

# Classificação de sinais cerebrais sob a ótica dos *Large Language Models*: aplicação na reabilitação clínica

Victor Hugo Gonçalves Gomes de Novais

Instituto de Pesquisas Tecnológicas do Estado de São Paulo  
São Paulo, Brasil  
victor.novais@ensino.ipt.br

Adriano Galindo Leal

Instituto de Pesquisas Tecnológicas do Estado de São Paulo  
São Paulo, Brasil  
leal@ipt.br

## ABSTRACT

The recent advancements in natural language processing (NLP) have introduced novel artificial intelligence models for data classification, extending their scope to analyzing brain signals acquired via electroencephalogram (EEG). Among these developments, the transformer architecture, which has become available in recent years, has provided researchers with a powerful model to explore and evaluate its capabilities in various EEG-related studies, including developing new assistive devices tailored for individuals with impaired communication skills. This work leverages the transformer model to classify P300 event-related potentials on publicly available EEG data, aiming to benchmark its accuracy against established algorithms documented in literature. Upon conducting the case study, the results reveal that the transformer achieves a noteworthy accuracy rate of 95%, indicating its viability as a classifier for P300-based spellers.

**Keywords:** EEG, P300, Brain-computer interface, Assistive devices, LLM, Artificial Intelligence

## 1 INTRODUÇÃO

Desde os primeiros experimentos de eletroencefalografia (EEG) conduzidos por Hans Berger na década de 1920 [1], os sinais cerebrais vêm tendo destaque nos mais diversos campos da ciência, incluindo, recentemente, o da computação. Atualmente, sabemos que estes impulsos elétricos gerados pelo cérebro podem ser processados e utilizados para os mais diversos fins, incluindo o controle de dispositivos assistivos como próteses, soletradores digitais e cadeiras de rodas motorizadas [2], [3].

Por meio da chamada interface cérebro-máquina (ICM), os computadores passaram a ter o papel de analisar estes sinais e decodificá-los para manejo destes dispositivos, com especial aplicação na área de reabilitação clínica [4]. Neste sentido, pessoas portadoras de distúrbios neuromotores podem ser mais independentes e, conseqüentemente, ter uma melhor qualidade de vida. Atualmente, apesar da facilidade

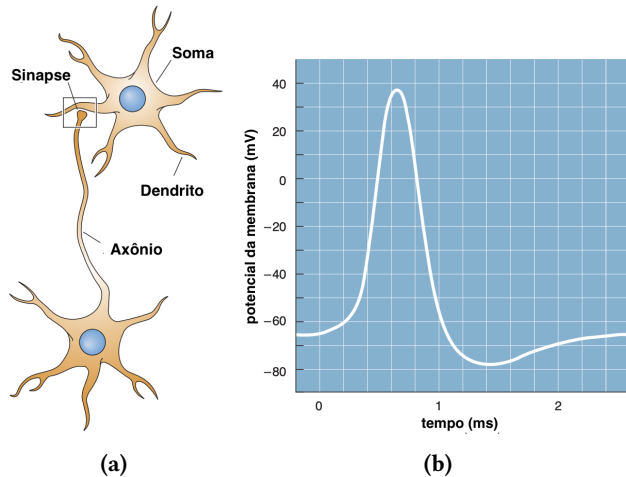
Reconhecemos o apoio financeiro e institucional do Instituto de Pesquisas Tecnológicas do Estado de São Paulo e da Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) através dos processos n°s #2019/01664-6 #2020/09850-0 #2017/50343-2 e #2021/11905-0.

em capturar e analisar os sinais cerebrais por meio do EEG (tal fato consolidado pelos avanços tecnológicos na área de neurociência), um dos atuais desafios enfrentados é a correta classificação para a posterior transformação na intenção desejada [3]. Nesta perspectiva, ao analisar os esforços científicos ao longo das décadas, é possível perceber que diversas iniciativas foram propostas para a resolução deste problema, incluindo a aplicação de técnicas alinhadas à inteligência artificial (IA), em especial com o uso de *machine learning* e *deep learning*. Entre os principais métodos utilizados para realizar a classificação dos sinais cerebrais, podemos citar, entre outros: *linear discriminant analysis* (LDA) [5], [2], [6], *support vector machine* (SVM) [6], *random forest* [7] e redes neurais convolucionais (CNN) [8]. Embora estas técnicas possam suprir os requisitos de acurácia necessários, a recente evolução da área de processamento de linguagem natural, em especial com os *transformers* [9], trouxe à luz uma nova maneira de conseguir lidar com as especificidades do sinal de EEG, permitindo capturar relações temporais com maior facilidade através do mecanismo de atenção [3].

Entre os diversos tipos de sinais que podem ser emitidos pelo cérebro, o P300 é um dos mais relevantes em pesquisas [5], [10], [11]. Este potencial evocado por eventos ocorre aproximadamente 300 milissegundos após um fenômeno incomum ser detectado pelo córtex visual no meio de ordinários. Devido a essa característica, o P300 é comumente utilizado nos chamados soletradores digitais, no qual o paciente é exposto a uma matriz de caracteres que piscam continuamente e deve focar sua atenção na letra desejada, que elicitava este tipo de sinal quando acende na tela do dispositivo.

Os soletradores digitais permitem que pessoas portadoras de severas distúrbios neuromotores, como a esclerose lateral amiotrófica (ELA) e a síndrome do encarceramento, possam se comunicar apenas com o cérebro, que continua suas tarefas cognitivas normais mesmo com a paralisia total do corpo.

A principal contribuição acadêmica deste trabalho reside na aplicação de conceitos oriundos da área de NLP, mais especificamente os *Large Language Models* (LLM), na área da reabilitação clínica, com o objetivo de validar se o uso da arquitetura *Transformer Neural Network* pode ser utilizada na classificação do sinal P300. Dessa forma, espera-se ampliar o conhecimento da efetividade desta nova técnica e compará-la aos classificadores tradicionalmente empregados na literatura.



**Figura 1.** a) Estruturas principais do neurônio. b) Potencial de ação ao longo do tempo. Adaptado de [12].

## 2 FUNDAMENTAÇÃO TEÓRICA

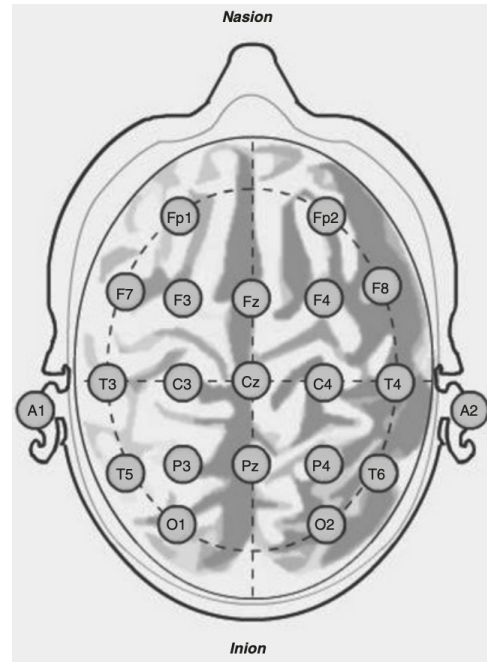
### 2.1 Potencial de ação e EEG

Os neurônios localizados no cérebro são os responsáveis por transmitir informações em formato de impulsos elétricos, também conhecidos pelo termo "potencial de ação". Por meio de técnicas e equipamentos diversos, como o EEG, estes potenciais são passíveis de captura e análise para entender o comportamento do cérebro durante um determinado cenário.

A geração do impulso elétrico é originária de uma intrincada cadeia de ações bioquímicas que ocorrem no neurônio, produzindo uma diferença de potencial da ordem de milivolts (mV). Um neurônio em repouso possui um potencial de repouso por volta de  $-65$  mV, que é alterado pelo início de um estímulo sináptico, despolarizando (tornando-se mais positivo) para aproximadamente  $+40$  mV e posteriormente repolarizando de volta pra seu estado de repouso [12]. Após o impulso percorrer o neurônio através de seu axônio, uma nova sinapse é feita com o dendrito do próximo neurônio, criando uma extensa rede de tráfego de informações em formato elétrico [12].

Atualmente, há diversos métodos para a leitura da atividade cerebral, destacando-se, principalmente, a eletroencefalografia e a ressonância magnética. O EEG é um dos métodos mais simples em termos de usabilidade, praticidade e custo, além de ser não-invasivo, de forma que apenas um capacete com eletrodos embutidos é vestido na cabeça do paciente. Embora o EEG possua estas características positivas, ele sofre de alguns empecilhos, como a vulnerabilidade à ruídos e a atenuação do sinal devido ao crânio.

Os grupos de neurônios localizados na parte mais externa do cérebro (córtex) são os responsáveis pelos potenciais de ação detectados na eletroencefalografia, de modo que os eletrodos posicionados no escalpo do paciente podem reconhecer a mudança de voltagem ao longo do tempo. Estes



**Figura 2.** Posicionamento dos eletrodos de EEG no sistema internacional 10-20. Extraído de [13].

eletrodos são comumente posicionados de acordo com o sistema internacional 10-20, que delimita as regiões corretas para cada um de forma padronizada. Esta distribuição é necessária pois a atividade cerebral gravada pelo EEG varia significativamente em cada área [13].

No sistema 10-20, os eletrodos com números pares são localizados na parte direita do escalpo, enquanto que os ímpares estão na parte esquerda. Na linha sagital média há 3 eletrodos simbolizados com a letra "z", indicando a posição "zero". Cada letra indica a área cerebral imediatamente abaixo do eletrodo, sendo: Fp (Frontopolar), F (Frontal), T (Temporal), C (Central), P (Parietal) e O (Occipital). Além disso, comumente há 2 eletrodos localizados nas orelhas e que servem como sinal de referência (A1 e A2) [13]. O esquemático da distribuição dos eletrodos está destacado na figura 2.

O sinal capturado pelos eletrodos é composto por um conjunto de frequências de onda e possíveis ruídos, como a rede elétrica, interferências eletromagnéticas do ambiente e contaminações originárias do movimento muscular e/ou ocular. Deste modo, a etapa de pré-processamento é obrigatória para que se possa analisar as características desejadas.

Este artigo irá se basear exclusivamente em dados provenientes de leituras de EEG, devido à ampla oferta deste tipo de dado e pela facilidade de análise.

### 2.2 Potenciais relacionados a eventos e o P300

Existe um conjunto de eventos específicos que elicitam padrões característicos no EEG, por exemplo, ao reconhecer

um rosto ou ao ser exposto a estímulos auditivos e/ou visuais. Estes padrões são conhecidos como potenciais relacionados a eventos (ERP na sigla em inglês) e são utilizados em dispositivos de ICM. Cada tipo de evento gera uma resposta cerebral diferente com características bem definidas e, entre as principais respostas, pode-se citar: *slow cortical potentials* (SCP), *sensorimotor rhythm* (SMR) e o P300 [14].

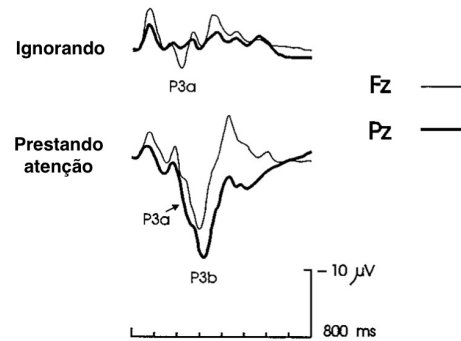
Em termos de aplicação na reabilitação clínica, estes potenciais relacionados a eventos podem ser utilizados em dispositivos assistivos, pois dependem apenas do poder cognitivo dos pacientes [14]. Em especial, pacientes portadores da síndrome do encarceramento são capazes de usufruir extensamente de dispositivos como os soletradores digitais, que podem se basear na elicitação do P300.

O P300, chamado também de P3, é uma das diversas respostas elicitadas pelo cérebro em razão de um evento específico e é um dos mais importantes para avaliação da atividade cognitiva [15]. Para essa resposta ser identificada no cérebro, o usuário deve ser submetido a uma série de eventos comuns e alguns poucos incomuns, de forma que a percepção deste evento raro dispare o P300. Por essa característica, o P300 também é conhecido na literatura como "*oddball paradigm*".

Os primeiros trabalhos relacionados à detecção do P300 datam da década de 1960, com os experimentos conduzidos por Chapman e Bragdon em 1964 [16] e, em 1965, por Sutton et al. [17]. Em ambos os experimentos, os pesquisadores encontraram uma alta deflexão positiva do sinal após aproximadamente 300 milissegundos da exposição a um evento (nestes casos, a exibição de números, sons e luzes), indicando uma resposta cognitiva relacionada a estes estímulos. Por conta destas características, essa resposta foi nomeada posteriormente como P300, indicando a positividade ("P") e o tempo médio de latência do sinal [18].

Uma das premissas para elicitar esse tipo de resposta é a capacidade de atenção e engajamento do usuário, que deve estar consciente e ciente dos eventos raros que irão ocorrer. Neste sentido, a atenção que cada usuário consegue oferecer difere, de forma que cada pessoa pode prestar menos ou mais atenção de acordo com seu estado mental [19]. Outros fatores também impactam na geração, amplitude e/ou latência do P300, como a idade e a presença de desordens neurais [18]. A figura 3 demonstra uma típica resposta P300 no cenário de atenção ativa, além da comparação com o sinal padrão quando o evento é ignorado.

Atualmente, sabe-se que o P300 é independente do tipo de estímulo aplicado, sujeitando-se apenas à raridade do evento. Sinais visuais, auditivos e somatosensoriais elicitam respostas P300 similares, caso esses se baseiem na mesma tarefa lógica. Ao analisar potenciais relacionados a eventos, é importante notar que um sinal pode conter diversas respostas concorrentes derivadas de outros fenômenos, como os conhecidos N2, P4, N400 [19], entre diversos outros que precisam ser filtrados corretamente.



**Figura 3.** Exibição da resposta P300 em duas situações: com o usuário ignorando o evento ou prestando atenção ativamente. Nota-se a maior amplitude do sinal após aproximadamente 300 milissegundos quando há atenção, representado por "P3b". O registro foi feito a partir de dois eletrodos centrais (Fz e Pz), localizados na área frontal e parietal. Observa-se, também, que a amplitude é maior no eletrodo parietal. Adaptado de [18].

### 2.3 Soletradores digitais

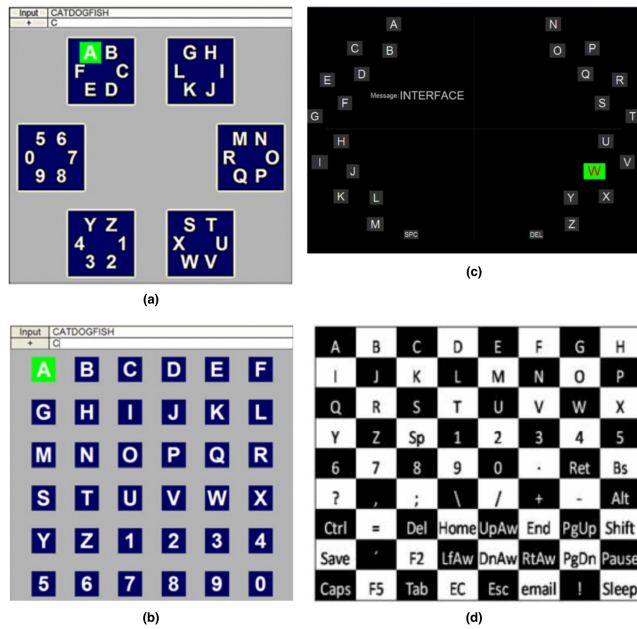
O P300 pode ser o meio utilizado para o funcionamento dos soletradores digitais, que pertencem à categoria de dispositivos assistidos relacionados à reabilitação de comunicação.

A primeira iniciativa em utilizar o sinal P300 em seleção de itens exibidos no monitor de um computador foi feita em 1988, por Farwell e Donchin [5]. Os mesmos pesquisadores foram também responsáveis pelo primeiro experimento utilizando um soletrador digital utilizando o chamado paradigma linha/coluna, no qual os voluntários eram expostos a 36 caracteres alfanuméricos dispostos em uma matriz 6×6. Os participantes deviam prestar atenção em 1 dos caracteres enquanto cada linha e coluna piscava em uma ordem randômica. Em uma rodada de 12 piscadas (uma para cada linha e coluna), a letra escolhida piscava apenas 2 vezes (uma para a linha e uma para a coluna). Estes 2 eventos raros no meio de 12 possíveis, era o que elicitava a resposta P300 [5]. A figura 4 demonstra um voluntário do estudo utilizando o soletrador digital proposto.

Um dos principais desafios enfrentados na utilização do soletrador digital é a velocidade de comunicação. De acordo com os estudos conduzidos por Wolpaw et al. [2], esta taxa varia entre 1 e 2 palavras por minuto, o que pode ser um processo moroso para os usuários. Para a melhoria deste aspecto, alguns novos paradigmas de exibição de caracteres foram propostos como, por exemplo: agrupamento [20], soletrador de caractere único em matriz [20] ou lateralizado [21], matriz em xadrez [22], uso de imagens [23], entre outros. A figura 5 exhibe as propostas de alguns paradigmas.



**Figura 4.** Voluntário utilizando capacete de EEG em frente a uma tela contendo 36 caracteres. Extraído de [5].



**Figura 5.** a) Paradigma de agrupamento. Extraído de [20]. b) Paradigma de caractere único matricial. Extraído de [20]. c) Paradigma de caractere único lateralizado. Extraído de [21]. d) Paradigma de matriz em xadrez. Extraído de [22]

### 2.4 Transformers

Para que o EEG seja utilizado em dispositivos assistivos como o soletrador digital, é necessário que haja uma etapa de classificação do sinal, com a intenção de identificar se ele é ou não do tipo P300. A partir desta informação, o sistema é capaz de cruzar a elicitação do sinal com o tempo em que ocorreu, resultando na escolha da letra desejada.

A classificação proposta neste artigo se baseia na arquitetura *transformers* que, dentre os avanços na área de processamento de linguagem natural, ocupa lugar de destaque. Anunciada em 2017, esta arquitetura apresenta uma solução para lidar com longas entradas de dados, de forma a preservar o contexto e manter a memória do que foi processado por meio do uso do mecanismo de atenção [9].

Mediante a estes avanços, diversas ferramentas foram desenvolvidas para processamento de texto e, a mais expoente delas, o ChatGPT [24] se tornou a principal vitrine desta tecnologia. Embora a arquitetura *transformer* seja primariamente utilizada para tarefas que envolvam NLP, ela também pode ser aproveitada em outras atividades como, por exemplo, a classificação de sinais cerebrais [3].

Com a característica de retenção de contexto, os *transformers* conseguem capturar relações temporais presentes no sinal do EEG, o que é importante para conseguir extrair características do sinal que dependam do tempo. Modelos de *transformers* integrados com EEG já foram utilizados para classificação de emoções, classificação de fala imaginada e classificação de estágios do sono [25].

De acordo com Song et al. os modelos baseados em *transformers* possuem grande capacidade de retenção de informação em largas janelas de tempo, mas não são qualificados para aprender características locais, que também são necessárias para a classificação do EEG. Uma das soluções apresentadas por Song et al. para resolver este problema é utilizar, antes da aplicação do *transformer*, uma rede neural convolucional (CNN) [3]. Este ponto será investigado nos trabalhos futuros dos autores deste artigo pois, em face dos avanços recentes na área de IA, a aprendizagem dos LLMs está mais capacitada.

## 3 METODOLOGIA EXPERIMENTAL

A metodologia seguida por este trabalho é o estudo de caso, no qual um conjunto de dados público será utilizado para analisar o desempenho da classificação pelo *transformer*.

O artigo originário dos dados públicos obtidos usou os modelos *Bayesian Linear Discriminant Analysis* (BLDA) e *Fisher’s Linear Discriminant Analysis* (FLDA) para detectar o P300 [23]. Neste sentido, este trabalho se difere pela aplicação de outro modelo, com o intuito de, com a mesma base de dados, explorar uma nova maneira de classificação.

### 3.1 Conjunto de dados

O *dataset* utilizado neste estudo de caso é originário dos experimentos de Hoffman et al. [23], que foram conduzidos em uma população de 5 portadores de deficiência e 4 indivíduos saudáveis. Para cada um deles, foi exibida uma matriz contando com 6 imagens (televisão, telefone, lâmpada, porta, janela e rádio) que piscavam em ordem randômica com intervalo de 400 ms entre os estímulos. Cada *flash* de imagem

durava 100 ms e durante os 300 ms restantes, nenhuma outra piscava [23].

O EEG foi capturado a uma taxa de amostragem de 2048 Hz a partir de 32 eletrodos posicionados no escalpo no padrão internacional 10-20. Todos os portadores de deficiência utilizavam cadeira de rodas mas possuíam diferentes habilidades de comunicação e controle muscular. Devido a uma grave hipofonia e flutuações no nível de alerta, a comunicação com o voluntário 5 foi de grande dificuldade, o que acarretou na remoção de seus dados do conjunto. Os outros 4 indivíduos eram estudantes de doutorado recrutados do laboratório de Hoffman et al. [23].

Cada indivíduo participou de 4 sessões, sendo 2 realizadas em cada dia. Cada sessão era composta de 6 rodadas, uma para cada imagem. A duração média das rodadas foi de 1 minuto, enquanto que a duração de uma sessão foi de aproximadamente 30 minutos, contando com o tempo de preparação dos equipamentos [23].

### 3.2 Objetivo

Utilizando o conjunto de dados descrito, o objetivo deste estudo de caso é comparar a efetividade da arquitetura *transformer* em relação a algoritmos de classificação do sinal P300 comumente utilizados na literatura (SVM, LDA, SWLDA e CNN). Ao final, espera-se concluir se a acurácia atingida pelo *transformer* é equiparável aos algoritmos tradicionais. A análise será feita para cada indivíduo separadamente, pois, como mencionado anteriormente, o sinal P300 pode variar de forma significativa em função das características biofisiológicas de cada pessoa.

### 3.3 Pré-processamento

Como indicado anteriormente, os dados provenientes de EEG precisam passar por filtros com a intenção de extrair o sinal desejado. Para realizar a etapa de pré-processamento, bem como todas as outras etapas posteriores, será utilizada a linguagem de programação *Python* em conjunto com a biblioteca MNE [26], criada especialmente para lidar com dados originários de magnetoencefalografia (MEG) e EEG.

Como primeira etapa, é necessário analisar os artefatos (ruídos) presentes no sinal do EEG. Estes podem ser divididos em dois tipos:

- **Fisiológicos:** são os artefatos relacionados ao movimento muscular (principalmente da cabeça e do pescoço), ocular (mudança do campo de visão ou ato de piscar) e, mais raramente, batidas do coração e respiração.
- **Não-fisiológicos:** são os artefatos relacionados ao ambiente no qual o EEG está sendo registrado, como a interferência eletromagnética proveniente da rede elétrica.

Antes de efetuar a remoção dos artefatos não fisiológicos, é necessário verificar a densidade espectral de potência (PSD) no domínio da frequência. Dessa forma, será possível

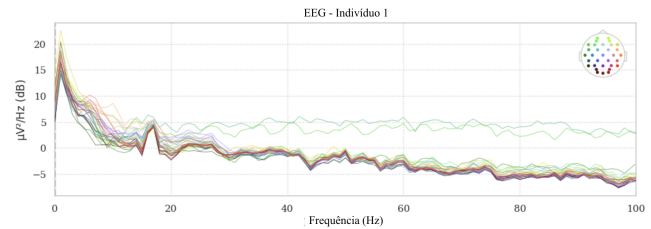


Figura 6. Densidade espectral de potência do indivíduo 1

detectar possíveis anomalias eletromagnéticas no sinal. A figura 6 demonstra o PSD para o indivíduo 1.

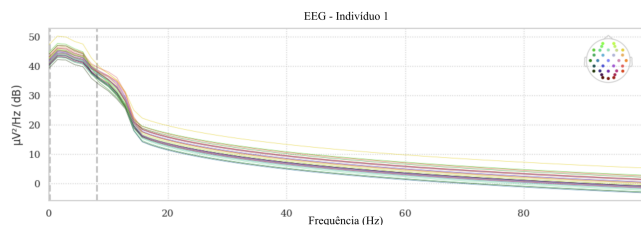
A interferência eletromagnética aparece, na Suíça (local de obtenção dos dados [23]), na faixa de 50 Hz (no Brasil, este valor é de 60 Hz), e também na forma de harmônicos (múltiplos da frequência de 50 Hz). Como é possível observar na figura 6 (para o indivíduo 1), não há ruídos oriundos de interferências, desse modo não há artefatos não-fisiológicos a serem removidos. Esta mesma análise foi feita para os 8 indivíduos presentes no conjunto de dados.

Para os artefatos fisiológicos há diferentes técnicas de remoção. Um dos modos de efetuar esse passo é removendo os canais indesejados para o estudo e que são os mais afetados pelos movimentos oculares (Fp1 e Fp2). Além disso, o sinal P300 aparece com maior potência na região centro-parietal do cérebro, o que permite que alguns canais específicos sejam selecionados para aprimorar o desempenho do modelo de classificação. Para este estudo, os canais foram escolhidos com base nos resultados obtidos por Krusienski et al. [27]: Oz, Fz, Cz e Pz. A partir dessa seleção de canais, o impacto dos artefatos fisiológicos foi reduzido.

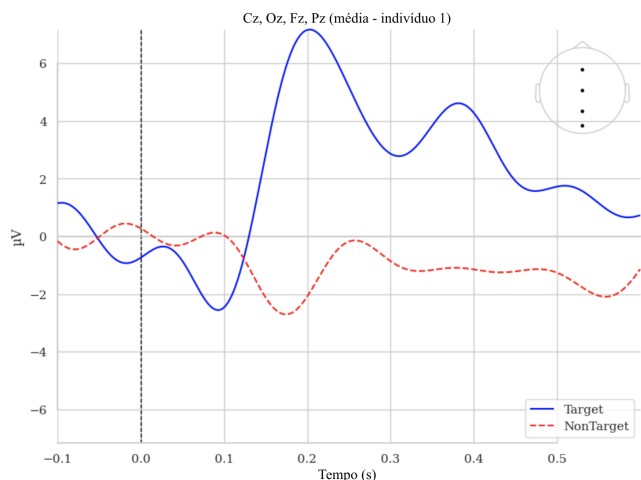
O próximo passo da etapa de pré-processamento é efetuar o filtro de frequência. Este filtro visa destacar bandas de frequência específicas associadas a diferentes atividades cerebrais, como ondas delta (0 - 4 Hz), theta (4 - 8 Hz), alpha (8 - 13 Hz), beta (13 - 30 Hz) e gama (> 30 Hz). Ao aplicar filtros de frequência adequados, a qualidade dos dados EEG é aprimorada, facilitando a análise e interpretação dos padrões cerebrais. De acordo com estudos anteriores, o P300 está intimamente ligado às bandas delta e theta, de forma que uma boa janela de filtro de frequência é entre 0 Hz e 8 Hz [28]. A figura 7 demonstra o PSD do indivíduo 1 após a aplicação deste filtro.

### 3.4 Segmentação

Para que seja possível classificar o EEG, é necessário dividir o sinal contínuo em segmentos temporais menores, chamados de épocas (*epochs*). Cada *epoch* é rotulada de acordo com a condução do experimento, marcando qual evento que ocorreu para cada segmento: *Target* (quando o indivíduo gera o sinal P300 com base na letra/imagem indicada pelos pesquisadores) e *Non-Target* (estímulo de controle não relacionado à letra/imagem alvo).



**Figura 7.** PSD do indivíduo 1 após aplicação do filtro de frequência



**Figura 8.** Epoch com a média dos eventos nos canais Cz, Oz, Fz e Pz (indivíduo 1). Para este indivíduo, o P300 aparece aproximadamente aos 200 ms.

Como visto anteriormente, o P300 costuma aparecer aproximadamente 300 ms após o estímulo, podendo variar para mais ou para menos de acordo com as características de cada pessoa. Por esse motivo, a segmentação foi feita em uma janela de 600 ms. A figura 8 exibe a média de cada evento, por epoch, dos 4 canais escolhidos (Oz, Fz, Cz e Pz), no indivíduo 1.

### 3.5 Criação de dados sintéticos

Devido ao baixo número de eventos (*target* e *non-target*) presentes nas *epochs* (aproximadamente 22 e 112, respectivamente, por indivíduo [23]), foi necessária a criação de dados sintéticos para que os modelos de classificação consigam generalizar de forma apropriada e, conseqüentemente, tenham uma melhor capacidade preditiva.

A estratégia adotada neste estudo envolveu a geração sintética de dados de *epochs* de EEG individualizadas. Utilizando como base as *epochs* de EEG de cada indivíduo, a técnica manipulou os dados existentes aplicando fatores aleatórios de escala de amplitude, deslocamento temporal e inclusão de ruídos. Esses parâmetros aleatórios foram ajustados para

**Tabela 1.** Resultados da classificação (percentual de acurácia)

	CNN	SVM	LDA	SWLDA	Transformer
Ind. 1	96,3	94,3	92,2	88,0	97,5
Ind. 2	97,2	91,1	93,3	83,0	93,9
Ind. 3	97,0	92,0	96,3	93,7	94,2
Ind. 4	96,6	94,8	92,1	93,3	95,1
Ind. 6	96,3	95,3	94,3	93,9	93,4
Ind. 7	97,4	96,3	95,9	92,0	96,8
Ind. 8	97,5	96,5	95,6	93,0	94,6
Ind. 9	96,3	95,3	95,0	93,7	94,5
Média	96,8	94,4	94,3	91,3	95,0

criar novos conjuntos de dados sintéticos que preservassem as características dos sinais de EEG de cada indivíduo.

Ao final da criação destes conjuntos de dados sintéticos, cada indivíduo permaneceu com, aproximadamente, 1.800 eventos *target* e 2.100 *non-target*.

## 4 CLASSIFICAÇÃO

Após o pré-processamento e a ampliação do conjunto de dados, a etapa de classificação foi realizada utilizando os seguintes algoritmos: CNN (com o modelo EEGNet [29]), SVM, LDA, SWLDA e *Transformer Neural Network* (com o modelo EEGConformer [3]). O desempenho dos modelos foi medido através da métrica de acurácia, que é um indicador fundamental que avalia a precisão obtida, ou seja, a habilidade do modelo em acertar as classes dos dados.

A tabela 1 exibe os resultados obtidos para cada um dos algoritmos e indivíduos, além da média total entre indivíduos.

## 5 LIMITAÇÕES

Entre as limitações deste estudo estão:

- **Conjunto de dados:** como visto, a quantidade de eventos originalmente disponíveis é insuficiente para conseguir treinar modelos de redes neurais com boa capacidade de generalização, o que tornou necessária a geração de dados sintéticos;
- **Engenharia de atributos:** a classificação foi executada após o filtro de frequência e a seleção de canais, porém a aplicação de mais técnicas de seleção de características do sinal (*feature engineering*), tanto no domínio do tempo, quanto no da frequência, pode aprimorar a classificação.

## 6 TRABALHOS FUTUROS

Os autores deste estudo pretendem conduzir dois novos experimentos: o primeiro, no primeiro semestre de 2024, com o objetivo de criar um protótipo de interface de soletrador com imagens de atividades comuns do dia-a-dia e testá-lo com os próprios autores como voluntários, e um segundo

em 2025, com o objetivo de conduzir um estudo com pacientes portadores de doenças neurodegenerativas utilizando o mesmo protótipo. Ressalta-se que o segundo experimento será regulado pelo CEP (Comitê de Ética em Pesquisa), Órgão Federal para monitoramento de pesquisas envolvendo humanos.

O objetivo de ambos os estudos é aprofundar o conhecimento no funcionamento do sinal P300 e avaliar o desempenho da arquitetura de redes neurais do tipo *transformer* como classificador, bem como entender o impacto desta tecnologia na vida de pessoas que necessitam de dispositivos assistivos para poder se comunicar.

## 7 CONCLUSÃO

A partir dos resultados obtidos no presente estudo, é possível perceber que o *Transformer Neural Network* atinge um desempenho competitivo em relação aos algoritmos comumente utilizados na classificação do P300 no sinal de EEG, ficando em segundo lugar geral com média da acurácia (95%) entre os 8 indivíduos do estudo, perdendo apenas para a CNN (96,8%).

Dessa forma, é factível afirmar que a utilização de algoritmos baseados em NLP podem, de fato, encontrar lugar em tarefas relacionadas à classificação de EEG, abrindo novas portas para a criação de dispositivos assistivos que possuam maior capacidade de inferência e possam levar maior qualidade de vida para as pessoas que buscam recuperar a capacidade de se comunicar com o mundo exterior.

## Agradecimentos

Os autores gostariam de expressar sua mais profunda gratidão e apreciação ao Anderson Ribeiro Correia, Liedi Legi Bariani Bernucci, Fabricio Araujo Mirandola, Maria Cristina Machado Domingues, Denis Bruno Virissimo, Alessandro Santiago dos Santos, Silvia Elisabete Ferrari, Eduardo Luiz Machado, Dirce A. Rosaboni, Terezinha e Maria A. Chinellato.

## Referências

- [1] Sumit Ghulyani, Yashasvi Pratap, Sumit Bisht, and Ravideep Singh. Brain computer interface boulevard of smarter thoughts. *International Journal of Advanced Research in Artificial Intelligence*, 1(7), 2012.
- [2] Jonathan R Wolpaw, Niels Birbaumer, Dennis J McFarland, Gert Pfurtscheller, and Theresa M Vaughan. Brain-computer interfaces for communication and control. *Clinical Neurophysiology*, 113(6):767–791, 2002.
- [3] Yonghao Song, Qingqing Zheng, Bingchuan Liu, and Xiaorong Gao. Eeg conformer: Convolutional transformer for eeg decoding and visualization. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 31:710–719, 2023.
- [4] Sergio Machado, Fernanda Araújo, Flávia Paes, Bruna Velasques, Marlo Cunha, Henning Budde, Luis Basile, Renato Anghinah, Oscar Arias-Carrión, Mauricio Cagy, Roberto Piedade, Tom Graaf, Alexander Sack, and Pedro Ribeiro. Eeg-based brain-computer interfaces: An overview of basic concepts and clinical applications in neurorehabilitation. *Reviews in the neurosciences*, 21:451–68, 01 2010.
- [5] F. Nijboer, E.W. Sellers, J. Mellinger, M.A. Jordan, T. Matuz, A. Furdea, S. Halder, U. Mochty, D.J. Krusienski, T.M. Vaughan, J.R. Wolpaw, N. Birbaumer, and A. Kübler. A p300-based brain-computer interface for people with amyotrophic lateral sclerosis. *Clinical Neurophysiology*, 119(8):1909–1916, 2008.
- [6] H. Mirghasemi, R. Fazel-Rezai, and M. B. Shamsollahi. Analysis of p300 classifiers in brain computer interface speller. In *2006 International Conference of the IEEE Engineering in Medicine and Biology Society*, pages 6205–6208, 2006.
- [7] Faraz Akram, Ahmed Alwakeel, Mohammed Alwakeel, Mohammad Hijji, and Usman Masud. A symbols based bci paradigm for intelligent home control using p300 event-related potentials. *Sensors*, 22(24), 2022.
- [8] W-L Mao, H I K Fathurrahman, Y Lee, and T W Chang. Eeg dataset classification using cnn method. *Journal of Physics: Conference Series*, 1456(1):012017, jan 2020.
- [9] Ashish Vaswani, Noam Shazeer, Niki Parmar, Jakob Uszkoreit, Llion Jones, Aidan N. Gomez, Lukasz Kaiser, and Illia Polosukhin. Attention is all you need. *CoRR*, abs/1706.03762, 2017.
- [10] Reza Fazel-Rezai, Brendan Allison, Christoph Guger, Eric Sellers, Sonja Kleih, and Andrea Kübler. P300 brain computer interface: current challenges and emerging trends. *Frontiers in Neuroengineering*, 5, 2012.
- [11] T.M. Vaughan. Guest editorial brain-computer interface technology: a review of the second international meeting. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 11(2):94–109, 2003.
- [12] Mark F. Bear, Barry W. Connors, and Michael A. Paradiso. *Neuroscience: Exploring the Brain*. Wolters Kluwer, 4th edition, 2015.
- [13] O. Mecarelli. *Clinical Electroencephalography*. Springer International Publishing, 2019.
- [14] Andrea Kübler, Adrian Furdea, Sebastian Halder, Eva Maria Hammer, Femke Nijboer, and Boris Kotchoubey. A brain-computer interface controlled auditory event-related potential (p300) spelling system for locked-in patients. *Annals of the New York Academy of Sciences*, 1157(1):90–100, 2009.
- [15] Rui Zhong, Mengmeng Li, Qingling Chen, Jing Li, Guangjian Li, and Weihong Lin. The p300 event-related potential component and cognitive impairment in epilepsy: A systematic review and meta-analysis. *Frontiers in Neurology*, 10, 2019.
- [16] Robert M. Chapman and Henry R. Bragdon. Evoked responses to numerical and non-numerical visual stimuli while problem solving. *Nature*, 203:1155–1157, 1964.
- [17] Samuel Sutton, Margery Brearen, Joseph Zubin, and E. R. John. Evoked-potential correlates of stimulus uncertainty. *Science*, 150(3700):1187–1188, 1965.
- [18] TW Pictor. The p300 wave of the human event-related potential. *Journal of Clinical Neurophysiology*, 9(4):456–479, OCT 1992.
- [19] E. Donchin. *Cognitive Psychophysiology: Event-Related Potentials and the Study of Cognition: The Carmel Conferences Volume I*. Psychology Revivals. Taylor & Francis, 2022.
- [20] Jiahui Pan, Yuanqing Li, Zhenghui Gu, and Zhuliang Yu. A comparison study of two p300 speller paradigms for brain-computer interface. *Cognitive Neurodynamics*, 7, 12 2013.
- [21] Gabriel Pires, Urbano Nunes, and Miguel Castelo-Branco. Comparison of a row-column speller vs. a novel lateral single-character speller: Assessment of bci for severe motor disabled patients. *Clinical Neurophysiology*, 123(6):1168–1181, 2012.
- [22] G. Townsend, B.K. LaPallo, C.B. Boulay, D.J. Krusienski, G.E. Frye, C.K. Hauser, N.E. Schwartz, T.M. Vaughan, J.R. Wolpaw, and E.W. Sellers. A novel p300-based brain-computer interface stimulus presentation paradigm: Moving beyond rows and columns. *Clinical Neurophysiology*, 121(7):1109–1120, 2010.
- [23] Ulrich Hoffmann, Jean-Marc Vesin, Touradj Ebrahimi, and Karin Dierschens. An efficient p300-based brain-computer interface for disabled subjects. *Journal of Neuroscience Methods*, 167(1):115–125, 2008. Brain-Computer Interfaces (BCIs).

## XV Computer on the Beach

10 a 13 de abril de 2024, Balneário Camboriú, SC, Brasil

---

- [24] OpenAI. Chatgpt, 2023. Acessado em 07/11/2023.
- [25] Jin Xie, Jie Zhang, Jiayao Sun, Zheng Ma, Liuni Qin, Guanglin Li, Huihui Zhou, and Yang Zhan. A transformer-based approach combining deep learning network and spatial-temporal information for raw eeg classification. *IEEE Transactions on Neural Systems and Rehabilitation Engineering*, 30:2126–2136, 2022.
- [26] Alexandre Gramfort, Martin Luessi, Eric Larson, Denis A Engemann, Daniel Strohmeier, Christian Brodbeck, Roman Goj, Mainak Jas, Teon L Brooks, Lauri Parkkonen, and Matti Hamalainen. Mne software for processing meg and eeg data. *NeuroImage*, 86:446–460, 2014.
- [27] D.J. Krusienski, E.W. Sellers, D.J. McFarland, T.M. Vaughan, and J.R. Wolpaw. Toward enhanced p300 speller performance. *Journal of Neuroscience Methods*, 167(1):15–21, 2008. Brain-Computer Interfaces (BCIs).
- [28] Kyra Kannen, Behrem Aslan, Cindy Boetzel, Christoph S. Herrmann, Silke Lux, Helena Rosen, Benjamin Selaskowski, Annika Wiebe, Alexandra Philipsen, and Niclas Braun. P300 modulation via transcranial alternating current stimulation in adult attention-deficit/hyperactivity disorder: A crossover study. *Frontiers in Psychiatry*, 13, 2022.
- [29] Vernon J. Lawhern, Amelia J. Solon, Nicholas R. Waytowich, Stephen M. Gordon, Chou P. Hung, and Brent J. Lance. Eegnet: A compact convolutional network for eeg-based brain-computer interfaces. *CoRR*, abs/1611.08024, 2016.