos específicos foram:

- 1) Estudar e compreender o modelo de uma interface cérebro-máquina baseada no, potencial visualmente evocado em regime permanente (BCI-SSVEP) e as funcionalidades da Rede Neural escolhida.
- 2) Realizar as etapas pertinentes ao tratamento do sinal cerebral, como pré-processamento e extração de características.
- 3) Implementar a classificação dos estímulos da BCI-SSVEP usando um classificador linear e outro não-linear.
- 4) Comparar o desempenho da BCI-SSVEP com as diferentes técnicas de classificação.
- 5) Avaliar o desempenho da BCI-SSVEP em diferentes cenários da Rede Neural.

Este trabalho está particionado em cinco capítulos. O Capítulo 1, com a Introdução, faz a imersão do leitor sobre o tema da BCI, mostrando os objetivos propostos, a relevância do tema escolhido e os principais trabalhos que serviram de base para o estudo. No Capítulo 2 são discutidos os conceitos envolvendo a aprendizagem de máquina. O Capítulo 3 aborda as metodologias escolhidas para implementação de uma BCI-SSVEP. Os resultados obtidos através dos cenários simulados da BCI-SSVEP são discutidos no Capítulo 4, e o Capítulo 5 apresenta a conclusão do trabalho.

1.2 Revisão Bibliográfica

Neste estudo utilizou-se uma base de dados de sinais EEG coletada no âmbito do projeto Xtreme (LEITE et al., 2018). Desta forma a etapa de aquisição do sinal cerebral não foi implementada. Contudo, para compreender o funcionamento da BCI foi usado o livro do (GRAIMANN, 2009) que aborda aspectos sobre a origem de aquisição do sinal além das contribuições reunidas no trabalho de (CARVALHO et al., 2015) onde foi concebida uma BCI-SSVEP completa.

O foco principal deste trabalho encontra-se no processamento digital do sinal, mais precisamente na etapa de classificação, onde ocorre a tomada de decisão que irá apontar para qual estímulo o indivíduo estava focando sua atenção. Os principais estudos na área trazem uma abordagem de BCI-SSVEP usando classificadores lineares, que são capazes de resolver o problema da BCI com resultados satisfatórios, como foi relatado em (DE JESUS JR, 2016). No entanto, esta metodologia se mostra pouco flexível e sem muitas possibilidades de ajustes de parâmetros para melhorar o desempenho final dos sistema.

Estudos mais recentes têm introduzido o conceito de Redes Neurais aplicadas à BCI-SSVEP como algo promissor. Dentre elas estão o Perceptron Multicamadas (MLP, do inglês *Multi Layer Perceptron*) e as Redes Neurais Convolucionais (CNN, do *Convolutional Neural Network*). O trabalho de (RAVI, 2019) faz uma análise comparativa entre o algoritmo de detecção por correlação canônica (CCA), técnica similar ao classificador linear de mínimos quadrados no que diz respeito ao número de parâmetros para se ajustar, com a CNN para melhorar a percepção de potenciais SSVEP que possuem estímulos concorrentes, ou seja, de frequências próximas. O autor obteve uma taxa de acertos de 75% usando a CNN, contra 67% com o CCA, no modelo de BCI-SSVEP *offline* similar ao implementado neste trabalho.

Outro trabalho que reforça o uso de Redes Neurais foi o de (MULLER, 2010), que também utilizou a CNN, para um exoesqueleto controlado pelo cérebro em condições ambulatoriais, onde os autores encontraram para arquitetura CNN implementada, índices de classificação de 99,28% e 94,03% nas condições estática e ambulatorial (exoesqueleto se movimentando), respectivamente.

Como o processamento digital de sinais da BCI é realizado via algoritmos, foi usado o livro do (NIELSEN, 2015) que apresenta no Capítulo 6 várias sugestões de código para a implementação da CNN em linguagem *Python*. Além disso, para implementação e testes da estrutura do modelo que representasse a base de dados, afim de obter-se bons resultados e em tempo hábil de processamento, foi preciso compreender seu funcionamento e (GOODFELLOW, 2016) explica em seu livro o modelo matemático das Redes Neurais

que são inspirados no neurônio biológico.

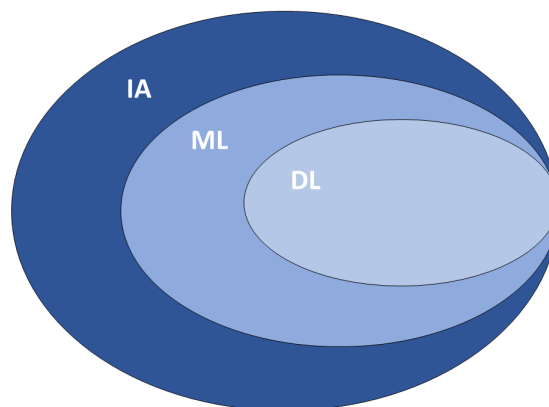
Assim, prevendo-se que no futuro o *hardware* da BCI possa ser implementada, deve-se otimizar ao máximo a etapa de processamento dos dados para que o sistema tenha uma resposta mais rápida quando um comando for acionado. Neste ponto, o trabalho de (RAMPAZZO et al., 2019) mostrou que as Redes Neurais também apresentam mais uma vantagem, pois permitem reduzir etapas de processamento, como a extração de características, usando os dados brutos na entrada.

2 Aprendizado de Máquina

Segundo (GOODFELLOW, 2016), um algoritmo de ML é capaz de aprender a partir de um conjunto de dados. Desta forma, se seu desempenho melhora então diz-se que ele aprendeu. Geralmente, as métricas usadas para atingir a melhor performance são baseadas em equações matemáticas, como o erro quadrático mínimo e determinação de superfícies de separação. Os algoritmos de DL também aprendem por dados, por isso diz-se que estes são uma subárea do ML, como mostra o diagrama da Figura 3.

Dentro destes dois conjuntos de técnicas existem as aborgadens de aprendizagem, as mais comuns são: O Aprendizado Supervisionado, em que o algoritmo recebe os dados de treinamento rotulados, desta forma o aprendizado é direcionado a resposta esperada na saída. E o Aprendizado Não-Supervisionado, onde os dados de treinamento não são rotulados, com isso o algoritmo precisa relacionar os dados para obter a saída. Além destas técnicas, também existe o aprendizado semi-supervisionado, por reforço, etc. Neste trabalho, foram usadas duas técnicas de classificação de dados por aprendizado supervisionado, descritas a seguir.

Figura 3 – Subconjuntos da área de inteligência artificial.



Fonte: Própria.

2.1 Classificador Linear baseado no Método dos Mínimos Quadrados

O Classificador Linear Baseado no Método dos Mínimos Quadrados (CL) funciona como um organizador, separando as respostas aos estímulos como peças em uma caixa. Porém, para que o método seja possível de ser empregado é preciso garantir que as características sejam linearmente separáveis e que é possível definir hiperplanos que separem as classes de estímulos visuais. Segundo descrito em (THEODORIDIS; KOUTROUMBAS, 2009), para o caso simples de apenas duas classes pode-se assumir que exista uma função linear para o hiperplano do tipo:

$$y(X) = \omega^T X + \omega_0 \quad (2.1)$$

Onde:

- $\omega = [\omega_1, \omega_2, \dots]$ é o vetor de pesos;
- ω_0 é o limiar do hiperplano;
- X é o vetor de amostras das classes.

A Equação (2.1) se assemelha à equação de uma reta, cujos coeficientes são definidos de forma matricial. Para $\omega_0 = 0$ tem-se a equação do hiperplano que cruza a origem, como:

$$y(X) = \omega^T X \quad (2.2)$$

Fazendo manipulações matriciais para colocar em evidência o vetor de pesos ω tem-se:

$$\omega = Y^{-1}X \quad (2.3)$$

Como Y é a saída desejada do classificador e não é conhecida o problema é resolvido considerando uma margem de erro mínimo entre a saída desejada e a verdadeira, usando o método dos mínimos quadrados dada por:

$$E^2 = (Y - X^T \omega)^2 = 0 \quad (2.4)$$

Objetivando o erro mínimo, este será dado pela soma do erro de cada elemento, como:

$$\sum_{i=1}^N e^2 = \sum_{i=1}^N (y_i - x_i^T \omega)^2 \quad (2.5)$$

Para uma função quadrática, o ponto de mínimo ocorre durante a inflexão, onde a derivada parcial da Equação (2.5) em relação a ω é nula, obtendo-se:

$$\sum_{i=1}^N x_i (y_i - x_i^T \omega) = 0 \quad (2.6)$$

Reescrevendo-se a Equação (2.6):

$$\sum_{i=1}^N x_i y_i = \sum_{i=1}^N x_i x_i^T \omega \quad (2.7)$$

Na forma matricial, tem-se:

$$X^T Y = (X^T X) \omega \quad (2.8)$$

Assim, para obtenção do vetor de pesos

$$\omega = (X^T X)^{-1} X^T Y \quad (2.9)$$

onde $(X^T X)^{-1} X^T$ é a pseudo-inversa da matriz dos atributos X extraídos pela FFT, portanto é invertível. Logo, o vetor de saída é determinado por:

$$Y = X \omega \quad (2.10)$$

Após o treinamento do classificador linear, foi realizada a validação (usando o restante da base de dados), desta forma pode-se testar a capacidade de generalização do algoritmo e a eficiência do classificador linear. Por isso, nas simulações feitas dividiu-se aleatoriamente as amostras destinadas ao treinamento e a validação, em dois conjuntos disjuntos (BISHOP, 2006).

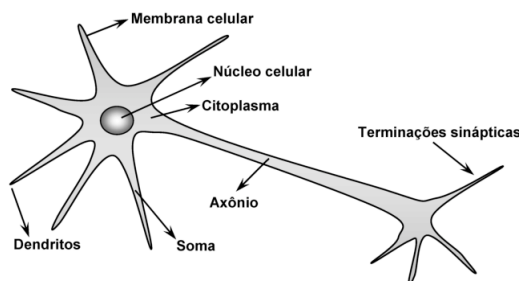
2.2 Rede Neural Convolutacional

2.2.1 Modelo do Perceptron

As Redes Neurais Artificiais (RNAs), sejam elas Convolucionais, *Multilayer Perceptrons*, Recorrentes, etc. levam este nome pois, seu funcionamento é inspirado no neurônio biológico do cérebro humano, e são bastante motivadoras, pois são capazes de manipular um grande número de informação, sendo muito usadas para o reconhecimento de padrões e aprendizado de máquina (HAYKIN et al., 2009).

As RNAs são sistemas paralelos compostos por camadas, que por sua vez são compostas por nós. Em linhas gerais, pode-se considerar que a estrutura e funcionamento dos neurônios está dividido em três partes cruciais: corpo celular (ou soma), dendritos e o axônio (BRAGA, 2000). Os dendritos são responsáveis por receber os impulsos nervosos de outros neurônios e encaminhá-los ao corpo celular, onde a informação será processada e um potencial de ativação indicará se o neurônio deverá gerar o impulso. Os axônios são o canal por onde os impulsos percorrem até as terminações sinápticas (FLAUZINO, 2010). Veja na Figura 4 uma representação do neurônio humano.

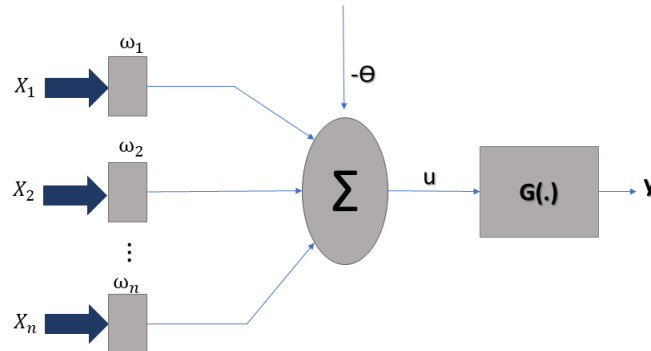
Figura 4 – Neurônio biológico.



Fonte: (FLAUZINO, 2010).

Este sistema é eficiente e proporciona ao ser humano um nível de aprendizado extremo, pois estima-se que o cérebro humano possua 10^{11} neurônios. O modelo matemático proposto para a representação do neurônio e usado no aprendizado de máquina, é chamado *perceptron* (BRAGA, 2000). Este pode ser considerado o modelo mais simples de uma rede neural, constituído por um camada com n -amostras de entrada e uma única saída.

Figura 5 – Representação do perceptron.



Fonte: Própria.

A Figura 5 apresenta o esquema de um Perceptron, no qual as amostras de entrada $[x_1, x_2, \dots, x_n]$, que contém as informações dos estímulos, é mapeada e ponderada pelos pesos sinápticos $[\omega_1, \omega_2, \dots, \omega_n]$. O ponto de soma introduz o limiar de ativação $[-\theta]$, análogo ao potencial de ativação do neurônio biológico; esta soma u é repassada para a função de ativação que produzirá uma saída y .

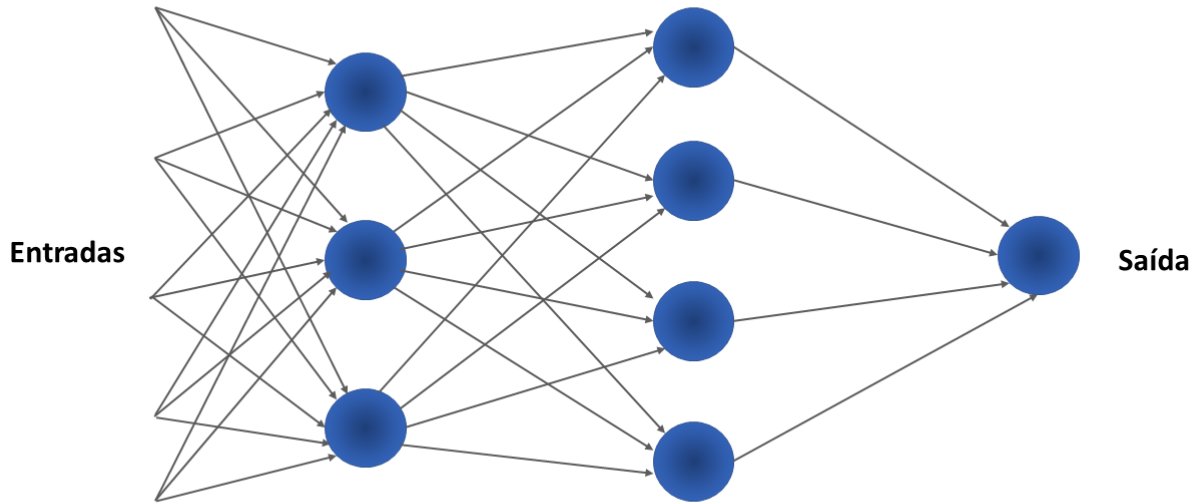
O modelo matemático que representa o *perceptron* é dado pelas expressões (FLAUZINO, 2010):

$$u = \sum_{i=1}^n w_i x_i - \theta \quad (2.11)$$

$$y = g(u) \quad (2.12)$$

onde x_i são as amostras de entradas da rede, w_i é o peso ligado à i -ésima entrada, $-\theta$ é o limiar de ativação, $g(\cdot)$ é a função de ativação e u é o potencial de ativação. A Figura 6, mostra a representação de uma RNA do tipo MLP, composta de uma camada de entrada, duas camadas intermediárias com 3 e 4 neurônios e apenas um neurônio de saída.

Figura 6 – Representação de uma RNA.



Fonte: Própria.

2.2.2 Camada de Convolução

As Redes Neurais Convolucionais levam esta nomenclatura, pois o método matemático implementado pelo código da estrutura é a convolução, que é uma operação matemática entre duas funções representada pela integral de convolução, da seguinte forma (OPPENHEIM, 2010):

$$y(t) = x_1(t) * x_2(t) = \int_{-\infty}^{\infty} x_1(\tau)x_2(t - \tau)d\tau \quad (2.13)$$

A integral da Equação (2.13) é definida para sinais de tempo contínuo, o sinal EEG é um sinal analógico, onde t pode assumir valores reais não negativos, porém quando este sinal é lido pelo equipamento de aquisição ele é submetido aos processo de amostragem, logo a operação equivalente a Equação (2.13) para um sinal amostrado é a soma de convolução, dada por:

$$y[n] = x_1[n] * x_2[n] = \sum_{k=-\infty}^{\infty} x_1[k]x_2[n - k] \quad (2.14)$$

Na prática, os dados de entrada da CNN são uma matriz de dados $I[m, n]$ que é convoluída por um *Kernel* de dimensão $K[m, n]$. Como as amostras são finitas, o somatório infinito torna-se:

$$y[i, j] = (I * K)[i, j] = \sum_m \sum_n I[m, n]K[i - m, j - n] \quad (2.15)$$

Usando a propriedade comutativa do somatório, tem-se:

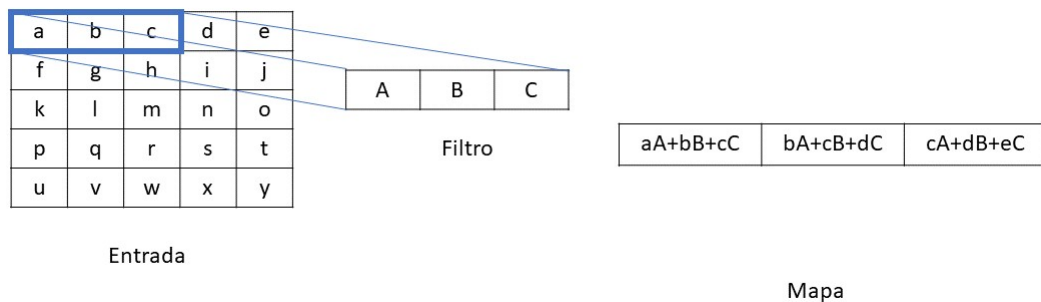
$$y[i, j] = (I * K)[i, j] = \sum_m \sum_n I[i - m, j - n]K[m, n] \quad (2.16)$$

Segundo (GOODFELLOW, 2016), o algoritmo por trás das bibliotecas da CNN realiza a operação de correlação cruzada ao invés da convolução, representada por:

$$y[i, j] = (I * K)[i, j] = \sum_m \sum_n I[i + m, j + n]K[m, n] \quad (2.17)$$

Desta forma, pode-se observar a semelhança entre as Equações (2.16) e (2.17) exceto pela reflexão do *Kernel* na primeira equação. A Figura 7 representa a operação de convolução em 2-D sem a reflexão do *Kernel* para uma matriz de entrada bidimensional 5×5 , e um kernel também bidimensional 3×3 . Os quadrados demarcados de azul e vermelho, representam os dados de entrada que foram convoluídos pelo *Kernel*, o resultado mapeado é lançado na saída, com as cores correspondentes de cada neurônio. A estes quadrados dar-se o nome de campo receptivo local, observe que algumas informações são compartilhadas pelos neurônios, assim o algoritmo consegue reconhecer padrões pela repetição dessas informações.

Figura 7 – Representação gráfica da operação de convolução para uma entrada 1D.



Fonte: Própria.

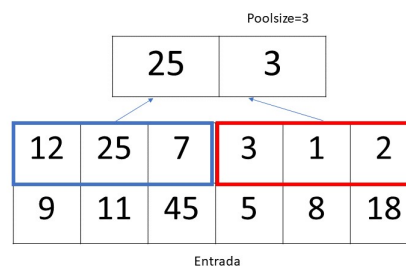
Essas peculiaridades da CNN, como a forma que os neurônios se conectam, a convolução com o *Kernel* criando os campos receptivos locais e o compartilhamento das informações, tornam este tipo de classificador não-linear eficiente e rápido. A equivariância da saída em relação à entrada ocasionada pelo processo de convolução, torna o sistema mais robusto e eficiente reduzindo operações (RAMPAZZO et al., 2019).

2.2.3 Camada de *Pooling*

Existem três tipos de camadas que compõem uma rede neural: a camada de entrada, camada oculta (ou intermediária) e a camada de saída. Cada uma possui um determinado número de neurônios definido pelo usuário, no entanto diferente dos outros tipos de RNA's, nas redes convolucionais apenas a camada de saída é totalmente conectada. Ou seja, os neurônios das camadas ocultas e de entrada não se conectam com todos os outros neurônios da camada seguinte.

As camadas de *pooling* são usadas entre as camadas intermediárias e servem para agrupar os dados em uma mesma representação mesmo que eles tenham sofrido translações, no caso das séries temporais, elas são usadas para resumir os dados de saída da camada convolucional. Existem algumas funções de agrupamento, como a *Averagepooling*, *Maxpooling* e a *GlobalAveragepooling*. A função de *Maxpooling* por exemplo, reduz os dados considerando o valor máximo em uma janela de tamanho *poolsize*, como é exemplificado na Figura 8.

Figura 8 – Representação gráfica da operação de *pooling* 1D.



Fonte: Própria.

O salto da janela de pooling é dada por um número inteiro que representa a quantidade de passos, se não for definido então ele se moverá com o mesmo valor de *poolsize*.

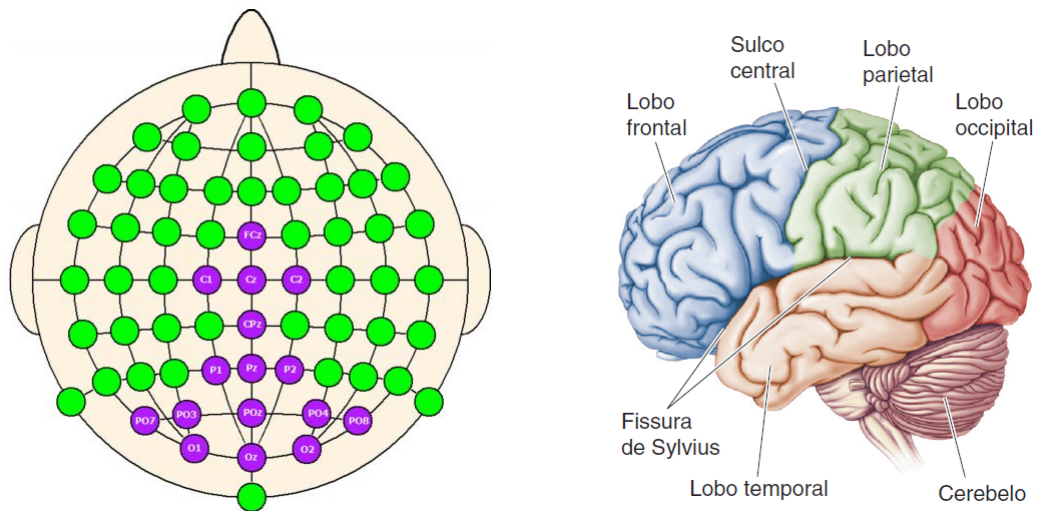
3 Metodologias

Neste Capítulo é detalhado os métodos escolhidos para o modelo da BCI-SSVEP simulada neste estudo, bem como a relevância de cada etapa de implementação.

3.1 Aquisição do Sinal

A primeira etapa do sistema BCI consiste em realizar a aquisição do sinal cerebral, neste trabalho utilizou-se a mesma base de dados do trabalho de (LEITE et al., 2018). A coleta dos sinais cerebrais foi realizada de maneira não invasiva, por meio de um eletroencefalograma, em que 16 eletrodos foram dispostos no escalpo do indivíduo seguindo padrão internacional 10-10. Os eletrodos foram predominantemente dispostos na região occipital e parieto occipital, aonde espera-se que o potencial evocado visualmente seja mais expressivo, conforme mostra a Figura 9.

Figura 9 – Disposição dos eletrodos por região do cérebro durante o processo de aquisição.



Fonte: (CARVALHO et al., 2015) e (BEAR, 2002).

Para gerar o potencial evocado visualmente, cada voluntário foi submetido a um experimento que consistia em estímulos visuais dipostos em um monitor de fundo preto com quatro quadrados brancos que cintilavam nas frequências de 6, 10, 12 e 15 Hz em que o voluntário focava sua atenção em um dos quadrados. O procedimento foi feito intercalando 8 coletas de 12 segundos de duração para cada estímulo, com intervalos de descanso regulados pelo próprio voluntário. A frequência de amostragem do sistema foi de 256 Hz. Ao todo participaram do experimento 32 voluntários saudáveis. Mais detalhes sobre o processo de aquisição dos sinais cerebrais pode ser consultado em (LEITE et al., 2018).

3.2 Pré-processamento

Os sinais cerebrais lidos durante a aquisição são da ordem de microvolts e, mesmo com a boa resolução do equipamento de aquisição, há uma baixa relação sinal ruído devido às interferências internas, como a respiração, batimentos cardíacos, piscar de olhos, e externas ao indivíduo, como cabeamento do equipamento de coleta, interferentes eletromagnéticos do ambiente, iluminação não controlada, entre outros (LUDWIG et al., 2009). A etapa de pré-processamento visa melhorar a qualidade do sinal de maneira a tornar mais nítido o potencial evocado visualmente e facilitar as demais etapas, impactando na qualidade do sistema BCI.

O pré-processamento se inicia no equipamento de aquisição, quando o sinal cerebral é submetido a um filtro passa-faixas com frequências de corte entre 5 e 50 Hz e um filtro *nocth* para eliminar a componente de 60 Hz oriunda da interferência externa da rede elétrica (TAVARES et al., 2015). Na sequência, ocorre o pré-processamento digital onde foi implementado o filtro de Referência de Média Comum (CAR, do inglês *Common Average Reference*), que consiste na subtração do valor lido em cada eletrodo pela média temporal dos 16 eletrodos, eliminando os ruídos que são comuns entre todos os eletrodos (DE JESUS JR., 2016).

Sendo definido pela Equação 3.1:

$$V_i^{CAR} = V_i^{EL} - \frac{1}{16} \sum_{j=1}^{16} V_j^{EL} \quad (3.1)$$

Onde:

- V_i^{CAR} é o valor filtrado por CAR do i -ésimo eletrodo.
- V_i^{EL} o valor inicial de tensão eletrodo i .

3.3 Extração de Características

A etapa de extração de características é responsável por indicar ao classificador, de maneira compacta e numérica, as características do sinal que permitem diferenciar e identificar para qual estímulo o indivíduo estava focando sua atenção. Seria possível usar os dados filtrados diretamente na classificação, mas isso aumenta a dimensionalidade do problema. A extração de características, assim como todas as etapas anteriores visa melhorar o desempenho e acurácia da BCI (COSTA et al., 2015).

Em análise de sinais é muito comum observar as formas de onda de uma função $g(t)$, por exemplo, que pode representar uma grandeza elétrica. Onde a variável independente é o tempo, entretanto há casos em que é mais conveniente observá-las em relação a frequência. Este artifício é possível empregando o método matemático da Transformada de Fourier que permite a visualização espectral de um sinal temporal (GOMES, 2001).

A partir do espectro do sinal cerebral, é possível obter informações relevantes sobre suas componentes de frequência, amplitude e fase. Em particular, a análise das amplitudes de cada componente de frequência que compõem o espectro do sinal permite definir para qual estímulo o voluntário estava sendo submetido. Neste trabalho, empregou-se o algoritmo da Transformada Rápida de Fourier (FFT, do inglês *fast Fourier transform*) para desenvolver esta análise (OPPENHEIM; SCHAFER, 2014).

Um outro aspecto importante para o bom desempenho da BCI é o tempo que o voluntário deve fixar seu olhar no estímulo para ativar um determinado comando para a aplicação a ser controlada. Este tempo determina a dinâmica de funcionamento de todo o sistema e quão rápido os algoritmos de processamento de sinais devem atuar desde o pré-processamento até a classificação. Neste trabalho, adotou-se segmentos de 3 s do sinal temporal para análise.

A janela deve ser desenvolvida visando uma ponderação entre a latência de resposta da aplicação, a dinâmica do sistema, a velocidade de processamento disponível e a acurácia da BCI. Quanto menor a janela temporal, menor também é a densidade espectral nas frequências evocadas, dificultando a etapa de extração de atributos e classificação o que pode impactar na redução de desempenho do sistema. Entretanto, o usuário leva menos tempo submetido ao estímulo aumentando a velocidade de resposta e interação com o sistema BCI.

3.4 Classificação

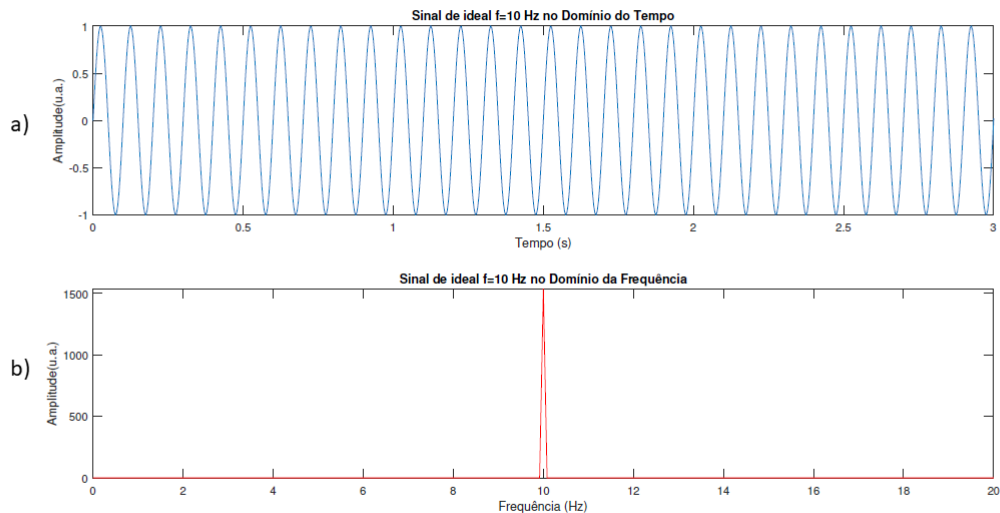
A classificação dos estímulos é o núcleo do sistema BCI, pois é através dela que o usuário é capaz de controlar as aplicações desejadas. Esta etapa possui particularidades, uma delas diz respeito à separação das amostras, se elas são linearmente ou não-linearmente separáveis. Este critério define o tipo de classificador que melhor se adapta ao problema. O método comumente utilizado para categorizar os dados SSVEP é o CL, como em (TAVARES et al., 2015). Neste trabalho foram realizados três cenários: CL com os dados extraídos pela FFT *vs.* CNN 1-D com os dados extraídos pela FFT e CNN 1-D com os dados brutos.

4 Resultados e Discussão

4.1 Pré-Processamento e Extração de Características

Os sinais cerebrais da base de dados fornecida possuem uma baixa relação sinal ruído oriundos de interferências internas e externas, algo muito recorrente durante a etapa de aquisição. Por isso, a necessidade de realizar-se o pré-processamento, que consistiu na filtragem dos ruídos comuns pelo método CAR. Em um sistema BCI-SSVEP ideal o sinal amostrado apresenta o comportamento de uma senóide no domínio do tempo, que corresponde a um impulso na frequência evocada quando observado no domínio da frequência. A Figura 10 ilustra este comportamento idealizado de um sinal oscilante na frequência de 10 Hz. Entretanto, no sistema real há a presença de ruídos que podem dificultar a identificação da frequência evocada.

Figura 10 – Sinal ideal a) no domínio do tempo e b) no domínio da frequência.

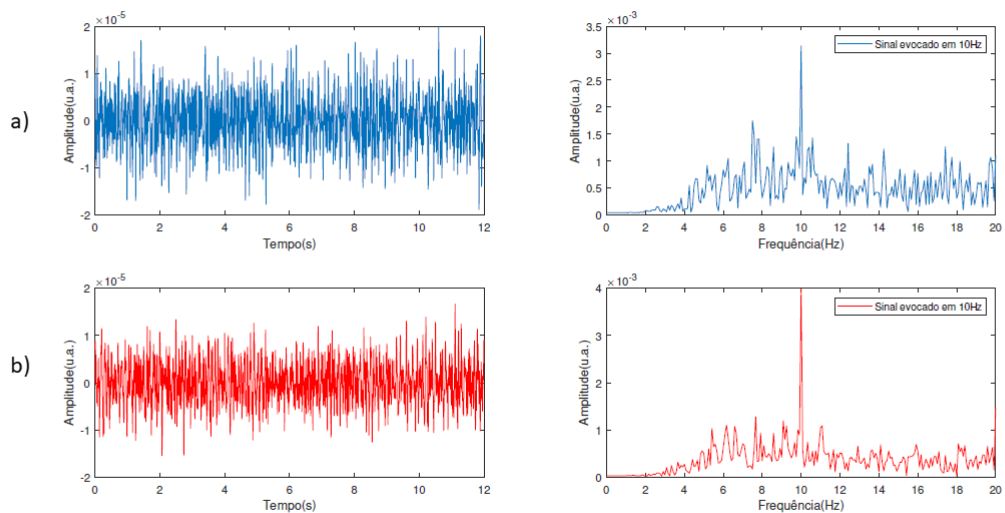


Fonte: Própria.

A Figura 11 apresenta um sinal EEG evocado em 10 Hz coletado pelo eletrodo O_Z . É possível observar que o sinal no domínio do tempo e seu espectro, antes e após o processo de filtragem CAR. Observa-se que o sinal temporal sofreu uma redução em

amplitude após a filtragem, justamente em decorrência da remoção de parte dos picos de ruídos e do valor de *offset*. No domínio da frequência, o sinal filtrado apresenta uma atenuação da amplitude nas diversas frequências, evidenciando melhor o pico em 10 Hz que corresponde a resposta SSVEP na frequência desejada.

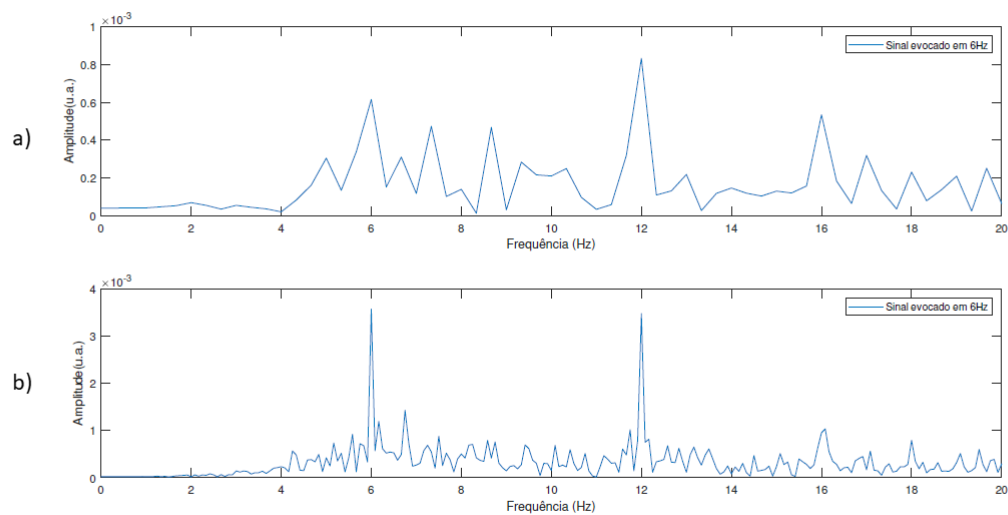
Figura 11 – Sinais EEG evocados em 10 Hz no domínio do tempo e frequência. a) Sinal sem filtragem, b) Sinal filtrado.



Fonte: Própria.

Após a filtragem, executou-se o janelamento do sinal de 12 s em segmentos de 3 s, de maneira a reduzir o tempo de extração das características e classificação da BCI-SSVEP. A Figura 12 apresenta o sinal EEG evocado em 6 Hz, pode-se observar o efeito resultante do janelamento sobre os picos de interesse (6 Hz e na harmônica de 12 Hz).

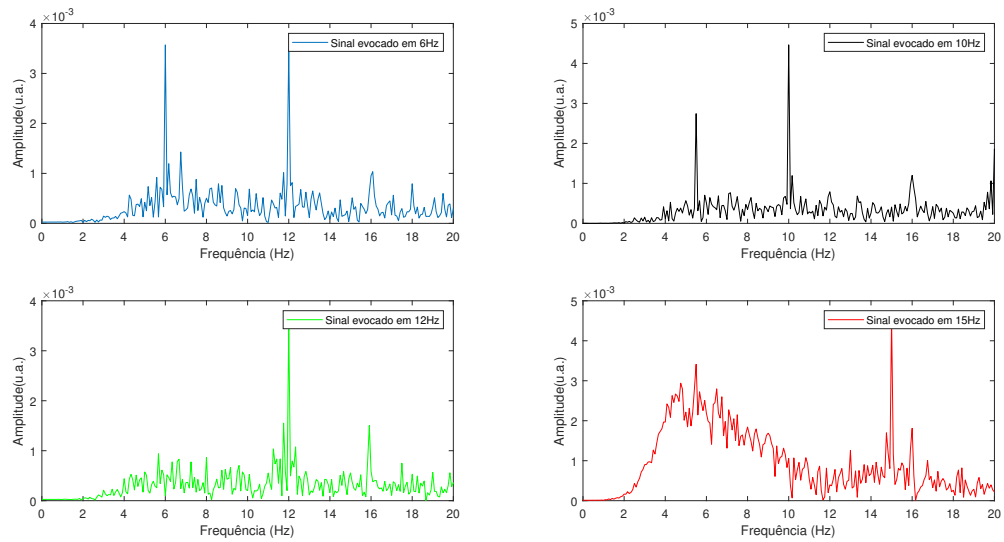
Figura 12 – Espectro do sinal cerebral EEG evocado em 6 Hz (a) 3 s e (b) 12 s.



Fonte: Própria.

Após o pré-processamento, realizou-se a extração das características por meio da FFT, afim de selecionar as informações relevantes ao sistema. A Figura 13 mostra os espectros para as 4 classes de frequência (6, 10, 12 e 15 Hz) do sinal do indivíduo 30 obtidos em Oz.

Figura 13 – Resposta SSVEP para as 4 frequências (6, 10, 12 e 15 Hz).



Fonte: Própria.

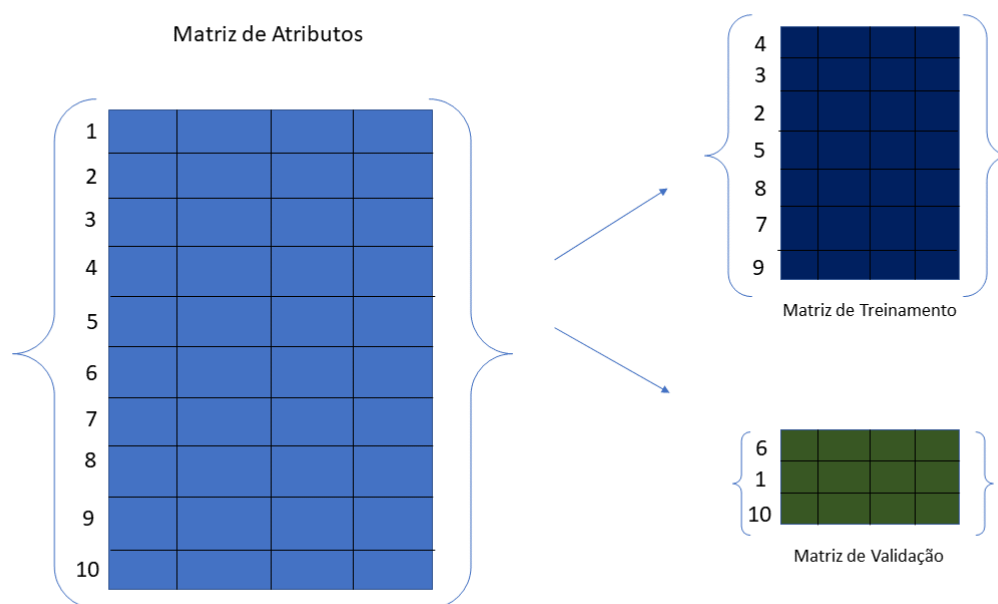
4.2 Classificação

4.2.1 Classificador Linear baseado no Método dos Mínimos Quadrados

O primeiro método de classificação implementado foi o Classificador Linear baseado no Método dos Mínimos Quadrados, doravante referenciado como CL, usando linguagem *Python*. A matriz de atributos extraídos pelo método da FFT foi usada para os dados de entrada do CL, além disto, uma prática muito comum quando se está lidando com análise de dados é dividir amostras em uma parte para treinamento e outra para a validação do modelo criado.

Neste estudo, fez-se uma divisão aleatória, por meio de um sorteio selecionando-se as linhas da matriz de atributos que foram destinadas para concatenação da matriz de treinamento e matriz de validação, tendo cautela para que as linhas não se repetissem. O sorteio é importante para estimar com maior precisão a eficiência e capacidade de generalização do classificador, foram realizados 10 sorteios para testes. Neste trabalho optou-se por usar 70% dos dados para treinamento e 30% para a validação. O desempenho médio foi considerado como parâmetro para avaliar o desempenho do sistema. A Figura 14 exemplifica o processo do rateio das matrizes de treinamento e validação, observe que as linhas da matriz de atributos são divididas para gerar as matrizes de treinamento e validação, sem repetição de dados.

Figura 14 – Rateio de treinamento e validação.



Fonte: Própria.

O CL realizou as predições e comparou-as com um vetor de rótulos, o número de acertos e erros para cada classe foi registrado em uma matriz de confusão (MC) que serviu para computar a análise de desempenho do classificador, esta técnica também foi usada na Rede Neural. A MC é gerada após cada classificação, apresentando em sua diagonal principal as taxas de acerto de cada classe e os demais elementos são os erros. A Tabela 1 mostra um exemplo de resultado da matriz de confusão para o Indivíduo 32. É possível identificar, neste caso, que o classificador encontrou dificuldades em diferenciar as classes de frequências 6 Hz e 12 Hz, por serem frequências harmônicas, por outro lado na frequência de 10 Hz obteve-se 100% de acerto.

Tabela 1 – Matriz de confusão para o Indivíduo 32.

	6 Hz	10 Hz	12 Hz	15 Hz
6 Hz	7	0	3	0
10 Hz	0	10	0	0
12 Hz	2	0	8	0
15 Hz	1	0	0	9

A Tabela 2 mostra o resultado médio referente aos 10 sorteios para as frequências de 6, 10, 12 e 15 Hz usando o CL para todos os sinais da base de dados. A média geral para a BCI-SSVEP usando o classificador linear foi de 47,4% com desvio padrão de 17,5%.

Este resultado demonstra que o sistema possui capacidade de atuar na discriminação, uma vez que a taxa de acerto randômica para 4 classes seria de 25%. Entretanto, a acurácia ainda é baixa para o controle satisfatório de algumas aplicações. Há de se considerar ainda que a taxa de acerto varia bastante entre os 32 indivíduos, devido à peculiaridade neuro-fisiológica de cada um.

Tabela 2 – Taxa de acertos para as 4 classes pelo método do CL.

Indivíduo	Acurácia(%)	Indivíduo	Acurácia(%)
1	74.7	17	49.0
2	26.6	18	36.3
3	32.9	19	50.4
4	38.8	20	65.6
5	34.4	21	45.0
6	42.7	22	25.7
7	26.0	23	32.4
8	79.4	24	34.3
9	31.5	25	32,4
10	51.3	26	44.2
11	53.4	27	63.6
12	78.9	28	49.1
13	31.9	29	49.6
14	55.6	30	83.9
15	32.6	31	33.6
16	47.4	32	85.6

4.2.2 Rede Neural Convolutacional

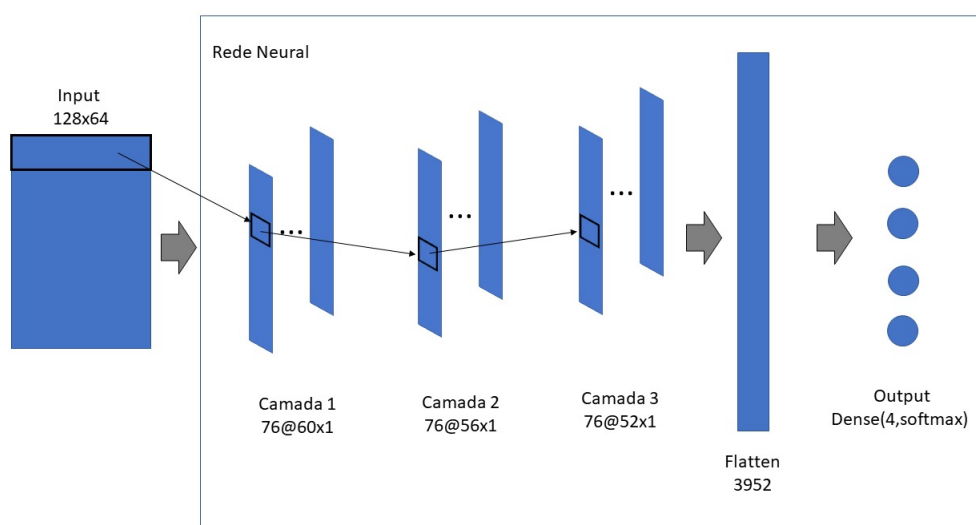
A CNN foi implementada também em *Python* usando o pacote *keras* que é uma biblioteca do *TensorFlow*, uma plataforma de código aberto usada para criar e treinar modelos de *machine learning* (GULLI; KAPOOR; PAL, 2019). Para modelá-la foi preciso configurar as camadas do modelo e em seguida compilá-lo. Para isso foi usada a função `model = Sequential()`, que cria a pilha de camadas. A esta pilha adicionou-se as camadas usando a função `model.add()` cada camada possui parâmetros disponíveis que auxiliaram na estruturação da CNN que apresentou melhor desempenho para a BCI.

Alguns destes são:

- *filters*: um inteiro, que representa o número de filtros de saídas na convolução.
- *kernel_size*: inteiro (ou lista), especifica o comprimento da janela de convolução.
- *activation*: função de ativação da camada.
- *input_shape*: formato dos dados da camada de entrada.

Neste trabalho, utilizou-se a redes neurais convolucionais de 1 dimensão, assim como os artigos de (RAVI, 2019) e (MULLER, 2010). Para isto, utilizou-se a função *Conv1D*, que inicializa um *kernel* que será convoluido em apenas um dimensão, espacial ou temporal, no estudo em questão na temporal. A Figura 15 exemplifica a estrutura de uma CNN em uma 1 dimensão, composta por 3 camadas convolucionais, uma camada de *flatten* e uma camada densa com 4 neurônios de saída. No exemplo em questão foi utilizado um *kernel* de dimensão 5×1 e 76 filtros.

Figura 15 – Arquitetura de uma CNN composta de uma camada de entrada, 3 camadas convolucionais e uma camada de saída.



Fonte: Própria.

Depois de criar as camadas da rede, a configuração estabelecida pode ser compilada, usando o comando *model.compile*, com otimizador ajustado para *'Adam'*, função de perda ajustada para *'categorical_crossentropy'* e métrica configurada para *'accuracy'*. Na simulação da BCI-SSVEP empregou-se a acurácia que determina a proporcionalidade entre as predições e os rótulos através de uma matriz de confusão. Por último, fez-se o treinamento da CNN com a função *model.fit()*, durante esta fase foi usado uma parte da matriz de atributos, extraídos pela FFT, para treinamento, exatamente 70% dos dados e os outros 30% para validação.

- **Primeiro teste:**

Para testar a CNN foram usados os dados já lapidados da matriz de atributos extraídos pelo método da FFT e divisão aleatória de 70% dos dados para treinamento e 30% para validação. A configuração usada foi 1 camada de entrada, 1 camada oculta, com janela de convolução de tamanho 3×1 e função de ativação 'ReLU' e 1 camada densa na saída com função de ativação 'Softmax'. Foram realizados 10 partições aleatórias da base de dados para gerar os conjuntos disjuntos de treinamento e validação. Esta estrutura foi embasada em estudos anteriores. O resultado da média entre as 4 classes para cada um dos 32 voluntários é mostrado na Tabela 3. A taxa de acerto média para todos os voluntários foi de 52,8% e desvio padrão de 19%.

Tabela 3 – Taxa de acertos para CNN-1, dados lapidados.

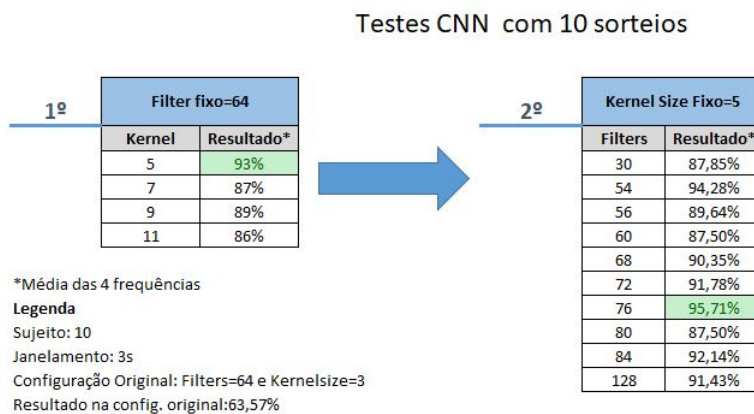
Indivíduo	Acurácia(%)	Indivíduo	Acurácia(%)
1	85.7	17	56.1
2	28.6	18	52.5
3	41.4	19	55.7
4	40.7	20	73.6
5	30.7	21	47.1
6	55.7	22	22.5
7	28.9	23	48.9
8	74.3	24	48.2
9	27.9	25	28.6
10	63.6	26	44.3
11	61.4	27	72.5
12	82.9	28	48.2
13	27.5	29	58.2
14	62.1	30	93.9
15	42.5	31	40.7
16	57.9	32	87.5

Para melhorar a taxa de acerto média variou-se o tamanho do *kernel*, o número de filtros, a função de ativação e o número de camadas ocultas. Para o segundo teste, foram escolhidos os indivíduos 7 e 10 para realizar simulações com a variação dos parâmetros da CNN, pois o primeiro obteve uma taxa de acertos para as 4 classes de apenas 28,9% e o segundo ficou um pouco acima da média com 63,6%.

• Segundo teste:

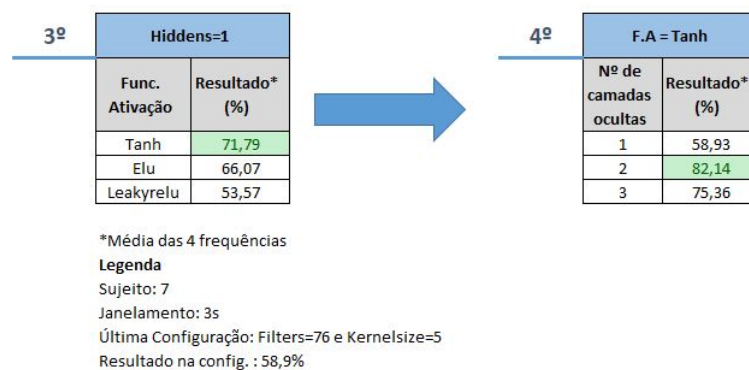
Neste segundo teste as simulações consistiram em duas etapas: a primeira foi variando somente o número de filtros e o tamanho da janela de convolução. Variou-se o tamanho do kernel de 5 a 11, considerando os valores ímpares. Então, foi fixado o kernel que forneceu o melhor desempenho, e foi testado diferentes valores para número de filtros, até a quantidade de 128, conforme mostra a Figura 15. A segunda simulação foi realizada iniciando dos resultados da etapa anterior, alterando as funções de ativação e o número de camadas ocultas como mostra a Figura 17.

Figura 16 – Primeira etapa de testes com a CNN-dados lapidados.



Na Figura 16, apenas variando o tamanho da janela de convolução, a taxa de acertos para o indivíduo 10 aumentou de 63,6% para 93%. Após, mantendo o tamanho da janela fixo em 5 e modificando o número de filtros percebeu-se que para valores menores a média reduziu até que o melhor ponto foi com *filters* = 76, em que o indivíduo atingiu a média de 95,7% e para valores maiores que este e menores que 128, o resultado saturou.

Figura 17 – Segunda etapa de testes com a CNN-dados lapidados.



Na sequência, fixou-se o $kernel_size = 5$ e $filters = 76$, alterou-se a função de ativação entre as especificadas na Figura 17. Nesta etapa utilizou-se o indivíduo 7, pois o indivíduo 10 obteve um resultado bastante superior na etapa precedente. Pode-se observar que ao definir estes valores para os dois parâmetros em questão, o indivíduo 7 obteve uma melhora em sua média de 28,9% (configuração inicial) para 58,9%. Usando a função de ativação tangente hiperbólica (Tanh), seu desempenho subiu para 71,79%, no passo seguinte adicionou-se camadas intermediárias à rede, o melhor desempenho de 82,14% foi encontrado empregando-se 2 camadas.

A Tabela 4 exhibe a taxa de acertos média para as 4 classes de todos os 32 voluntários, usando a configuração final de: $kernel_size = 5$, $filters = 76$, $activation = 'Tanh'$ e 2 camadas ocultas. Com esta configuração, a média geral da taxa de acerto para todos os voluntários subiu de 52,8% para 87% com desvio padrão de 7,8%.

Tabela 4 – Taxa de acertos para CNN-2, dados lapidados.

Indivíduo	Acurácia(%)	Indivíduo	Acurácia(%)
1	93.2	17	87.5
2	78.2	18	89.6
3	84.3	19	88.6
4	76.8	20	98.6
5	76.0	21	88.6
6	88.6	22	80.0
7	80.4	23	81.8
8	96.8	24	72.9
9	82.5	25	68.9
10	92.1	26	88.6
11	91.0	27	92.5
12	97.5	28	85.4
13	78.6	29	91.4
14	92.1	30	99.6
15	82.5	31	89.3
16	90.7	32	100

4.2.3 Análise Comparativa: Classificador Linear vs. Rede Neural Convolutacional

Após os testes durante a implementação dos classificadores é possível comparar o desempenho para a BCI-SSVEP. Alguns pontos observados têm papel importante, são eles: tempo de execução do código, acurácia e média geral. A Tabela 5 mostra de maneira sintetizada a média geral de todos os 32 voluntários com o respectivo desvio padrão e tempo de execução. Todas as técnicas foram avaliadas sob as mesmas circunstâncias, ou seja, mantendo fixa a base de dados, o tempo de janelamento (3 segundos), número de amostras para validação (30%) e quantidade de sorteios (10).

Tabela 5 – Placar de desempenho dos classificadores.

Classificador	Média Geral (%)	Desvio Padrão (%)	Tempo de Execução (min)
CL	47	17.5	2,5
CNN	87	7.8	180

O CL é a técnica mais simples implementada a qual obteve um desempenho geral de 47% e seu tempo de execução foi o menor entre os classificadores. A média geral pode ser melhorada reservando menos amostras para validação e mais para treinamento, aumentando o número de sorteios para cada indivíduo ou adicionando mais técnicas de pré-processamento, por exemplo, a seleção de características, onde somente os eletrodos de maior potencial são considerados. Assim, a dimensão do problema ficaria reduzida e o CL encontraria mais facilmente o potencial evocado.

Contudo, estes ajustes se encontram no processamento do sinal e não no classificador propriamente dito, porém para avaliar os métodos de classificação foi preciso expô-los as mesmas circunstâncias. Uma vez que, o CL não apresenta tanta flexibilidade no ajuste de parâmetros, mas foi capaz de resolver o problema da BCI-SSVEP com desempenho aceitável.

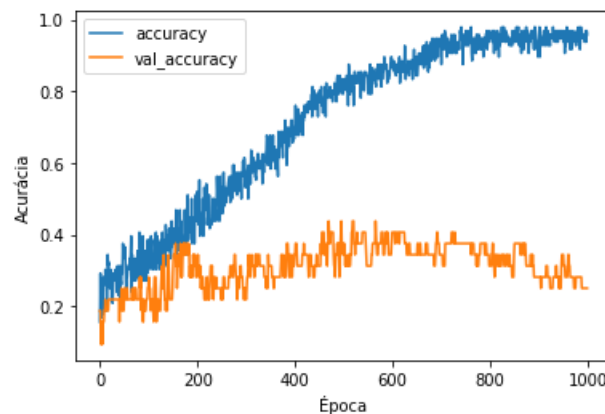
A CNN por sua vez, obteve um desempenho geral de 87%, pois possui um grande número de parâmetros que possibilitam o ajuste fino das superfícies de separação das classes, dentre eles se destacam: o tamanho da janela de convolução (*kernel size*), número de filtros por camada, número de camadas convolucionais e a função de ativação. Como dito no Capítulo 3, a CNN processa os dados por meio dos campos receptivos locais que se encontram na camada convolutacional, no algoritmo é possível alterar o tamanho deste campo mudando o valor do *kernel size*, o que possibilitou a autonomia dos testes para verificar qual melhor se encaixaria no problema da BCI-SSVEP.

O número de filtros e camadas convolucionais também aumentaram o desempenho da BCI-SSVEP, porém, o aumento desses dois parâmetros influenciou diretamente no tempo de processamento, já que a Rede Neural se torna cada vez mais profunda. A profundidade da CNN implica no refinamento dos dados já que a inclusão de múltiplas camadas faz com que o algoritmo resolva o problema por divisão e conquista, isso explica o aumento do desempenho (PONTI; COSTA, 2017). Por último, a função de ativação usada para obter o desempenho da Tabela 5 foi a *Tanh*.

4.2.4 CNN com os Dados Brutos

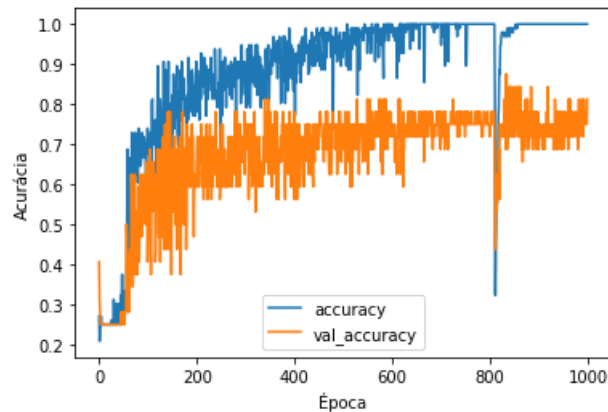
Este teste foi feito considerando os dados de entrada da rede CNN sem a etapa de extração de atributos pelo método da FFT, mas com a aplicação da filtragem CAR e o janelamento do sinal em 3 segundos. Primeiramente foi realizada a sintonia da rede que consistiu em ajustar os parâmetros de números de neurônios, quantidade de camadas convolucionais e tamanho da janela de convolução. Estes ajustes foram necessários para reduzir o *overfitting*, que nada mais é que um sobreajuste dos parâmetros da rede, ocasionando uma ótima acurácia durante o treinamento, mas péssimos resultados durante a validação. A Figura 19 abaixo exemplifica a ocorrência deste fenômeno em um dos testes realizados com o indivíduo 32.

Figura 18 – *Overfitting* - Sobreajuste dos dados.



Como pode-se observar no início da simulação até a época 200 as duas curvas crescem juntas, porém a partir deste ponto ocorre a divergência ficando os valores na validação em torno de 37%, o que não é satisfatório para o desenvolvimento da BCI. Existem diversas técnicas que tratam o *overfitting*, a mais importante delas é o ajuste da taxa de aprendizado (ou *learning rate*) do otimizador, como é explicado por (GOODFELLOW, 2016). Basicamente o que deve ser feito é reduzir ao máximo o número de parâmetros contudo, após o teste de sintonia da rede especificamente para o indivíduo 32 foi possível obter um resultado mais próximo do ideal, como pode ser observado na Figura 19.

Figura 19 – Acurácia do Indivíduo 32.



Foram realizadas 10 partições aleatórias para se obter uma média de acertos mais representativa. A acurácia de discriminação das quatro classes foi de 81,25% com desvio padrão de 10,82%. A topologia da rede implementada foi de 128 neurônios na camada de entrada, 1 camada de *maxpooling* com *poolsize* igual a 64, 4 camadas de convolução e 1 camada densa com 4 neurônios de saída. A mesma configuração foi testada para outros voluntários da base, porém não foram obtidos resultados satisfatórios o que se leva a observar que em um caso de implementação prática de um sistema operando em tempo real, por exemplo, seria preciso sintonizá-lo para o usuário algo que é bastante usual e aceitável. Entretanto, dos testes preliminares realizados, foi observado que o desempenho do sistema sempre ficava inferior ao obtido pela rede neural CNN recebendo na sua entrada as características extraídas pela FFT.

5 Conclusão

Durante o trabalho foram implementadas todas as etapas referentes ao processamento dos sinais cerebrais EEG, i.e., filtragem, janelamento e extração de atributos, além de duas técnicas de classificação. A filtragem permitiu remover parte dos ruídos do sinal, melhorando a SNR. O janelamento do sinal em 3 s permitiu realizar o processamento em um segmento do sinal que corresponderia a quantidade de dados em uma situação de funcionamento em tempo de execução do sistema, ou seja, ao tempo necessário de exposição do indivíduo ao estímulo. A extração de atributos via FFT permitiu selecionar as principais características do sinal, reduzindo a dimensionalidade do problema de classificação ao sintetizar as informações relevantes para a identificação do estímulo ao qual o voluntário focava sua atenção.

A implementação e análise das técnicas de classificação, que era o principal objetivo deste trabalho, permitiu solucionar o problema de classificação da BCI-SSVEP. Ambos foram capazes de distinguir e identificar os estímulos visuais apresentando desempenho satisfatório - CL 47% e CNN 87%. A CNN foi a que apresentou melhor acurácia entre os métodos, porém esta técnica é computacionalmente mais custosa que o Classificador Linear.

Também foi possível obter informações sobre as funcionalidades de implementação das Redes Neurais comparando-a com um método de classificação tradicional, baseado no Método dos Mínimos Quadrados e identificando os prós e contras de cada técnica de classificação. A análise da CNN com os dados brutos mostrou que se tratando de séries temporais a CNN tem um pouco de dificuldade de generalizar os dados para todos os voluntários da base de dados, devido a presença de muitos ruídos. Sendo necessário passar o sistema por uma sintonia para reduzir o número de parâmetros gerado pelo sobreajuste. De maneira que em, propostas futuras seria interessante simular mais configurações da CNN com os dados brutos, usando uma outra metodologia de análise, por exemplo, através de espectogramas do sinal cerebral.

Anexos

Linear Classifier vs. MLP for four-class discrimination in EEG signals

Larissa R. Azevedo¹, Victor D. Nascimento¹, Harlei M. A. Leite^{1,2}, Sarah N. Carvalho^{1,2}

¹Federal University of Ouro Preto (UFOP) – Brazil

²Brazilian Institute of Neuroscience and Neurotechnology (BRAINN) – Brazil

Introduction: Brain-computer interface (BCI) is a closed-loop system that establishes communication between a brain and an application of control or communication, making this innovative human-computer interaction useful in the development from assistive technologies to entertainment. In this study, we have analyzed and compared the performance of a BCI based on steady state visually evoked potentials (SSVEP) employing two approaches of classification - a linear discriminant and a neural network.

Materials and Methods: The brain signal was acquired by electroencephalography (EEG) from 32 health subjects, with 8 repetitions for every visual stimuli of 12 s each. Four visual stimuli were projected by a monitor at frequencies 6, 10, 12 and 15 Hz [Leite *et al.*, 2018]. The preprocessing consisted of application of a notch filter at 60 Hz, a passband filter (5 – 100Hz) and a digital filter based on Common Average Reference (CAR). The features of the EEG were extracted in windows of 2 s without overlap employing the Fast Fourier Transform (FFT) algorithm, considering the magnitude of signal at frequencies of stimulation. The classification was realized by a linear classifier based on the least squares (LC) and a Multi-Layer Perceptron (MLP) with two hidden layers, each with 512 neurons and the output layer with four neurons.

Results: Tables 1 and 2 show the average hit rate applying the 10-cross-fold validation for each subject using the BCI system with the linear classifier and the MLP, respectively. The average accuracy was 50.3±17.2 % for the LC and 52.1±16.4% for the neural network.

Table 1 - Accuracy employing the LC.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Accuracy (%)	86	19	27	49	36	37	28	83	35	62	65	88	29	53	24	58
Subject	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Accuracy (%)	55	40	57	64	45	25	31	39	33	43	69	53	53	92	40	91

Table 2 - Accuracy employing the MLP.

Subject	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Accuracy (%)	84	24	27	47	35	36	31	79	40	67	67	88	29	52	26	60
Subject	17	18	19	20	21	22	23	24	25	26	27	28	29	30	31	32
Accuracy (%)	56	51	48	70	60	29	47	43	32	53	75	43	44	91	43	89

Discussion/Conclusion: The results show that both classifiers are able to discriminate the four classes of visual stimuli with an accuracy of about 50%. The average performance employing the MLP is slightly higher than employing the linear classifier. However, the difference of accuracy is within the standard deviation presented. The high standard deviation is due to the large variation in performance among subjects, ranging from over 90 for Subject 30 to practically the randomization rate for the Subject 2. But, the performance of the subjects between the techniques was practically the same, the subjects with the highest variability (9 or more) were: 18, 21, 23 and 26 (with better performance with MLP) and 19, 28 and 29 (with better performance using linear). Thus, the simplicity of the linear classifier and its low computational cost make this approach more attractive for BCI-SSVEP systems.

Acknowledgements: The authors thank CNPq, Fapesp and UFOP for the financial support.

Reference: [1] LEITE, H. M. A., et al. Analysis of user interaction with a brain-computer interface based on steady-state visually evoked potentials: case study of a game. Computational intelligence and neuroscience, v. 2018, 2018.

Preliminary investigation of convolutional neural networks for BCI-SSVEP

Larissa R. Azevedo¹, Harlei M. A. Leite^{1,2} and Sarah N. Carvalho^{1,2}

¹Federal University of Ouro Preto (UFOP) – Brazil

²Brazilian Institute of Neuroscience and Neurotechnology (BRAINN) – Brazil

Introduction: The Steady-State Visually Evoked Potentials (SSVEP) is a popular paradigm used to conceive Brain-Computer Interfaces (BCI). In the non-invasive approach, the brain signal can be acquired by electroencephalography (EEG). Applications of a BCI-SSVEP range from assistive technology to entertainment, such as moving a wheelchair or augmented reality. In this study, we have analyzed and compared the performance of a BCI-SSVEP employing as classifier two scenarios of a Convolutional Neural Network (CNN): (1) employing features extracted from EEG signals and (2) using the raw EEG signals.

Materials and Methods: The database employed [1] consisting of EEG signals of one healthy subject. During the EEG data acquisition, the subject was exposed to four visual stimuli frequencies (6, 10, 12 and 15 Hz), with 8 repetitions of 12 s for each frequency. The preprocessing consisted of the application of a notch filter at 60 Hz, a passband filter (5 – 100Hz), the segmentation of signal in 3 s window and the filtering with a digital filter based on Common Average Reference. In the first classification scenario, the CNN was fed with the magnitudes of Fast Fourier Transform (FFT) extracted in evoked frequencies. In the second scenario, the CNN was fed with the raw data.

Results: Fig.1 shows the structure of CNNs in both scenarios evaluated. Table 1 shows the average accuracy of BCI-SSVEP for both scenarios evaluated, considering 10-cross-fold-validation and convolution with the Kernel at 1D. In Scenario 1 the average accuracy was 87,5±9,3 % and in Scenario 2 was 70,25±12,2%.

Fig. 1 – CNN topologies in Scenarios 1 and 2.

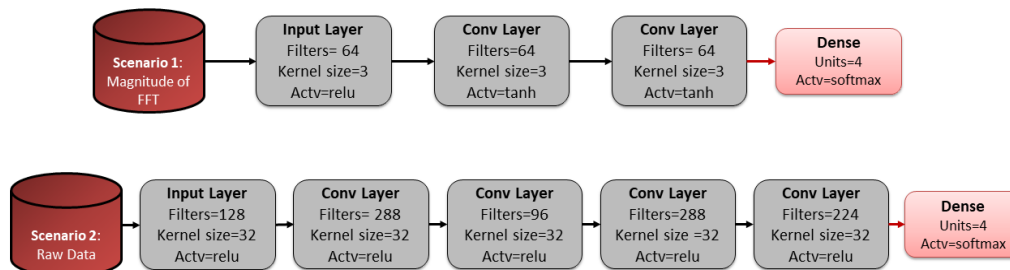


Table 1 – Average performance and standard deviation of BCI-SSVEP in the two scenarios.

Frequency (Hz)/ Average accuracy (%)	6	10	12	15	Total
Scenario 1	81.4	98.6	75.7	94.3	87.5±9.3
Scenario 2	84.3	61.7	80.0	55.0	70.3±12.2

Discussion/Conclusion: Our preliminary results showed that models of both scenarios can discriminate the four visual stimuli with satisfactory accuracy (higher than 70%). However, the CNN of Scenario 1 seems to be able to get slightly better accuracy than the CNN of Scenario 2. The strategy of including feature extraction, in our case via FFT, looks attractive in some applications and should be further investigated by expanding the analysis to other subjects of the database.

Acknowledgments: The authors thank FAPESP and UFOP for the financial support.

Reference: [1] LEITE, H. M. A., et al. Analysis of user interaction with a brain-computer interface based on steady-state visually evoked potentials: case study of a game. Computational intelligence and neuroscience, v. 2018, 2018.

Referências

- AZEVEDO, L. R.; LEITE, H. A.; CARVALHO, S. N. Preliminary investigation of convolutional neural networks for bci-ssvep. *Journal of Epilepsy and Clinical Neurophysiology*, Corpo, v. 26, p. 27–28, 2022. 3
- AZEVEDO, L. R. et al. Linear classifier vs. mlp for four-class discrimination in eeg signals. *Journal of Epilepsy and Clinical Neurophysiology*, Corpo, v. 26, p. 27–28, 2020. 3
- BEAR, M. F. e. a. *Neurociências: desvendando o sistema nervoso*. [S.l.]: Artmed editora, 2002. 4, 16
- BISHOP, C. M. *Pattern recognition and machine learning*. [S.l.]: springer, 2006. 10
- BRAGA, A. d. P. *Redes neurais artificiais: teoria e aplicações*. [S.l.]: Livros Técnicos e Científicos, 2000. 11
- CARVALHO, S. N. et al. Comparative analysis of strategies for feature extraction and classification in ssvep bcis. *Biomedical Signal Processing and Control*, Elsevier, v. 21, p. 34–42, 2015. 3, 6, 16
- COSTA, T. B. d. S. et al. Desenvolvimento de interface cérebro-computador baseada em potenciais evocados visualmente em regime estacionário. [sn], 2015. 18
- DE JESUS JR, e. a. B. J. Análise do potencial evocado visualmente em sinais cerebrais para implementação de bcis. [sn], p. 1–5, 2016. 6, 17
- FLAUZINO, R. A. e. A. *Redes neurais artificiais para engenharia e ciências aplicadas curso prático*. São Paulo: Artliber, 2010. 11, 12
- GOMES, A. T. *Telecomunicações: transmissão e recepção am/fm*. São Paulo, 2001. 18
- GOODFELLOW, I. e. a. *Deep learning*. [S.l.]: MIT press, 2016. 1, 6, 8, 14, 31
- GRAIMANN, B. e. a. Brain–computer interfaces: A gentle introduction. In: *Brain-computer interfaces*. [S.l.]: Springer, 2009. p. 1–27. 2, 3, 6
- GULLI, A.; KAPOOR, A.; PAL, S. *Deep learning with TensorFlow 2 and Keras: regression, ConvNets, GANs, RNNs, NLP, and more with TensorFlow 2 and the Keras API*. [S.l.]: Packt Publishing Ltd, 2019. 25
- HAYKIN, S. S. et al. *Neural networks and learning machines/Simon Haykin*. [S.l.]: New York: Prentice Hall,, 2009. 11
- LEITE, H. M. d. A. et al. Analysis of user interaction with a brain-computer interface based on steady-state visually evoked potentials: case study of a game. *Computational intelligence and neuroscience*, Hindawi, v. 2018, 2018. 6, 16, 17
- LUDWIG, K. A. et al. Using a common average reference to improve cortical neuron recordings from microelectrode arrays. *Journal of neurophysiology*, American Physiological Society, v. 101, n. 3, p. 1679–1689, 2009. 17

- MULLER, S. T. e. a. Interface cérebro-computador baseada em potenciais evocados visuais para comandar cadeira de rodas robótica autônoma. *Journal of Medical and Biological Engineering*, 2010. 6, 26
- NIELSEN, M. A. *Neural networks and deep learning*. [S.l.]: Determination press San Francisco, CA, 2015. v. 25. 6
- OMS, O. Relatório mundial sobre a deficiência. *Banco Mundial. São Paulo*, p. 334, 2011. 4
- OPPENHEIM, A. V. *Sinais e sistemas*. [S.l.]: Prentice-Hall, 2010. 13
- OPPENHEIM, A. V.; SCHAFER, R. W. *Discrete-time signal processing*. [S.l.]: Pearson Education, 2014. 18
- PONTI, M.; COSTA, B. Tópicos em gerenciamento de dados e informações. *SBC, Uberlândia*, p. 63–88, 2017. 31
- RAMPAZZO, W. et al. Aprendizado profundo e interfaces cérebro-computador. [sn], 2019. 7, 15
- RAVI, A. e. a. A convolutional neural network for enhancing the detection of ssvep in the presence of competing stimuli. In: IEEE. *2019 41st Annual International Conference of the IEEE Engineering in Medicine and Biology Society (EMBC)*. [S.l.], 2019. p. 6323–6326. 6, 26
- TAVARES, T. V. et al. Análise de técnicas de filtragem espacial em sistemas bci-ssvep. *Anais do Simpósio de Processamento de Sinais da Unicamp*, 2015. 17, 19
- THEODORIDIS, S.; KOUTROUMBAS, K. *Pattern Recognition, Edition*. [S.l.]: Academic Press, fourth edition Edition, 2009. 9