,

CENTRO UNIVERSITÁRIO FEI

RODRIGO PRIOR BECHELLI

INTERFACE CÉREBRO–COMPUTADOR PARA CLASSIFICAÇÃO DE BANCO DE IMAGENS DE ACERVOS MUSEOLÓGICOS

São Bernardo do Campo

2018

RODRIGO PRIOR BECHELLI

INTERFACE CÉREBRO–COMPUTADOR PARA CLASSIFICAÇÃO DE BANCO DE IMAGENS DE ACERVOS MUSEOLÓGICOS

Tese de Doutorado apresentada ao Centro Universitário FEI para obtenção do título de Doutor em Engenharia Elétrica, orientado pela Profa. Dra. Maria Claudia Ferrari de Castro.

São Bernardo do Campo

2018

Prior Bechelli, Rodrigo.
Interface Cérebro-Computador para classificação de banco de imagens de acervos museológicos / Rodrigo Prior Bechelli. São Bernardo do Campo, 2018.
104 f. : il.

Tese - Centro Universitário FEI.
Orientadora: Prof.ª Dra. Maria Claudia Ferrari de Castro.

1. Interface cérebro-computador. 2. Recuperação de Imagens Baseada em Conteúdo. 3. Eletroencefalograma. 4. Potencial Evocado. 5. Aprendizado de Máquina. I. Ferrari de Castro, Maria Claudia, orient. II. Título.

Elaborada pelo sistema de geração automática de ficha catalográfica da FEI com os dados fornecidos pelo(a) autor(a).

centro universitário FEI

APRESENTAÇÃO DE TESE
ATA DA BANCA EXAMINADORA

Programa de Pós-Graduação Stricto Sensu em Engenharia Elétrica

Doutorado

PGE-10

Aluno: Rodrigo Prior Bechelli **Matrícula:** 514204-7

Título do Trabalho: Interface cérebro-computador para classificação de banco de imagens de acervos museológicos.

Área de Concentração: Processamento de Sinais e Imagens

Orientador: Prof.ª Dr.ª Maria Claudia Ferrari de Castro

Data da realização da defesa: 28/05/2018

ORIGINAL ASSINADA

Avaliação da Banca Examinadora

São Bernardo do Campo, 28 / 05 / 2018.

MEMBROS DA BANCA EXAMINADORA	
Prof.ª Dr.ª Maria Claudia Ferrari de Castro	Ass.: _____
Prof. Dr. Aldo Artur Belardi	Ass.: _____
Prof. Dr. Fabio Gerab	Ass.: _____
Prof. Dr. João Ricardo Sato	Ass.: _____
Prof.ª Dr.ª Esther Luna Colombini	Ass.: _____

A Banca Examinadora acima-assinada atribuiu ao aluno o seguinte:

APROVADO ☒ REPROVADO ☐

VERSÃO FINAL DA TESE

ENDOSSO DO ORIENTADOR APÓS A INCLUSÃO DAS RECOMENDAÇÕES DA BANCA EXAMINADORA

_____

Aprovação do Coordenador do Programa de Pós-graduação

_____

Prof. Dr. Carlos Eduardo Thomaz

À minha esposa Regina e meu filho Otávio.

AGRADECIMENTOS

À Profa. Dra. Maria Claudia Ferrari de Castro, por todo o apoio, dedicação e energia para a realização de todas as etapas necessárias desse projeto. Pela amizade, conselhos, aulas e ensinamentos em todos esses anos.

À Profa. Dra. Maria Beatriz Borba Florenzano, por todo o apoio, pelos diálogos e pela dedicação em todos esses anos de trabalho no Labeca e no Museu de Arqueologia e Etnologia da USP.

À Profa. Dra. Maria Isabel Fleming, pela presença e sugestões propostas na banca do Exame de qualificação. Agradeço ainda pela sua amizade e pelos conselhos.

Ao Prof. Dr. Plinio Thomaz Aquino Jr., por me aconselhar sempre que preciso e pela ajuda com o desenvolvimento desta pesquisa.

Ao Prof. Dr. Aldo Belardi pelo apoio desde a época do mestrado e pelas opiniões sempre sinceras. A contribuição com as wavelets foi essencial para a realização dessa pesquisa.

Ao Prof. Dr. Fábio Gerab, pelo apoio e as conversas sobre estatística.

Ao Prof. Dr. Renato Camargo Giacomini por sempre manter as portas abertas para mim e pela confiança.

Ao Prof. Dr. Carlos Eduardo Thomaz por me auxiliar no desenvolvimento do projeto de pesquisa e por me aceitar no programa de pós-graduação.

Aos amigos Rodrigo Piva, Renato Barelli, Douglas Rizzo e André Mendes, por compartilharem as alegrias e desafios da ciência.

À Ana Paula Tauhyl, uma grande amiga, pela ajuda sempre que precisei.

Ao Labeca e todos os seus integrantes, pela disposição e ajuda essencial para a execução desse projeto.

Aos colegas André Santos, Julia Arrabaça, Rodrigo Dória, por fazerem parte da primeira turma do curso de mestrado na FEI e compartilharem mais essa etapa importante da minha vida.

À equipe da pós-graduação: Adriana, Márcia e Ricardo sempre nos auxiliando da melhor maneira possível.

Ao pessoal que cuida do laboratório de elétrica da FEI, principalmente o Acácio que sempre me auxiliou.

À equipe do CGI da FEI, pela eficiência, atenção e disposição.

Aos funcionários da Biblioteca da FEI, pela paciência e vontade de ajudar, sempre.

Aos amigos da FEI da turma de 98'.

Ao Rodolfo, Odete e Bia, pela compreensão, ajuda, apoio e pelo carinho.

Ao meu pai Arnaldo (in memoriam), que sempre esteve ao meu lado em todas as iniciativas.

À minha irmã Carina, pelo imenso carinho, amor e incentivo para a realização deste trabalho.

Ao Otávio, meu filho, que é a energia que impulsiona as minhas conquistas.

À Regina, pela orientação, por seu amor, compreensão, força e equilíbrio necessários para a realização desse trabalho. Por cuidar de nosso filho e me dar o tempo para realizar esta pesquisa.

E por fim, mas não menos importante. À minha mãe Jandira que sempre acredita nos empreendimentos da minha vida. Este trabalho é dedicado à você.

Meus agradecimentos à FEI, MAE, USP e CAPES.

E a todos que não foram citados, mas que auxiliaram direta ou indiretamente para a realização deste trabalho.

“Assim falou e Atena, a deusa, olhos-azuis,
não desobedeceu. Dos píncaros do Olímpo
desceu precipitada e rápida alcançou
as agílimas naves dos Aqueus. Ali
topou com Odisseu, equiparado a Zeus
em argúcia. Parado, nem tocava a nau
belo–convés, escura: a dor no coração.”

Homero (CAMPOS, 2003)

RESUMO

Nas últimas décadas, o volume de informações armazenadas digitalmente cresce em taxas nunca antes vistas e boa parte destas informações são geradas pelo armazenamento amplo de imagens, vídeos e documentos digitalizados. Um dos sistemas mais robustos e de baixo custo para a aquisição de sinais cerebrais, aplicados em sistemas de Interface Cérebro–Computador (ICC), são os dispositivos de Eletroencefalograma (EEG). O processamento dos sinais de EEG, através de inteligência artificial, utilizando um conjunto de *wavelets* para uma análise espacial e temporal dos sinais cerebrais, apresenta uma alternativa para aplicações de Recuperação de Imagens Baseada em Conteúdo (RIBC) como forma de classificação e ordenação de um conjunto de imagens. Neste trabalho é proposta uma metodologia para classificar e representar uma relação entre sinais cerebrais e conjuntos de imagens apresentadas para diferentes grupos de voluntários, utilizando um método de captura de sinais cerebrais via EEG, como indicativo de uma compreensão de mais alto nível sobre a imagem ou grupo de imagens. Através de um conjunto de testes em ambiente controlado, via sinais de EEG, utilizando um conjunto de respostas cerebrais de potencial evocado em um grupo de voluntários, foi identificado imagens alvo entre uma série de imagens. O trabalho apresenta resultados significativos de classificação de sinais cerebrais para a aplicação em bancos de dados de imagens, uma comparação de aplicação entre diferentes tipos de grupos de usuários (especialistas e não especialistas) e diferentes bancos de dados de imagens de pesquisa científica.

Keywords: Interface cérebro-computador, Recuperação de Imagens Baseada em Conteúdo, Eletroencefalograma, Potencial Evocado, Aprendizado de Máquina

ABSTRACT

In the last decades, the volume of digitally stored information has grown at rates never seen before and much of this information is generated by the extensive storage of images, videos and digitized documents. One of the most robust and low-cost systems for the acquisition of brain signals, applied in *Brain Computer Interface* (BCI) systems, are the EEG devices. The processing of the EEG signals, through artificial intelligence, using a set of wavelets for a spatial and temporal analysis of brain signals, presents an alternative for *Content Based Image Retrieval* (CBIR) applications as classification and ranking to a set of images. This work proposes a methodology to classify and represent a relation between brain signals and sets of images presented to different groups of volunteers, using a method of capturing brain signals via EEG, as indicative of a higher level comprehension about the image or group of images. Through a set of controlled environment tests, via EEG signals, using a set of evoked potential brain responses in a group of volunteers, target images were identified among a series of images. The work presents significant results of classification of brain signals for application in image databases, a comparison of application between different types of user groups (specialists and non-specialists) and different databases of scientific research images.

Keywords: Brain Computer Interface, Content Based Image Retrieval, EEG, Evoked Related Pontential, Machine Learning

LISTA DE ILUSTRAÇÕES

Ilustração 1 – Nausitoo e cidades pesquisadas pelo Laboratório de Estudos sobre a Cidade Antiga (LABECA) 18
Ilustração 2 – Modelo de Recuperação de Imagens Baseada em Conteúdo. 22
Ilustração 3 – Técnicas de redução de descontinuidades semânticas. 24
Ilustração 4 – Exemplos de imagens da base de dados Banco de dados de imagens de pesquisa Caltech 101 (Caltech 101). 26
Ilustração 5 – Exemplos de imagens da base de dados Banco de dados de imagens de pesquisa do Projeto Labeca (Nausitoo). 27
Ilustração 6 – Perfil do cérebro humano e suas regiões. 28
Ilustração 7 – Modelo de neurônio para a transmissão do impulsos elétricos. 29
Ilustração 8 – Disposição de eletrodos conectados no escalpo da cabeça em relação ao cérebro. 30
Ilustração 9 – Etapas de conversão de um sinal analógico para digital. 31
Ilustração 10 – Representação da montagem de eletrodos 10–20 com a apresentação padrão para 75 eletrodos com as respectivas posições de eletrodos de referência: (a) e (b) representações tridimensionais das posições no couro cabeludo, e (c) configuração dos eletrodos em uma representação planificada do couro cabeludo, os eletrodos marcados em cinza apresentam uma configuração para 20 eletrodos e 2 eletrodos de referência. 33
Ilustração 11 – Decomposição de um sinal de EEG (canal O1) nas diferentes bandas clássicas de frequência: delta (1–3 Hz), teta (4–6 Hz), alfa (7–13 Hz), beta (15–25 Hz) e gama (35–45 Hz). As escalas no eixo y são diferentes com o objetivo de realçar os sinais de alta frequência e baixa amplitude. 34
Ilustração 12 – Principais componentes de um ERP. 35
Ilustração 13 – Relação entre ondas geradoras de ERP. 35
Ilustração 14 – Acurácia de um teclado virtual por Potencial Evocado positivo que ocorre aproximadamente 300*ms* após o estímulo (P300) em função do número de apresentações de estímulos visuais. 37
Ilustração 15 – Comparação da Transformada de Wavelets em relação Transformadas de *Fourier*. (a) sinal estacionário e (b) sinal não-estacionário. 41
Ilustração 16 – Comparação entre as resoluções de tempo e frequência. 42
Ilustração 17 – Decomposição de 4 níveis de Wavelet. 43
Ilustração 18 – Estimativa dos coeficientes da regressão de *ridge*. 46
Ilustração 19 – Fluxo de aquisição de dados de EEG e método de classificação de imagens. 56
Ilustração 20 – Método de RSVP. 60

Ilustração 21 – Média de sinal potencial evocado P300 em cada eletrodo para o Voluntário05 com a base Caltech 101: a) Visão de todos os eletrodos e b) Detalhe de um único eletrodo com as marcações de tempo a partir da apresentação da imagem. 64
Ilustração 22 – Níveis de decomposição adotados para cada wavelet. 65
Ilustração 23 – Valores de $F1$, por voluntário, agrupados por família de wavelet. 67
Ilustração 24 – Boxplot do valor de $F1$ agrupado por família de wavelet. 68
Ilustração 25 – Teste de multicomparação de Tukey entre todas as famílias de wavelets. . 68
Ilustração 26 – Comparação dos resultados do classificador para a família DB32 para todos os voluntários. 70

LISTA DE TABELAS

Tabela 1 – Valores de uma série temporal com os cálculos de média e diferença (em negrito). 40
Tabela 2 – Hardware utilizado. 49
Tabela 3 – Detalhes dos métodos de RSVP utilizados. 50
Tabela 4 – Banco de Dados de Imagens utilizado. 51
Tabela 5 – Detalhes sobre os voluntários utilizados. 52
Tabela 6 – Detalhes sobre os diferentes processamentos de EEG utilizados. 53
Tabela 7 – Detalhes sobre os voluntários. 58
Tabela 8 – Dimensão dos vetores por característica. 66
Tabela 9 – Tabela de comparação dos resultados do classificador para a família DB32 para todos os voluntários. 71

LISTA DE ABREVIATURAS

ACC	Acurácia.
ANOVA	Análise de Variância.
AR	Aprendizado por reforço.
AUC	Área abaixo da curva Característica de Operação do Receptor (ROC).
BCI	*Brain Computer Interface*.
BIOR	Biortogonal.
Caltech 101	Banco de dados de imagens de pesquisa Caltech 101.
CBIR	*Content Based Image Retrieval*.
COIF	Coiflets.
DB	Daubechies.
DB32	Daubechies de ordem 32.
DFT	*Discrete Fourier Transform*.
DMEY	Aproximação discreta da wavelet de Meyer.
DWT	*Discrete Wavelet Transform*.
EEG	Eletroencefalograma.
ERP	*Event-Related brain Potential*.
FAPESP	Fundação de Amparo à Pesquisa do Estado de São Paulo.
FEI	Fundação Educacional Inaciana.
FN	*False Negatives*.
FP	*False Positives*.
HAAR	Haar.
HSD	*Honest Significant Difference*.
IA	inteligência artificial.
ICC	Interface Cérebro–Computador.
LABECA	Laboratório de Estudos sobre a Cidade Antiga.
MAE	Museu de Arqueologia e Etnologia.
ML	*machine learning*.
Nausitoo	Banco de dados de imagens de pesquisa do Projeto Labeca.
P300	Potencial Evocado positivo que ocorre aproximadamente 300*ms* após o estímulo.
PR	Precisão.
RBIO	Biortogonal Reversa.
RC	*Recall*.
RIBC	Recuperação de Imagens Baseada em Conteúdo.
ROC	Característica de Operação do Receptor.
RSVP	Apresentação Visual Serial Rápida, do inglês *Rapid Serial Visual Presentation*.

SISNEP	Sistema Nacional de Ética em Pesquisa.
SNR	*signal-to-noise ratio*.
STFT	*Short-Time Fourier Transform*.
SVM	*Support Vector Machines*.
SYM	Symlets.
TN	*True Negatives*.
TP	*True Positives*.
UMESP	Universidade Metodista de São Paulo.
USP	Universidade de São Paulo.
VEP	*Visual Evoked Potential*.

LISTA DE SÍMBOLOS

α	parâmetro de otimização do classificador de Ridge.
a	parâmetro de escalmento da *Wavelet*.
b	parâmetro de translação da *Wavelet*.
Cl^-	ânions de cloreto.
$F1$	resultado de F1.
K^+	íons de potássio.
Na^+	íons de sódio.
ψ	Wavelet mãe.
χ_c^2	valor crítico de *qui-quadrado*.
χ^2	teste de independência *qui-quadrado*.

SUMÁRIO

1 INTRODUÇÃO 17
1.1 MOTIVAÇÃO 18
1.2 OBJETIVOS DA PESQUISA 19
1.3 ESTRUTURA DO TRABALHO 20
2 RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO 21
2.1 COMPONENTES DE UM SISTEMA DE CBIR 21
2.2 DESCONTINUIDADES SEMÂNTICAS 23
2.3 BIOSINAIS EM APLICAÇÕES DE CBIR 24
2.4 BANCOS DE IMAGENS PARA COMPARAÇÃO DE CBIR 25
2.4.1 Caltech 101 26
2.4.2 Nausitoo 26
3 O CÉREBRO 28
3.1 SINAIS CEREBRAIS 29
3.2 ELETROENCEFALOGRAMA 30
3.3 SISTEMA 10–20 DE EEG 31
3.4 BANDAS DE FREQUÊNCIA 32
3.5 POTENCIAIS EVOCADOS 32
3.6 P300 E INTERFACE CÉREBRO-COMPUTADOR 36
3.7 PLATAFORMAS DE BCI 37
4 TRANSFORMADA DE WAVELETS 40
5 APRENDIZADO DE MÁQUINA 44
5.1 CLASSIFICADOR DE RIDGE 45
6 TRABALHOS CORRELATOS 48
6.1 CLASSIFICAÇÃO DE CBIR VIA EEG 48
6.2 HARDWARE DE AQUISIÇÃO DE EEG 48
6.3 MÉTODO DE APRESENTAÇÃO DAS IMAGENS 48
6.4 BANCOS DE DADOS DE IMAGEM 51
6.5 SOFTWARE DE BCI 52
6.6 VOLUNTÁRIOS E CARACTERÍSTICAS 52
6.7 PRÉ-PROCESSAMENTO E PROCESSAMENTO DOS DADOS 53
6.8 ANÁLISE CRÍTICA 54
7 MATERIAL E MÉTODOS 56
7.1 MATERIAL 57
7.1.1 Banco de dados de imagens 57
7.1.2 Aquisição de EEG 57
7.2 VOLUNTÁRIOS 58
7.3 MÉTODO 59

7.3.1 Procedimento experimental 59
7.3.2 Processamento dos dados 60
7.3.3 Avaliação dos sinais de P300 61
7.3.4 Classificação individual das imagens alvo 61
7.3.5 Seleção da família de wavelets 62
7.3.6 Comparação entre os grupos de especialistas e não especialistas 62
7.3.7 Comparação entre as bases de dados para o grupo de especialistas 62
7.3.8 Teste do classificador 63
7.4 COMITÊ DE ÉTICA EM PESQUISA CIENTÍFICA 63
8 RESULTADOS 64
8.1 AVALIAÇÃO DOS SINAIS DE P300 64
8.2 CLASSIFICAÇÃO INDIVIDUAL DAS IMAGENS ALVO 65
8.2.1 Seleção da família de wavelets 65
8.2.2 Comparação entre os grupos de especialistas e não especialistas 69
8.2.3 Comparação entre as bases de dados para o grupo de especialistas 69
8.2.4 Teste de generalização do classificador 70
9 DISCUSSÃO 72
10 CONCLUSÃO 76
11 TRABALHOS FUTUROS 77
REFERÊNCIAS 78
APÊNDICE A – ICELERA IBLUE52 PYTHON INTEGRATION 87
APÊNDICE B – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO 89
ANEXO A – RELATÓRIO DE APROVAÇÃO DO COMITE DE ÉTICA . 94
ANEXO B – CERTIFICADO DE CALIBRAÇÃO ICELERA IBLUE52 . . 99

1 INTRODUÇÃO

Nas últimas décadas, o volume de informações armazenadas digitalmente cresce em taxas nunca antes vistas e boa parte destas informações são geradas pelo armazenamento amplo de imagens, vídeos e documentos digitalizados (CHEN et al., 2006a).

Um dos maiores desafios de se trabalhar com grandes volumes de dados, é ampliar a forma de indexar, armazenar e extrair as informações armazenadas em arquivos que não contém textos escritos, utilizando assim diferentes tipos de tecnologias para ampliar a compreensão de como tratamos o conhecimento obtido por outras formas que não a descrita por texto.

Soluções tradicionais relacionadas ao tratamento das características inerentes às imagens, como por exemplo, metodologias de visão computacional e inteligência artificial são amplamente aplicadas no contexto de extração de informação destas imagens, porém, é também conhecido que estes tipos de abordagem trazem consigo a necessidade de avaliação das descontinuidades semânticas (DESERNO; ANTANI; LONG, 2009).

Uma alternativa é a extração de informações não descritivas a partir de sinais cerebrais como um caminho adicional para a compreensão das informações contidas em um conjunto de imagens, e determinar uma compreensão de mais alto nível sobre a imagem ou grupo de imagens (LIANG et al., 2010). Utilizando o retorno desta informação como fonte de recurso para a melhoria da busca individual de cada imagem e como base de conhecimento para um possível treinamento de um sistema de aprendizado de máquina (RUI; HUANG, 1998), buscando reduzir as descontinuidades existentes entre o que é visualizado e o que é percebido (LI et al., 2010).

Um dos sistemas mais robustos e de baixo custo para a aquisição de sinais cerebrais, aplicados em sistemas de ICC, são os dispositivos de EEG. Algumas vantagens que podem ser descritas sobre esse tipo de dispositivo são: a mobilidade em relação a dimensão dos equipamentos; o baixo consumo de energia que possibilita sua operação através de baterias; o uso de conexões sem fio que ampliam a gama de possíveis testes; a rápida curva de aprendizado relacionada a montagem do equipamento, onde em poucos ciclos de utilização é possível aprender a montar e capturar sinais; o uso de dispositivos de EEG de baixo custo e de fácil acesso tanto para a montagem quanto para a manutenção.

O processamento dos sinais de EEG, através de inteligência artificial, utilizando um conjunto de *wavelets* para uma análise espacial e temporal dos sinais cerebrais, apresenta uma alternativa para aplicações de RIBC como forma de classificação e ordenação de um conjunto de imagens.

Partindo de bancos de dados de imagens de pesquisa, e posteriormente utilizando um banco de dados de imagens arqueológicas, com o auxílio de sinais de EEG, o projeto busca, através da obtenção dos indicadores de reposta cerebral à um estímulo, desenvolver uma interface cérebro computador para classificar estes sinais.

1.1 MOTIVAÇÃO

Este projeto de pesquisa teve sua concepção inicial em 2012 após um trabalho junto ao Laboratório de Estudos sobre a Cidade Antiga (LABECA) [1], criado em 2006 no Museu de Arqueologia e Etnologia (MAE) da Universidade de São Paulo (USP) como um projeto temático financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP). O laboratório tem como um de seus objetivos pesquisar novas formas de interpretação para os vestígios materiais relativos ao uso e à organização do espaço na Grécia antiga (ALDROVANDI; KORMIKIARI; HIRATA, 2011).

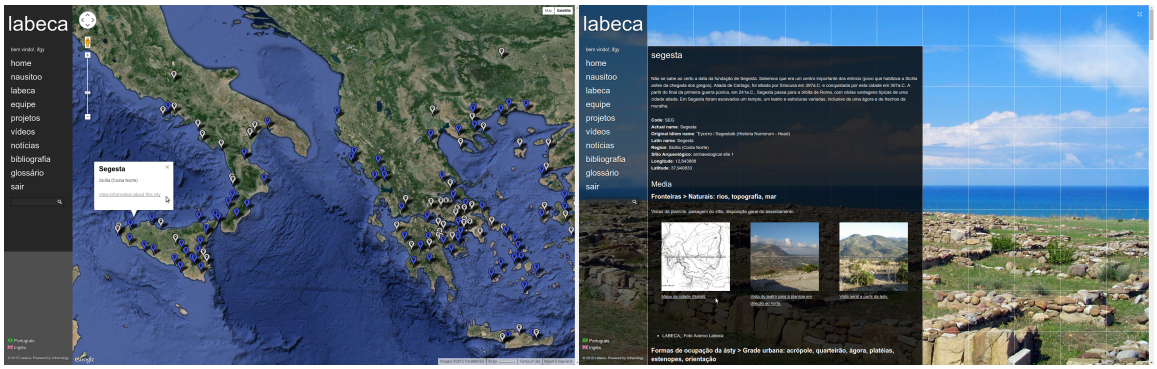
O primeiro site do LABECA foi ao ar em 2007, fornecendo informações do projeto para a comunidade científica e em paralelo ocorria a organização dos primeiros bancos de dados de imagens. Este site foi revisto no ano de 2008, para incluir uma seleção de imagens dos bancos de dados, desenvolvidos no ano anterior, e disponibilizar este material para o público em geral, mas principalmente para professores e pesquisadores.

Entre 2007 e 2009, o volume de imagens armazenadas e indexadas cresceu de aproximadamente 600 para 3500 arquivos. Nesse ponto, novas necessidades começaram a surgir no projeto, como por exemplo, utilizar o banco de dados por diferentes pessoas da equipe e uma abertura desses bancos, uma vez que alguns pesquisadores do grupo encontram-se em diferentes localidades do país e do exterior (FLORENZANO, 2009; 2010).

Atualmente, o banco de dados do LABECA (Figuras 1a e 1b) possui aproximadamente 7000 imagens de diversas fontes, como por exemplo: fotografias de campo, mapas, imagens de livros, páginas de obras de referência, entre outros tipos de informação visual da área pesquisada. A maior parte é classificada de forma manual pelos pesquisadores do projeto e agrupada pela posição dos sítios arqueológicos. Foi previsto no relatório de 2014 que o LABECA poderia atingir o volume de 14.000 imagens nos dois anos consecutivos (FLORENZANO, 2014).

Figura 1 – Nausitoo e cidades pesquisadas pelo LABECA

(a) Geolocalização das cidades (b) Detalhe das imagens e descrições



Fonte: Autor.

[1] http://labeca.mae.usp.br

Em 2011, foi proposto no projeto LABECA um novo campo de estudo que não havia sido previsto inicialmente. Esta nova linha de trabalho prevê o uso de imagens para construir o conhecimento sobre a antiguidade grega. O nome provisório dado a essa nova linha de pesquisa é *Arqueologia Visual* citado pela primeira vez em Aldrovandi, Kormikiari e Hirata (2011):

> "Para finalizar, queremos ainda chamar a atenção para um campo novo e que não havíamos previsto de início em nossos estudos sobre o ordenamento do espaço da Grécia antiga, que tem relação com a própria metodologia de organização reunida do Labeca. Trata-se do uso das imagens para a construção do conhecimento sobre a antiguidade grega. Esta é uma linha de pesquisa inédita que poderíamos definir provisoriamente como ***arqueologia visual***. A imagem em arqueologia tanto tem servido para ilustrar os estudos nessa área como para registrar achados e objetos. Antes, nos primórdios da nossa disciplina, o desenho preciso era indispensável. Em seguida, a fotografia científica passou a ocupar o lugar dos desenhos. Estes, entretanto, nunca foram abandonados para as reconstituições de pequenos ou de grandes objetos, incluindo-se aí esculturas, edifícios, plantas de cidades e assim por diante. Hoje em dia, recurso importante são as tomadas filmográficas. A imagem tem também assumido cada vez mais o papel de divulgadora de conhecimentos: filmes, DVDS e álbuns fotográficos são cada vez mais comuns, e o avanço tecnológico tem contribuído para uma divulgação de conhecimento científico arqueológico cada vez mais sofisticado. Por outro lado, já desde os anos de 1980, historiadores, inclusive da contemporaneidade, têm usado cada vez mais a imagem, estática ou não, como documento para o conhecimento do passado.
>
> Nossa experiência no Labeca, no entanto, tem apontado ainda uma outra via: maneiras de se construir o próprio conhecimento a partir do uso da imagem. A elaboração de um site, em que imagens têm um papel muito importante, e a realização de filmes no âmbito do Laboratório mostrou que ao fazer escolhas de imagens, ao editar um filme, estamos também construindo um tipo de conhecimento que é, por força, diferente daquele que elaboramos quando escrevemos. Esta é uma vertente que tem recebido nossa atenção, e que deverá ser aprofundada em etapa futura de nossas pesquisas."

O conceito de *arqueologia visual* traz implícita a necessidade de tratar de um problema observado nas últimas décadas. O volume de informações armazenadas digitalmente cresce em taxas nunca antes vistas e boa parte dessas informações corresponde a imagens, vídeos e documentos digitalizados (CHEN et al., 2006a).

Este trabalho proporcionará a ampliação do conceito de uso de imagens dentro da *arqueologia visual*, proposta anteriormente, através da aplicação de novas tecnologias. Dessa maneira, expande os horizontes de pesquisa do laboratório e contextualiza novas abordagens para obtenção de novos significados e compreensões para os materiais coletados pelo LABECA e seus pesquisadores.

1.2 OBJETIVOS DA PESQUISA

Neste trabalho é proposta uma metodologia para classificar e representar uma relação entre sinais cerebrais e conjuntos de imagens apresentadas para diferentes grupos de voluntários, utilizando um método de captura de sinais cerebrais via EEG e a extração de características via wavelets, como indicativo de uma compreensão de mais alto nível sobre a imagem ou grupo de imagens. Utilizando o retorno desta informação como fonte de recurso para a melhoria da

busca individual e como base de conhecimento para o treinamento de um sistema de aprendizado de máquina, busca-se reduzir as diferenças semânticas existentes entre o que é visualizado e o que é percebido.

Desta forma este trabalho pretende responder às seguintes perguntas:

a) Existe a possibilidade de traçar uma relação entre o comportamento dos sinais cerebrais e os resultados apresentados em diferentes conjuntos de voluntários (especialistas e não especialistas)?
b) Existe a possibilidade de traçar uma relação entre os sinais cerebrais e os resultados apresentados para os diferentes bancos de dados de imagens?
c) É possível desenvolver um modelo de teste e processamento de sinais EEG para complementar os métodos utilizados em sistemas de recuperação de imagens baseada em conteúdo para a área de arqueologia visual?
d) Através das informações capturadas via comportamento dos sinais cerebrais seria possível melhorar o desempenho (trabalhando descontinuidades) de um sistema de busca de imagens?
e) Criar uma nova proposta de processamento para a aplicação no segmento de arqueologia visual.

1.3 ESTRUTURA DO TRABALHO

Este trabalho está dividido nos seguintes capítulos:

a) O Capítulo 1 apresenta uma descrição introdutória sobre a motivação do trabalho e os aspectos relevantes para a pesquisa.
b) Os Capítulos 2, 3, 4 e 5 apresentam os fundamentos teóricos que serão utilizados ao longo do trabalho.
c) O Capítulo 6 apresenta um conjunto de trabalhos correlatos que relacionam EEG, inteligência artificial e recuperação de imagens baseadas em conteúdo.
d) O Capítulo 7 apresenta a metodologia, com todas as etapas, processos e equipamentos utilizados e desenvolvidos para a parte experimental da pesquisa.
e) Os capítulos 8, 9 e 10 apresentam os resultados da pesquisa, uma discussão sobre os temas abordados e suas conclusões.

2 RECUPERAÇÃO DE IMAGENS BASEADA EM CONTEÚDO

A Recuperação de Imagens Baseada em Conteúdo (RIBC), amplamente difundida em inglês como CBIR, é o resultado específico da busca de um conjunto de imagens, e o resultado semelhante ao objeto pesquisado (COX et al., 2000).

Uma das primeiras referências ao termo *Content Based Image Retrieval Systems* se dá no trabalho de Kato (1992). O termo foi definido para abranger a gestão de imagens em diversas áreas do conhecimento, como por exemplo: galerias de arte e gestão dos museus, projetos arquitetônicos e de engenharia, sensoriamento remoto e gestão dos recursos, sistemas de informação geográfica, gestão de banco de dados científicos, previsão do tempo e a aplicação da lei e investigação criminal.

Até então, existiam duas direções para abordar o assunto (GUDIVADA; RAGHAVAN, 1995), sendo a primeira, trabalhada como um sistema de banco de dados convencional, onde as informações eram extraídas manualmente e as consultas eram efetuadas seguindo uma representação de atributos pré-definidos para o contexto abordado. Já a segunda, sendo através da extração de aspectos e reconhecimento de objetos (LOWE, 1999) de forma automatizada aperfeiçoando a representação da primeira. Estes dois extremos apresentavam, e ainda apresentam, limitações, pois de um lado havia a capacidade humana de definir um modelo de banco de dados e do outro lado, um alto custo computacional para o processamento automático de todas as imagens de um conjunto. Daí, a necessidade constante de equilíbrio entre eles.

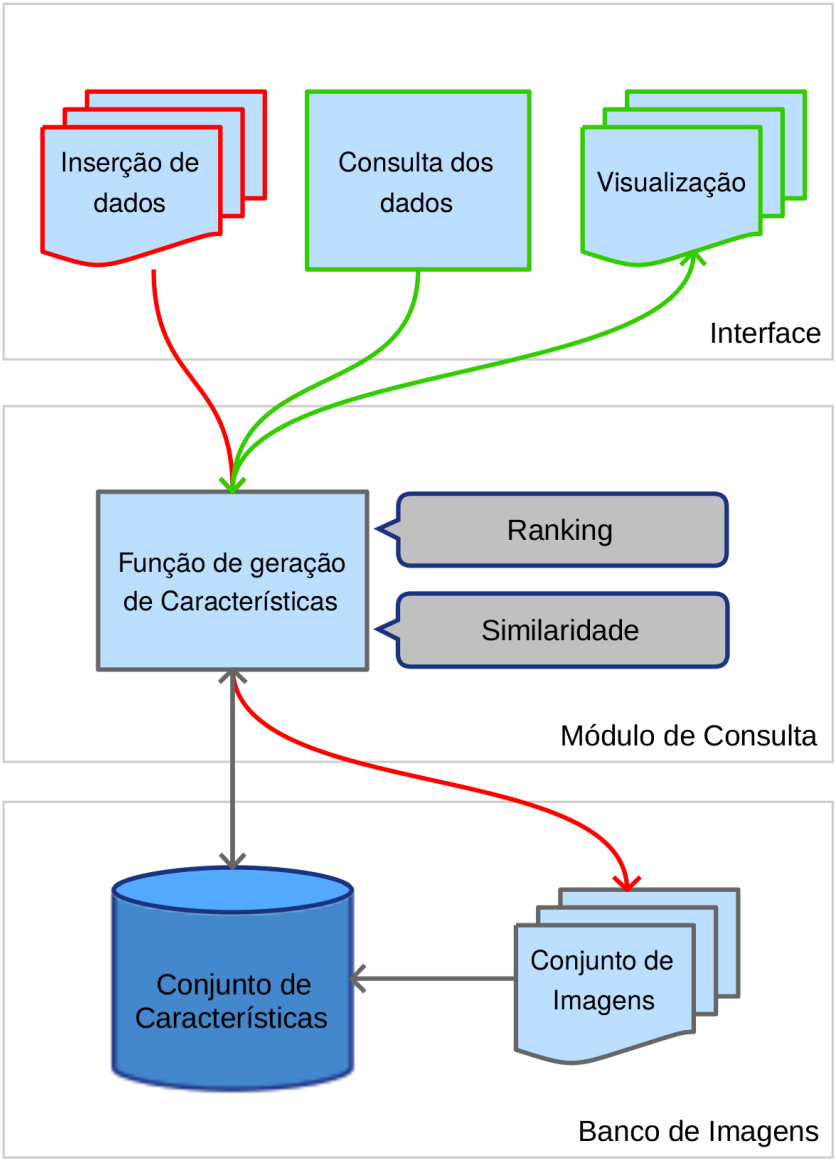
2.1 COMPONENTES DE UM SISTEMA DE CBIR

Na Figura 2, é possível observar um modelo geral dos principais grupos de sistemas que compõem uma solução de CBIR. A primeira camada é composta pela "interface" com o usuário, onde são inseridas as informações do banco, o local onde são feitas as consultas e o retorno dos dados consultados. A camada do "módulo de consulta" gera a classificação das imagens, tornando possível determinar um *ranking* e uma relação de similaridade entre os itens do banco de imagens. A camada do "banco de imagens" é responsável pelo armazenamento dos arquivos das imagens e do índice de dados gerados pela função de geração de características.

Atualmente, o banco de dados de imagens utilizado no projeto LABECA conta com uma topologia semelhante à descrita na Figura 2. Existe uma "interface" que permite ao usuário efetuar uma consulta por termos, há um "módulo de consulta" que processa o conjunto de dados, junto à camada do "banco de imagens", armazenando os respectivos arquivos de imagens.

O retorno das informações gerado pelas consultas ao banco é organizado em função das palavras pesquisadas em uma listagem pela posição da mesma na frase onde se encontra. Dessa forma, o sistema aumenta o peso da classificação em função da quantidade de acessos (cliques ou visualizações) nos termos retornados pelas buscas anteriores.

Figura 2 – Modelo de Recuperação de Imagens Baseada em Conteúdo.



Fonte: "Adaptado de" Torres; Falcão, 2006.

Este tipo de organização, é trivial em sistemas de busca textual, tanto que diversas empresas e grupos fornecem soluções para este tipo de serviço. Porém, nesse caso, mesmo existindo uma relevância dos resultados da busca, existe uma grande dependência da forma como as anotações foram adicionadas às imagens.

Assim sendo, existe um problema de descontinuidade entre o que está textualmente descrito e o que o conjunto de informações na imagem representa.

2.2 DESCONTINUIDADES SEMÂNTICAS

A bibliografia que trata de CBIR apresenta dois tipos de descontinuidades (*gaps*) (DESERNO; ANTANI; LONG, 2009): descontinuidade semântica (*semantic gap*) e descontinuidade sensorial (*sensory gap*). Por descontinuidade semântica (SMEULDERS; WORRING, 2000; MÜLLER et al., 2004; EAKINS; GRAHAM, 1999), entende-se que existe uma diferença entre a informação extraída dos dados visuais e a interpretação destes dados por uma pessoa. Por descontinuidade sensorial (SMEULDERS; WORRING, 2000), entende-se que existe uma diferença entre o objeto no mundo real e a representação computacional deste objeto.

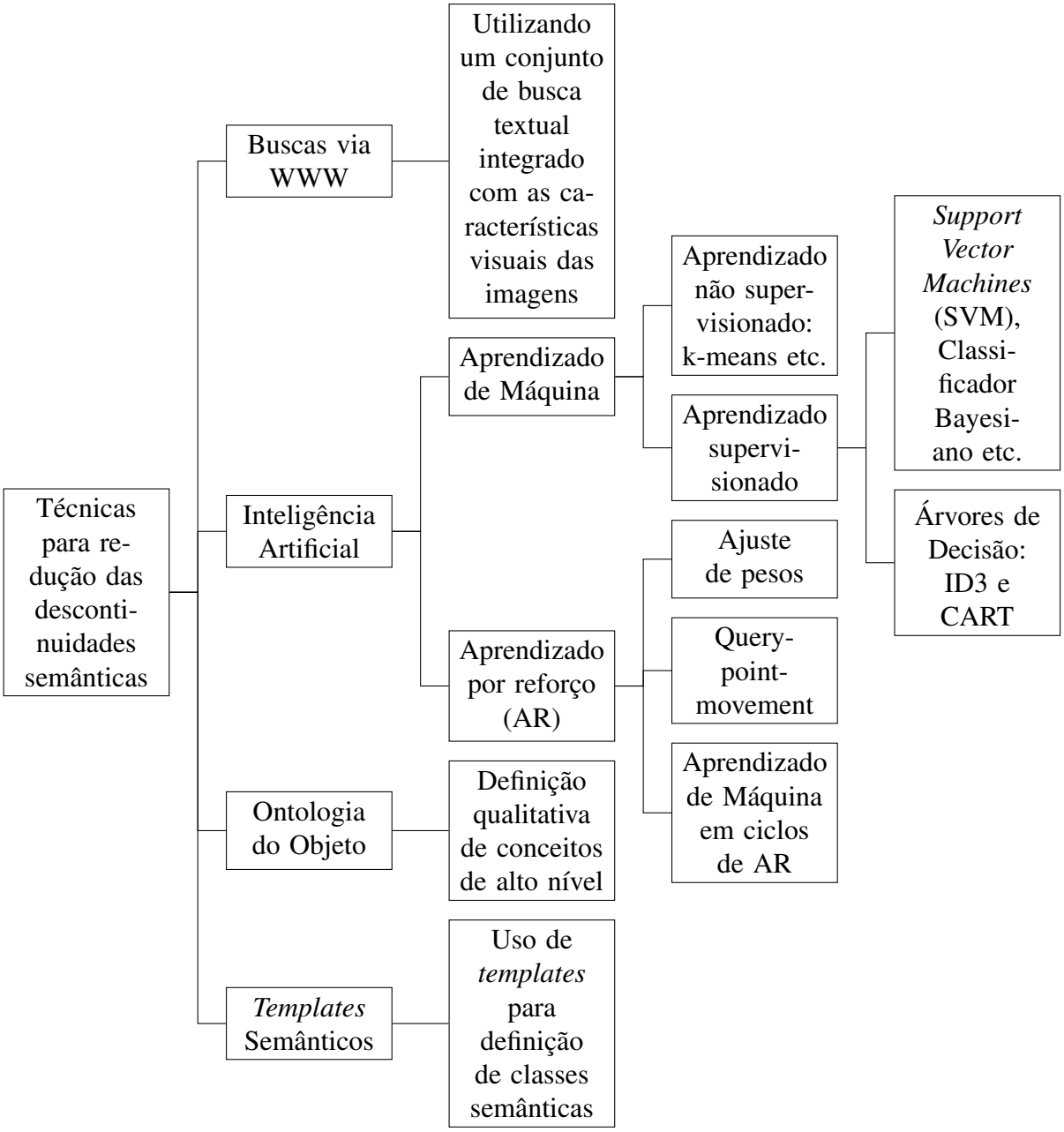
Uma das observações importantes, extraídas também desses artigos, é a classificação quanto ao tipo de pesquisa ou consulta de imagens. Sobre esse aspecto é possível determinar três níveis de abstração; o primeiro, relacionado aos aspectos *primitivos* como, por exemplo, uma cor, textura ou os limites de um objeto (SMITH; CHANG, 1997); o segundo, aos aspectos *lógicos* que definem a representação de um objeto (GUDIVADA; RAGHAVAN, 1995); e o terceiro, relacionado aos atributos *abstratos* que definem o contexto da imagem como um todo (EAKINS; GRAHAM, 1999).

O caminho mais importante se dá no equilíbrio entre a forma manual e a forma automatizada de pesquisa de imagens, e, principalmente, em determinar uma forma de capturar todos os aspectos (aspectos primitivos, aspectos lógicos e atributos abstratos). Nesse sentido, existem trabalhos abordando novas formas de compreender a interação entre uma pessoa e um sistema de CBIR para tentar expandir o conhecimento tanto do lado das imagens armazenadas (EAKINS; GRAHAM, 1999) quanto do usuário que pesquisa essas imagens (EAKINS, 2002).

Liu et al., 2007, apresenta um estudo das principais técnicas de se reduzir as descontinuidades semânticas em sistemas de CBIR. Um conjunto destas técnicas é apresentado na Figura 3. Avaliando os cinco grupos apresentados, a criação de ontologias de objetos e os *templates* semânticos tem uma grande dependência de técnicas de visão computacional, como por exemplo, quantização de cores ou avaliação de contornos de um conjunto de imagens; os outros três grupos, utilizam tecnologias, como aprendizado de máquina, que dependem de um treinamento prévio, e necessitam um aprendizado de modelos semânticos complexos.

Apesar dos avanços na tentativa de melhorar as interfaces humano computador baseadas em teclado e mouse, o processamento de um CBIR, que faz uso deste tipo de interface, não é suficiente para se obter as preferências de uma pessoa. Tem-se buscado outras formas de compreender quais são as intenções e interesses da pessoa por trás da máquina. Dessa forma, surgem novos estudos e linhas de pesquisa para outros mecanismos de redução das descontinuidades semânticas.

Figura 3 – Técnicas de redução de descontinuidades semânticas.



Fonte: "Adaptado de" Liu et al., 2007.

2.3 BIOSINAIS EM APLICAÇÕES DE CBIR

As pesquisas mais atuais apontam que uma forma de abordar o problema descrito nos parágrafos anteriores é trabalhar em um domínio específico (GUDIVADA; RAGHAVAN, 1995; EAKINS; GRAHAM, 1999), como por exemplo na área de medicina, onde existem trabalhos que tratam da busca e armazenamento de imagens (MÜLLER et al., 2004), da compreensão humana relacionada à pesquisa destas imagens (MÜLLER et al., 2004; DESERNO; ANTANI; LONG, 2009; LI et al., 2010) e de sistemas de apoio como acompanhamento da visão (*eye*

tracking) (MAIORANA, 2013) e retorno de informação (*relevance feedback*) (RUI; HUANG, 1998; COX et al., 2000) como mecanismos que permitem uma expansão dos limites impostos pela interface entre o usuário e o sistema que executa a pesquisa.

Por meio de um sensor de captura de informações relacionadas aos mecanismos da visão humana, algumas pesquisas apontam para a ampliação da interface com a máquina, sendo possível determinar o grau de atenção visual de um usuário para determinada imagem (SMEULDERS; WORRING, 2000; DUCHOWSKI, 2002; FU; CHI; FENG, 2006), composto pelo tempo em que uma região de interesse é observada (SCHULZE; FRISTER; SHAFAIT, 2013). Compreendendo o grau de atenção, é possível avaliar os processos que envolvem percepção e os conceitos compreendidos no momento em que uma imagem é visualizada por uma pessoa (LI et al., 2010).

Do ponto de vista mecânico de uma interface humano computador, algumas pesquisas demonstram que é possível utilizar a captura da visão para que sejam enviados comandos para um computador (MORIMOTO; MIMICA, 2005), como por exemplo, posição de um cursor na tela afim de substituir o mouse em função do ponto de observação.

Também na mesma linha, o uso de interfaces que visam aproveitar informações vindas do cérebro do usuário fornecem um conjunto maior de indicadores sobre o estado de atenção, resposta a determinados estímulos, estado emocional e respostas não intencionais. Este tipo de interface, chamado de ICC ou BCI permite de forma não invasiva obter informações sobre a atividade cerebral através da leitura de potencial elétrico (via EEG e magnetoencefalograma), concentração de oxigênio (via espectroscopia funcional na região do infravermelho, do inglês *functional near infrared spectroscopy*) e fluxo sanguíneo (via ressonância magnética funcional) em um determinado intervalo de tempo, e armazenamento dessas informações para um futuro processamento.

2.4 BANCOS DE IMAGENS PARA COMPARAÇÃO DE CBIR

Na maior parte das pesquisas relacionadas ao uso de imagens em CBIR, existe a necessidade de se comparar e avaliar os resultados obtidos pelas diferentes formas de se consultar um banco de imagens.

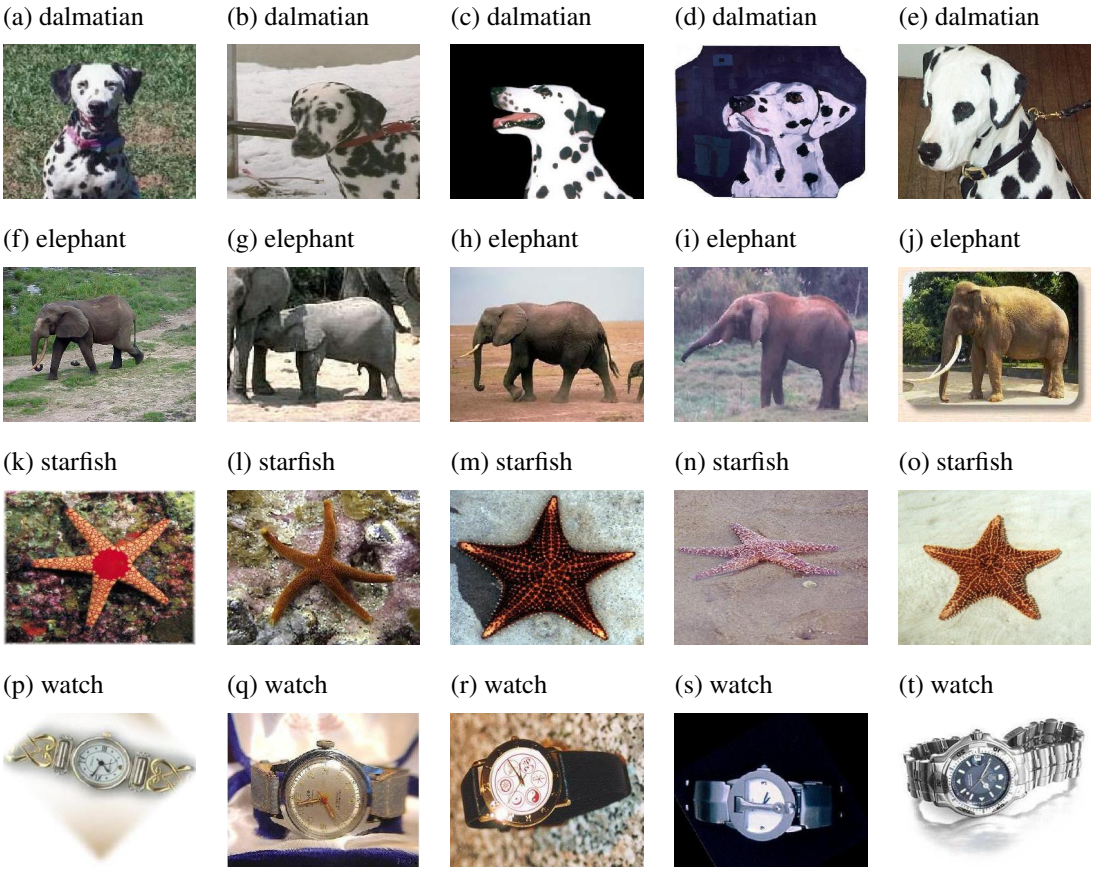
Para que fosse possível estudar os diferentes métodos de comparação, avaliação e resultados, diferentes grupos de pesquisadores passaram a fornecer abertamente seus bancos de dados de imagens, criados com métodos bem definidos de organização e pré-classificação.

Dessa forma, um pesquisador que tem a necessidade de avaliar um novo método, pode fazer uso de um banco de dados que já foi utilizado em diferentes pesquisas e comparar seus resultados com outras pesquisas já desenvolvidas. Um exemplo desse tipo de aplicação pode ser observado nas pesquisas dos autores Wang et al. (2009) e Ušćumlić, Chavarriaga e Millán (2013), que fazem uso do Caltech 101, conforme apresentado na Tabela 4 da página 51.

2.4.1 Caltech 101

Um exemplo de base de dados de pesquisa é chamada de Caltech 101[1] (FEI-FEI; FERGUS; PERONA, 2007). Uma base de dados de imagens naturais que possui aproximadamente 9144 imagens classificadas em 101 categorias, com uma variação de 40 à 800 imagens por categoria, com 50 imagens na maioria das categorias. Esta base de dados de imagens é utilizada, principalmente, para o treinamento de algorítimos de reconhecimento para aplicações de visão computacional. Alguns exemplos de imagens são apresentados na Figura 4[2].

Figura 4 – Exemplos de imagens da base de dados Caltech 101.



Fonte: Autor.

2.4.2 Nausitoo

Um segundo exemplo de banco de dados de pesquisa é o conjunto de imagens extraídas do banco de dados de imagens Nausitoo[3], desenvolvido no projeto temático LABECA (RE-

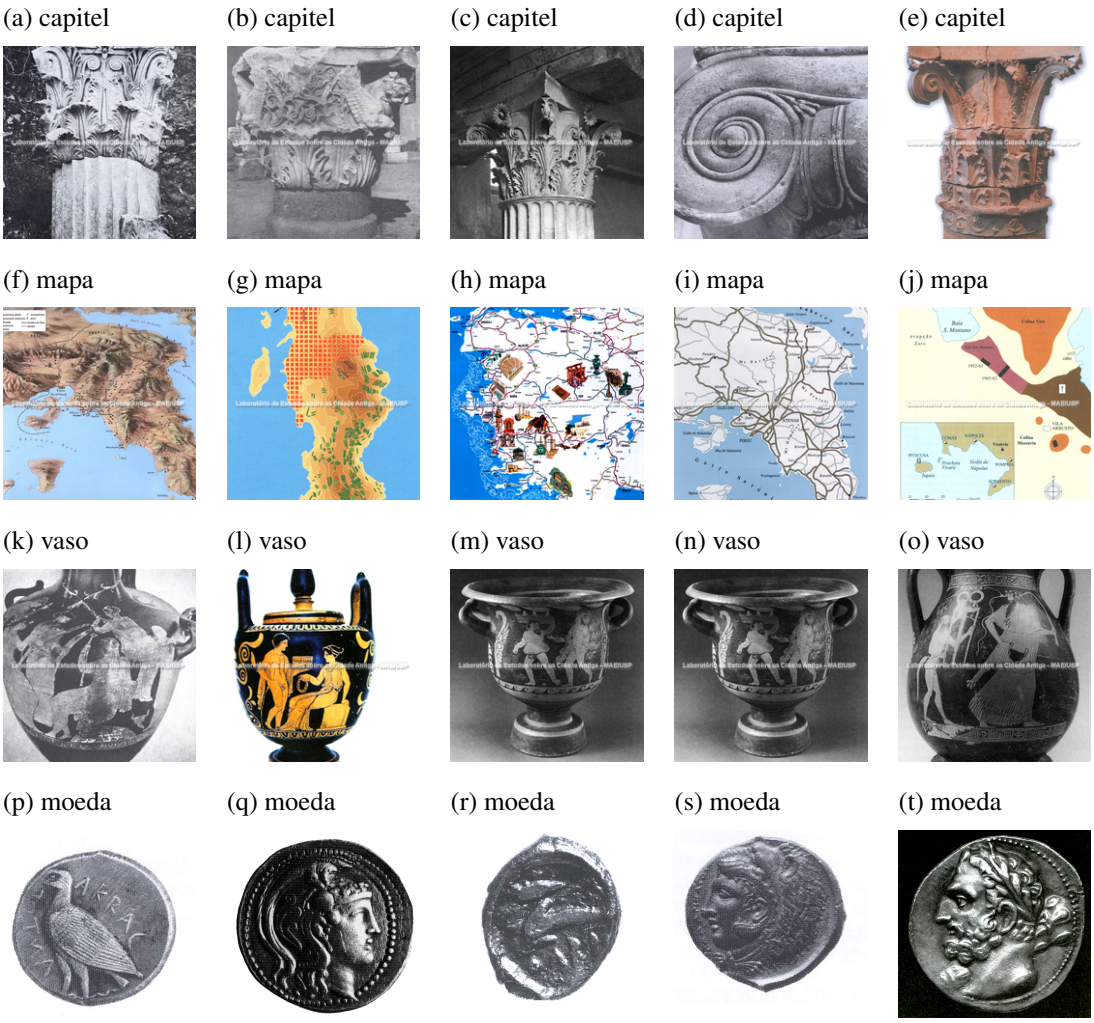
[1] `http://www.vision.caltech.edu/Image_Datasets/Caltech101/`

[2] Os nomes das categorias serão mantidos no idioma original ao longo de todo o texto com o objetivo de manter uma compatibilidade em caso de uma consulta aos arquivos originais da bases de dados de imagens.

[3] `http://labeca.mae.usp.br/city/`

ZENDE; BECHELLI, 2015). Este banco possui aproximadamente 7000 imagens relacionadas às pesquisas arqueológicas desenvolvidas pelo grupo do projeto e classificadas manualmente pelos pesquisadores do laboratório. Alguns exemplos de imagens são apresentados na Figura 5

Figura 5 – Exemplos de imagens da base de dados Nausitoo.

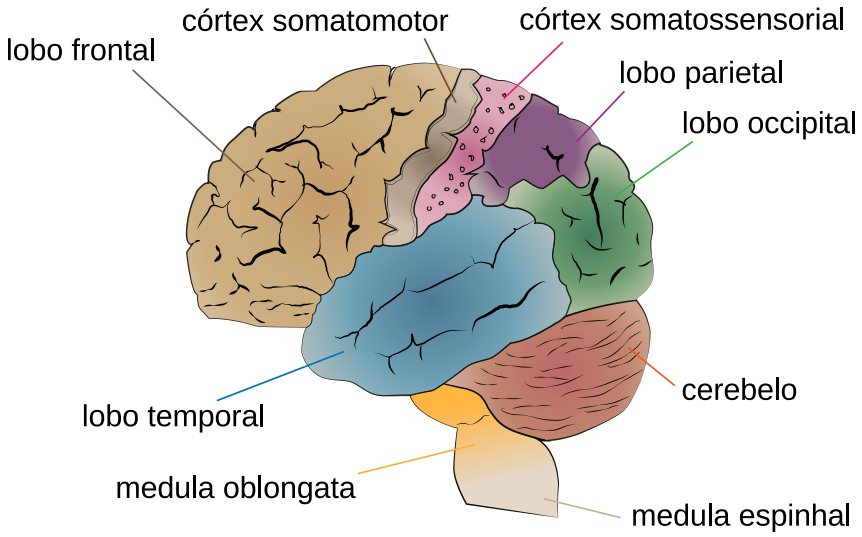


Fonte: Autor.

3 O CÉREBRO

O sistema nervoso central é um complexo sistema responsável por toda a malha que carrega a informação para as mais diversas partes do corpo humano, desde a coordenação dos sistemas como os responsáveis pela respiração até a forma que compreendemos o ambiente ao nosso redor. No ponto central desse sistema encontramos o cérebro que, através de uma visão mecanicista, representa a central computacional responsável pelas decisões e interações com a realidade (PINKER, 1998). O córtex cerebral é dividido em segmentos responsáveis pelas diferentes regiões do corpo, pelo envio de informações até o retorno sensorial dessas mesmas e pelo processamento destas informações (KHANDPUR, 2003).

Figura 6 – Perfil do cérebro humano e suas regiões.



Fonte: “Adaptado de” National Cancer Institute, 2013.

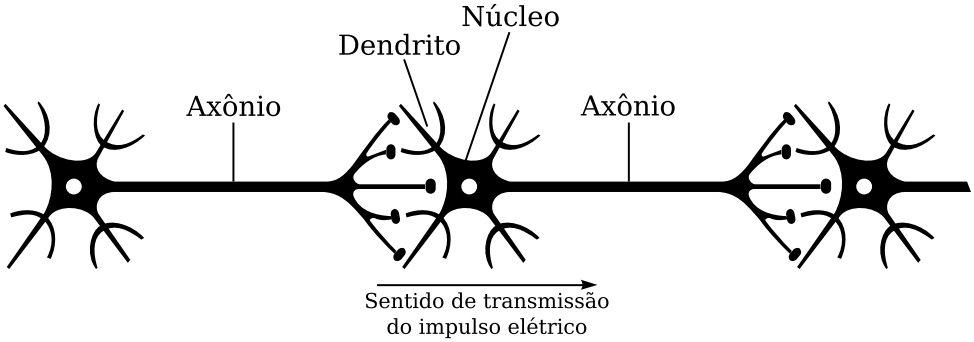
O cérebro, conforme a Figura 6, é representado pelas seguintes estruturas, que de maneira geral estão associadas as seguintes funções:

a) cerebelo: associado ao equilíbrio e coordenação dos movimentos do corpo;
b) córtex somatomotor: associado aos movimentos voluntários;
c) córtex somatossensorial: associado ao tratamento das informações sensoriais do corpo;
d) lobo frontal: associado à linguagem, planejamento (movimentos e ações) e memória;
e) lobo occipital: associado ao reconhecimento espacial e processamento dos estímulos visuais;
f) lobo parietal: associado aos estímulos sensoriais como a dor;
g) lobo temporal: associado à percepção dos estímulos auditivos e olfativos, emoções, memória e reconhecimento facial;

3.1 SINAIS CEREBRAIS

No caso do cérebro humano, esta diferença de potenciais em suas diferentes regiões é gerada através de um conjunto de processos eletroquímicos dentro das células do sistema nervoso, chamadas de neurônios, que compõe todas as estruturas cerebrais (DICKTER; KIEFFABER, 2014). O neurônio da Figura 7, é apresentado com todas as suas estruturas, onde os impulsos elétricos, gerados através de processo eletroquímicos, são gerados em seus núcleos e/ou recebidos de outros neurônios através dos dendritos, e em seguida transmitidos para um ou mais neurônios via o axônio e suas terminações que se conectam aos dendritos de outros neurônios, formando uma rede densa e complexa que controla o corpo.

Figura 7 – Modelo de neurônio para a transmissão do impulsos elétricos.



Fonte: "Adaptado de" Khandpur, 2003.

Um exemplo de processo eletroquímico que ocorre no cérebro é executado pela estrutura chamada "bomba de (Na^+/K^+)" (DRONGELEN, 2007), que controla a troca íons na membrana plasmática entre os ambientes de alta e baixa concentração iônica, controlando a diferença de concentração de íons de sódio (Na^+) e ânions de cloreto (Cl^-) fora de uma membrana em comparação com uma alta concentração de íons de potássio (K^+) dentro da membrana (DUFFY; IYER; SURWILLO, 1989). Estes sinais ocorrem de forma espontânea, como por exemplo os sinais do cérebro que comandam as funções biológicas como o batimento cardíaco, processam as informações visuais, processam os estímulos táteis entre outros tipos de processos espontâneos, e através de uma resposta a um estímulo gerando um potencial de ação, do inglês *action potential* (KHANDPUR, 2003). Esta ação, através de processos eletroquímicos de um conjunto de neurônios, gera um campo elétrico que constitui um biosinal que pode ser medido através das várias camadas entre o cérebro e o couro cabeludo.

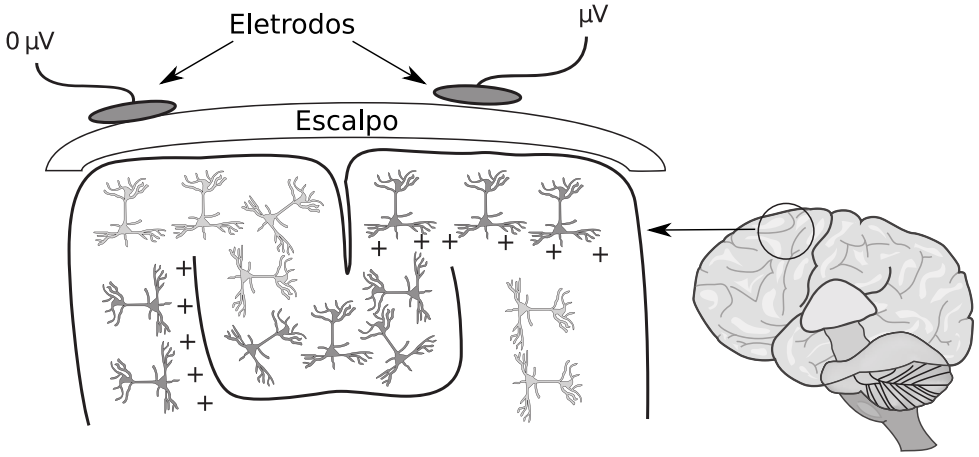
Luck (2014), descreve dois dos principais tipos de atividades elétricas que ocorrem nos neurônios: os potenciais de ação (*action potentials*) e os potenciais pós-sinápticos (*postsynaptic potentials*). O primeiro, é um pico de tensão que percorre o neurônio desde o núcleo ate a terminação do axônio, devido ao influxo de íons de sódio para dentro da membrana e fluxo de íons de potássio para fora, gerando a liberação de neurotransmissores, e o segundo é a tensão

que retorna, assim que estes neurotransmissores geram uma abertura na membrana, do neurônio pós sináptico, para a entrada de íons. Os potenciais pós-sinápticos, quando elicitados por um grade número de neurônios, podem ser medidos no topo da cabeça através de eletrodos.

3.2 ELETROENCEFALOGRAMA

Os sinais de EEG foram observados pela primeira vez em 1929 pelo psiquiatra alemão Hans Berger. O estudo demonstrou a existência de atividade elétrica no cérebro humano e que esta poderia ser medida através de amplificação dos sinais obtidos com eletrodos posicionados no couro cabeludo (LUCK, 2014). Nos anos consecutivos, a leitura de sinais de EEG passou a ser aplicada para diversas finalidades, desde a medicina até a engenharia. Os equipamentos de EEG medem a diferença de potencial elétrico em *Volts* que representa a diferença de cargas elétricas entre o ponto onde este sinal é mensurado e um segundo ponto de referência, conforme apresentado na Figura 8.

Figura 8 – Disposição de eletrodos conectados no escalpo da cabeça em relação ao cérebro.

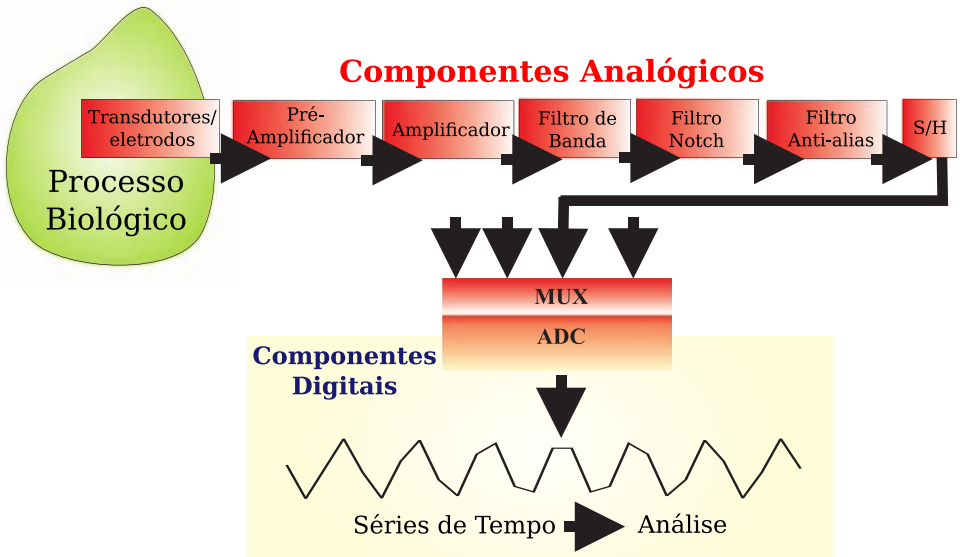


Fonte: “Adaptado de” Dickter; Kieffaber, 2014.

A obtenção deste sinal desde a sua origem eletroquímica, passando por um sinal analógico, até a sua conversão para um ambiente digital que permita o seu processamento, ocorre através de uma cadeia de circuitos e componentes analógicos e digitais conforme apresentado no modelo da Figura 9. Este modelo genérico apresenta uma cascata de aquisição de dados que parte da interface do eletrodo com o couro cabeludo até a conversão do sinal para um sinal digital. As duas primeiras etapas de pré-amplificação e amplificação são responsáveis por amplificar o sinal de um nível de $10^{-6}Volt$ ou $\mu Volt$ para um nível adequado de sinal para a conversão para digital. Este sinal é filtrado na etapa seguinte para que ocorra a remoção de artefatos que não fazem parte do biosinal, como por exemplo, interferências da rede elétrica ou do próprio circuito eletrônico e até mesmo de outros biosinais. Em seguida, o sinal é preparado

para a conversão de analógico para digital, onde um circuito faz a aquisição (*sampling*) e o armazenamento dos bits (*hold*) em paralelo com todos os eletrodos, também chamada de etapa de multiplexação, para o envio destes dados para o ambiente digital e posterior processamento online ou offline.

Figura 9 – Etapas de conversão de um sinal analógico para digital.



Fonte: "Adaptado de" Drongelen, 2007.

3.3 SISTEMA 10–20 DE EEG

A aquisição dos dados de EEG necessitava de um padrão para o posicionamento dos eletrodos ao longo do couro cabeludo dos voluntários, para uniformizar as diversas medidas obtidas entre os diversos voluntários do experimento e que atendesse os seguintes requisitos: (a) a estrutura de posicionamento deveria ser obtida através de uma medição específica, de possível reprodução e adaptável aos mais diversos tamanhos e formatos de crânio; (b) uma padrão que permitisse uma cobertura adequada do espaço ao longo de todo o topo da cabeça; (c) e provesse uma nomenclatura que simplificasse a comunicação entre o pessoal envolvido com a medição dos biosinais. Este padrão é descrito por Klem et al. (1958) e é chamado de padrão ou sistema 10–20.

Conforme apresentado na Figura 10, neste padrão os eletrodos recebem nomes relacionados com a área do cérebro em que estão posicionados ou próximos, conforme apresentado na Figura 6: Pré-frontal (FP), Frontal (F), Temporal (T), Parietal (P), Occipital (O), Central (C) e Auricular (A). Estes nomes são relacionados à números, onde os números pares representam o lado direito do cérebro, os números ímpares o lado esquerdo e a linha central do crânio recebe a letra "z" representando o número zero. Os números 10 e 20 do nome, representam a posição de 10% e 20% a partir das posições do osso do crânio: *nasion* e *inion*, a partir de onde são traçadas

as linhas de marcação para o posicionamento dos eletrodos. Cabe observar que, o ponto central no topo da cabeça (C_z) é posicionado a, exatamente, 50% das marcações da base do crânio. Os eletrodos na posição (A) são posicionados no lóbulo auricular e, normalmente, são utilizados como eletrodos de referência.

Ainda na Figura 10, os eletrodos marcados em cinza representam a proposta inicial do modelo que contemplava 22 eletrodos dispostos ao longo do couro cabeludo; porém, com o avanço dos equipamentos de medida trouxe a possibilidade de um volume muito maior de canais de medição, sendo necessária a expansão do sistema 10-20 para acomodar estes novos eletrodos, conforme apresentado na montagem de 75 eletrodos.

3.4 BANDAS DE FREQUÊNCIA

Um dos desafios ao se observar os sinais de EEG é o volume de informações obtidas nas leituras. A sobreposição de sinais é extremamente alta quando se imagina a quantidade neurônios e de funções que o cérebro humano executa de forma simultânea e sua variação ao longo do tempo, e portanto, o uso do sinal puro, extraído diretamente dos eletrodos, tem pouco valor. O que se observa na bibliografia correlata é a aplicação de métodos matemáticos e análises relacionadas ao tempo e frequência para a compreensão das funções cerebrais e o estudo da forma como esses sinais são compostos.

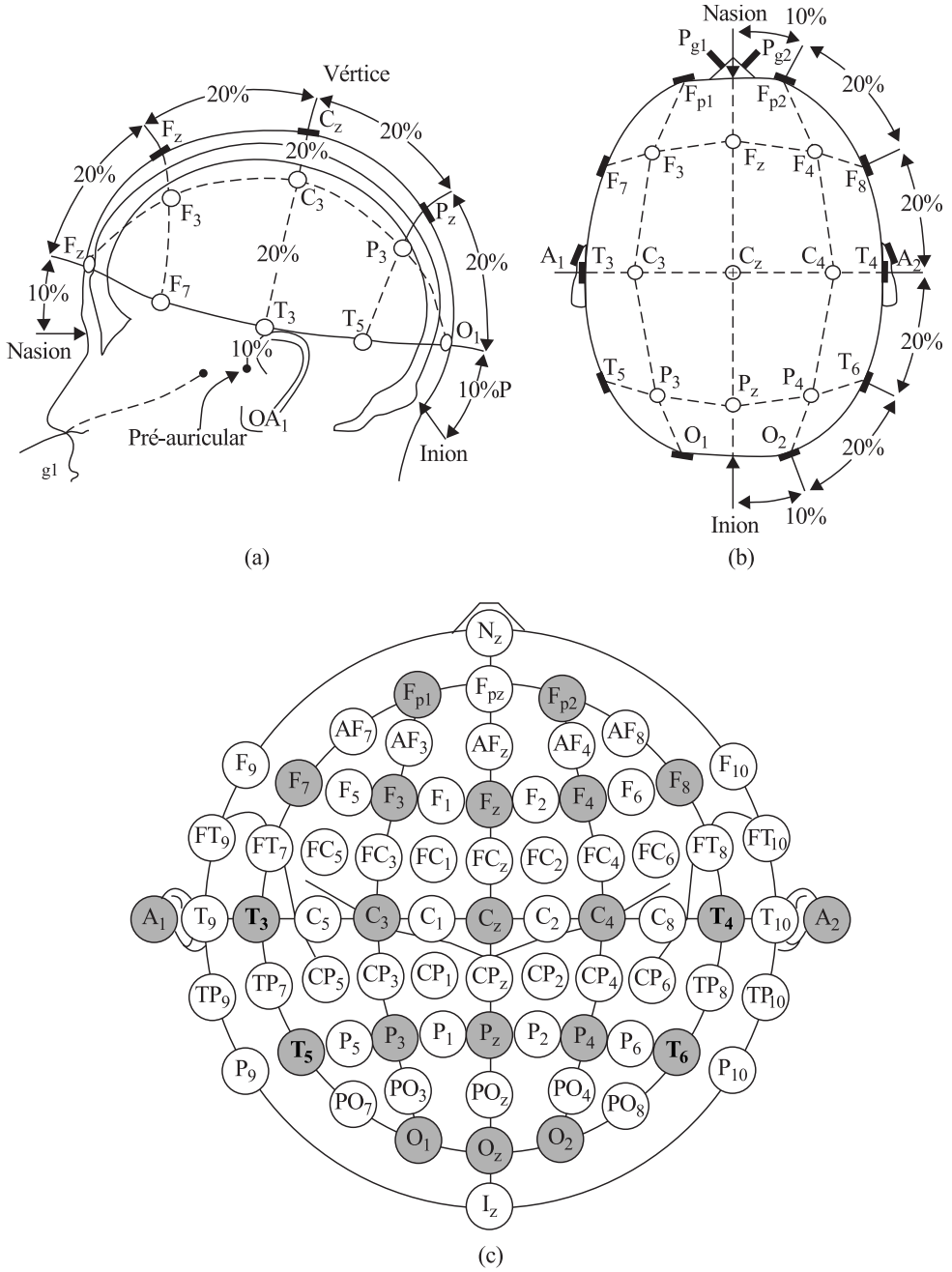
Um exemplo dessa complexidade pode ser observado na Figura 11, onde é apresentado uma medida de 10 segundos de EEG de um único eletrodo O_1 e os seus respectivos sinais, filtrados conforme o modelo de bandas de frequência: delta (1–3 Hz), teta (4–6 Hz), alfa (7–13 Hz), beta (15–25 Hz) e gama (35–45 Hz). Uma vez que o sinal é decomposto dessa forma, é possível identificar, visualmente, que diferentes ritmos cerebrais estão sendo executados de forma paralela e independente.

3.5 POTENCIAIS EVOCADOS

O cérebro humano apresenta uma variedade de sinais elétricos resultantes de fenômenos neurológicos. Destes sinais, no caso capturados por EEG, podem-se destacar algumas respostas fisiológicas que possuem características específicas de latência, amplitude e frequência captados nas medidas de tensão elétrica e que podem ser elicitados por estímulos externos (exógenos) e que não dependem da atenção do indivíduo, internos (endógenos) que dependem da capacidade de atenção do indivíduo ou motores que ocorrem durante a preparação para a execução de uma resposta motora (DONCHIN; RITTER; MCCALLUM, 1978; LUCK, 2014).

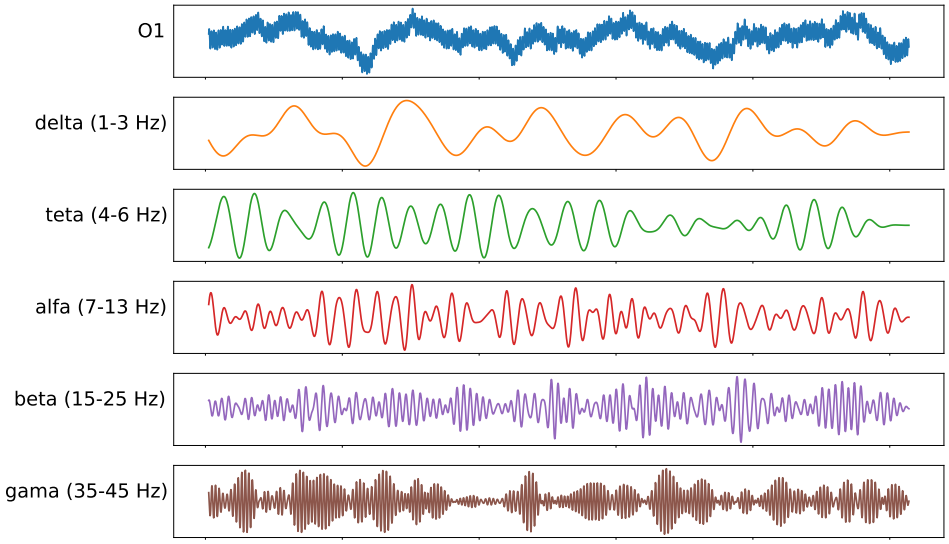
Entre os sinais que respondem à estímulos externos (exógenos), o potencial evocado visual, do inglês *Visual Evoked Potential* (VEP), é uma resposta à um estímulo visual que ocorre na região occipital do córtex cerebral. Um exemplo de VEP é o potencial evocado visual de estado estacionário (*Steady-State* VEP), onde a resposta gerada é uma modulação, em uma

Figura 10 – Representação da montagem de eletrodos 10–20 com a apresentação padrão para 75 eletrodos com as respectivas posições de eletrodos de referência: (a) e (b) representações tridimensionais das posições no couro cabeludo, e (c) configuração dos eletrodos em uma representação planificada do couro cabeludo, os eletrodos marcados em cinza apresentam uma configuração para 20 eletrodos e 2 eletrodos de referência.



Fonte: “Adaptado de” Sanei; Chambers, 2007.

Figura 11 – Decomposição de um sinal de EEG (canal O1) nas diferentes bandas clássicas de frequência: delta (1–3 Hz), teta (4–6 Hz), alfa (7–13 Hz), beta (15–25 Hz) e gama (35–45 Hz). As escalas no eixo y são diferentes com o objetivo de realçar os sinais de alta frequência e baixa amplitude.

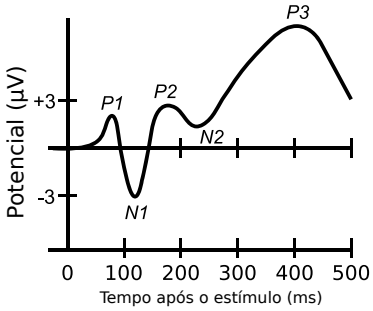


Fonte: “Adaptado de” Andrew A. Fingelkurts; Alexander A. Fingelkurts; Neves, 2010.

determinada frequência, criada por estímulos visuais, como por exemplo, uma luz piscando à uma determinada frequência específica que é replicada ou sintonizada pelo córtex cerebral e capturada via EEG (CAPATI; BECHELLI; CASTRO, 2016).

No caso dos sinais que respondem à estímulos internos (endógenos), os sinais de potencial relacionado à eventos, do inglês *Event-Related brain Potential* (ERP), apresentam alterações em momentos específicos relacionados à estímulos que criam um processo cognitivo ou de troca de contexto no processo cerebral (SUTTON et al., 1965; DONCHIN; ARBEL, 2009). A Figura 12 apresenta alguns dos principais componentes de um ERP, suas localizações no tempo após a ocorrência de um estímulo, a nomenclatura de cada um dos eventos, além da polaridade (positivos ou negativos) em que estes eventos ocorrem. Os principais sinais são: *P1*, também chamado de *P100*, tem seu pico de tensão à aproximadamente 100*ms* do momento do estímulo, mensurado nos eletrodos laterais da região occipital, é normalmente elicitado por atenção seletiva e pelo estado de excitação do voluntário; *N1*, também chamado de *N100*, é um tipo de resposta à manipulação de atenção e estímulos de atenção espacial, como estímulos visuais que exigem atenção em determinadas regiões do campo visual, como por exemplo, que ocorrem em regiões de uma tela; *P2* ocorre na região anterior e central do topo da cabeça e é elicitado pela relação de eventos alvo que são pouco frequentes. Estes experimentos, de maneira geral, apresentam uma sequência de sinais combinados de *P2-N2-P3* (LUCK, 2014).

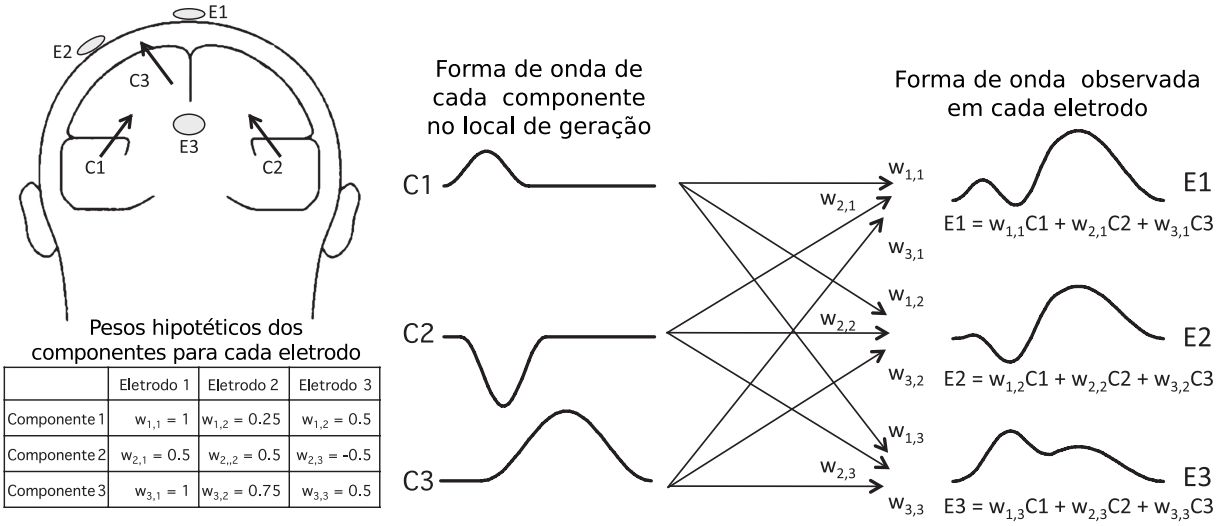
Figura 12 – Principais componentes de um ERP.



Fonte: "Adaptado de" Luck, 2014.

Luck (2014), também trata da geração individual das componentes de um ERP, chamadas de *source waveforms*, que correspondem ao local de geração da tensão para cada componente. Estas ondas são visualizadas nos eletrodos em função da soma de cada onda inicial multiplicada por um fator de peso. Uma representação visual do conceito, pode ser observada na Figura 13, onde são apresentados 3 eletrodos (E1, E2 e E3), 3 fontes de geração dos componentes (C1, C2 e C3) e um conjunto de fatores ($w_{C,E}$) que determinam o peso de cada componente em relação aos eletrodos dispostos no topo da cabeça. As setas representam os locais de geração de cada um dos componentes. A soma de cada um dos componentes, multiplicada pelo fator de peso, é observada em cada eletrodo como ondas resultantes diferentes. Nesse modelo, os pesos são determinados pela direção da onda, do local de geração ao eletrodo.

Figura 13 – Relação entre ondas geradoras de ERP.



Pesos hipotéticos dos componentes para cada eletrodo

	Eletrodo 1	Eletrodo 2	Eletrodo 3
Componente 1	$w_{1,1} = 1$	$w_{1,2} = 0.25$	$w_{1,2} = 0.5$
Componente 2	$w_{2,1} = 0.5$	$w_{2,,2} = 0.5$	$w_{2,3} = -0.5$
Componente 3	$w_{3,1} = 1$	$w_{3,2} = 0.75$	$w_{3,3} = 0.5$

Fonte: "Adaptado de" Luck, 2014.

3.6 P300 E INTERFACE CÉREBRO-COMPUTADOR

A aplicação dos potenciais elétricos gerados pelos neurônios em uma interface de comunicação computacional é o que deu origem ao termo BCI. A comunicação pode ser feita através de métodos invasivos, como por exemplo utilizando dispositivos implantados cirurgicamente (Eletrocorticograma) e não invasivos como a captura de sinais via EEG. No trabalho de Vidal (1973) encontramos uma das primeiras referências sobre o uso de sinais de EEG que poderiam ser processados de forma a criar um canal de comunicação com um computador.

Quase uma década depois, Farwell e Donchin (1988) apresentam um estudo sobre o uso de ERP, que faria uma pessoa com dificuldades motoras se comunicar, gerando sinais via um teclado virtual para um computador. Este software seria capaz de gerar uma resposta específica, uma vez que a pessoa "escolhe" uma determinada letra apenas olhando para ela no teclado virtual. O sinal de EEG gerado seria tratado de forma a determinar a existência de um ERP, e seria classificado de forma a identificar a letra escolhida usando apenas sinais cerebrais. Esta resposta representa a atividade cerebral de processamento da informação em função do evento de estímulo.

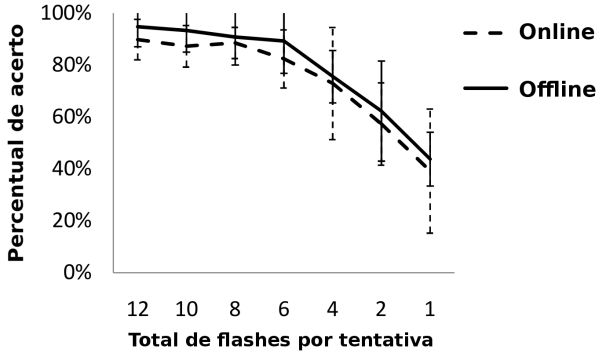
No caso de teclados virtuais operados por EEG o sinal de P300, uma deflexão positiva que ocorre 300*ms* após o estímulo, é o mais utilizado. A detecção deste tipo de sinal se dá pelo cálculo do valor médio de sinal de EEG em todos os eletrodos e entre todas as observações da letra desejada pelo voluntário. Uma vez calculados os sinais, é feita uma comparação entre os sinais gerados para a letra escolhida e as outras letras do teclado. O P300 é utilizado para especificar as imagens como alvo e não-alvo (*oddball paradigm*) (DONCHIN; RITTER; MCCALLUM, 1978; DONCHIN; ARBEL, 2009; CAPATI; BECHELLI; CASTRO, 2016).

Para a obtenção dos indicadores de potencial evocado, conforme apresentado na Figura 12, os sinais de EEG podem ser obtidos após a apresentação de cada um dos estímulos ao voluntário. Estes sinais, em resposta aos estímulos, são avaliados através do intervalo no tempo desde o início até alguns segundos após o estímulo. A curva obtida nesta "janela de tempo" é comparada com todas as janelas obtidas em todos os outros estímulos. Após a comparação, a letra que obtiver a maior quantidade de deflexões de P300, em relação ao conjunto total, é selecionada e o processo tem início novamente para a escolha da próxima letra.

Uma das principais críticas deste tipo de experimento é a necessidade de múltiplas repetições dos estímulos para que o sistema possa detectar a letra escolhida. Caso seja necessário um volume muito grande de repetições para a detecção individual de cada letra, a velocidade de escrita com um teclado deste tipo se torna inviável.

Na Figura 14, pode-se observar a relação entre o número de apresentações e o percentual de acerto para um teclado que opera online (em tempo real) e offline (simulação com dados gravados). Um ponto importante na curva apresentada é a variação percentual em relação a um único flash apresentado onde, nesse caso específico, a taxa de acerto é extremamente baixa e com uma grande variação nos resultados (DONCHIN; ARBEL, 2009).

Figura 14 – Acurácia de um teclado virtual por P300 em função do número de apresentações de estímulos visuais.



Fonte: "Adaptado de" Donchin; Arbel, 2009.

A baixa acurácia reportada por Donchin e Arbel (2009), apresentado na Figura 14, ocorre devido ao ruído inerente ao sinal EEG (FARWELL; DONCHIN, 1988). A solução mais utilizada para evitar este nível de relação sinal-ruído (*signal-to-noise ratio* (SNR)) é calcular uma média para todos os eventos alvo e outra para todos os eventos não-alvo, e dessa forma, compará- los para avaliar a existência de um sinal característico de ERP e seus marcadores, por exemplo, a existência de uma deflexão em 300*ms* a partir do início do estímulo (DONCHIN; RITTER; MCCALLUM, 1978; DONCHIN; ARBEL, 2009).

Blankertz et al. (2011) propôs uma outra abordagem para avaliar cada tentativa única de percepção dos eventos alvo e detectar um padrão de classificação individual contra eventos não-alvo. Este conceito é chamado de *single trial detection*. Com a aplicação de regressão linear sobre um filtro espacial, de forma a comparar todos os eletrodos no topo da cabeça, os sinais são processados com as janelas de eventos individuais, para que ocorra um filtro nos sinais e os eventos de ERP sejam destacados.

3.7 PLATAFORMAS DE BCI

Um cenário desenvolvido para BCI, impõe um conjunto de requisitos para que as informações adquiridas no decorrer do processo tenham a precisão adequada. Os requisitos percorrem diversas especialidades como: processamento de sinais, interação com o cenário de simulação, requisitos computacionais, neurofisiologia e desenvolvimento de software (RENARD et al., 2010).

Em função do grande número de especialidades descritas no parágrafo anterior, uma aplicação de BCI, no caso deste trabalho, voltada para captura de sinais por método não invasivo via EEG tem, necessariamente, que integrar o cenário de apresentação, a captura de EEG,

extração de características e as funções utilizadas para classificar os dados (BRUNNER et al., 2013).

Allison et al. (2012) apresentam um estudo de diversos grupos de pesquisa que desenvolveram plataformas integradas de BCI ao longo dos últimos anos. Alguns exemplos de softwares utilizados descritos em Brunner et al. (2013) e Delorme et al. (2011) são: OpenViBE, EEGLAB, BCI2000, TOBI Common Implementation Platform, BCILAB, BCI++, xBCI, BF++ e ERICA.

Com o objetivo de manter o projeto o mais independente possível de diferentes versões e tipos de sistemas operacionais, tentando utilizar, na medida do possível, softwares com código fonte livre e simplificar o desenvolvimento futuro, os softwares BCI2000, OpenViBE e EEGLAB foram avaliados.

No caso da plataforma BCI2000, o software foi facilmente instalado na plataforma Windows, porém, fica claro que o desenvolvedor se preocupa menos com o uso em outras plataformas. O EEGLAB é um pacote do software MATLAB e, nesse caso, apesar dos testes, o mesmo fica fora do requisito de software livre em função da dependência de um software comercial.

Uma das plataformas mais interessantes foi a OpenViBE, desenvolvida pelo grupo de pesquisa *French National Institute for Research in Computer Science and Control (INRIA)* na França. A plataforma OpenViBE [1] (BRUNNER et al., 2013) possui código fonte aberto com uma licença do tipo *GNU Lesser General Public License* [2] e pode ser utilizada em diferentes sistemas operacionais. Uma preocupação do grupo que desenvolve esta plataforma é a possibilidade de utilização por usuários com diferentes níveis de conhecimento em programação, possuindo uma interface com uma analogia de blocos de processamento e a possibilidade de integração com diferentes linguagens de programação como por exemplo: MATLAB, Python e Lua. O sistema modular do OpenViBE permite a integração com diferentes tipos de softwares, sendo extremamente flexível desde o suporte aos hardwares de captura de EEG até sistemas de geração de ambientes 3D para interação com os usuários. O grupo que desenvolve a plataforma também disponibiliza uma excelente documentação online com uma linguagem simples que atende usuários básicos e avançados, além de um conjunto de exemplos com diversos tipos de aplicações de BCI já desenvolvidas, o que acelera a curva de aprendizado da plataforma.

Outra plataforma pesquisada é o MNE-Python (GRAMFORT et al., 2013) que possui código fonte aberto com uma licença do tipo *BSD License*[3] e que também pode ser utilizada em diferentes sistemas operacionais. Este projeto fornece uma interface programática para o desenvolvimento de *pipelines* de processamento de dados de EEG. O pacote trabalha desde a fase de aquisição, armazenamento, organização, preprocessamento e processamento dos dados de EEG. A sua aplicação exige um nível de conhecimento de programação, o que cria uma barreira de entrada. Porém, permite alguns níveis de reprodutibilidade no desenvolvimento dos projetos, tanto na forma de execuções múltiplas do mesmo experimento, quanto na aplicação em diferentes conjuntos de dados. O que mais chama a atenção nesse pacote é a forma que os

[1] `http://openvibe.inria.fr`
[2] Versão 2 ou maior, `http://www.gnu.org/copyleft/lesser.html`
[3] 3–Clause BSD License, `https://github.com/mne-tools/mne-python#licensing`

dados permanecem organizados com relação, no caso de potenciais evocados, ao formato de janelamento dos dados e ao formato em que os sinais são relacionados ao tipo de estímulo. O pacote possui também um conjunto de filtros que podem ser aplicados diretamente aos sinais e uma interface de apresentação gráfica dos dados.

No desenvolvimento dos testes iniciais, foi utilizado a plataforma OpenViBE para integrar o equipamento Emotiv EPOC+ e criar um modelo do testes Apresentação Visual Serial Rápida, do inglês *Rapid Serial Visual Presentation* (RSVP). Esta plataforma foi selecionada em função de ser um software livre, multiplataforma, já possuía uma integração com o equipamento Emotiv EPOC+ (utiizado inicialmente na etapa de testes) e pela facilidade de desenvolvimento via interface gráfica.

Assim que o desenvolvimento inicial foi aprovado, os testes foram migrados para um processamento em Python e a aquisição dos dados via o software "Icelera iBlue52 integration" (BECHELLI, 2016b) e posteriormente processado utilizando o software MNE-Python.

4 TRANSFORMADA DE WAVELETS

O campo de pesquisa em Ondaletas, do inglês wavelets, é conhecida e amplamente discutida desde a sua reintrodução em 1985 (MEYER, 1993). Sua aplicação ocorre nas mais diversas áreas, sendo possível citar a compressão de imagens (DAUBECHIES, 1999), detecção de falhas (BELARDI et al., 2006), aplicações em biosinais (DRONGELEN, 2007), análise de dados de sismologia (VAN FLEET, 2011) e aplicações de matemáticas e de engenharia (MORETIN, 1999; BELARDI; CARDOSO; SARTORI, 2004a; 2004b).

Uma das formas de se compreender como as wavelets são aplicadas é utilizando como referência à Transformada Discreta de *Fourier*, do inglês *Discrete Fourier Transform* (DFT), onde o sinal, semelhante ao sinal de EEG ao longo do tempo apresentado na Figura 11, é decomposto em um conjunto de ondas senoidais que somadas representam o sinal original. No caso da Transformada Discreta de Wavelet, do inglês *Discrete Wavelet Transform* (DWT), a onda que é utilizada para decompor o sinal não precisa, necessariamente, ser uma onda senoidal e a forma como a decomposição é aplicada ao sinal faz com que as variações ao longo do tempo (sinais não–estacionários) possam ser visualizadas com uma granularidade maior quando comparada à DFT e também em relação às altas e baixas frequências que são representadas com níveis diferentes de detalhes (GRAPS, 1995).

O impacto dessa granularidade da DWT em relação à DFT pode ser observado na Figura 15, onde em (a) é apresentado uma curva da função estacionária $f = \sin(2\pi f_1 \cdot t) + \sin(2\pi f_2 \cdot t)$ para $t \in [0,5]$ e em (b) a curva da função não-estacionária $g_{[0,5]} = \sin(2\pi f_1 \cdot t)$ e $g_{[5,10]} = \sin(2\pi f_2 \cdot t)$. O sinal estacionário é constante no tempo e o sinal não-estacionário possui uma variação ao longo do tempo. Quando descritos no domínio das frequências de *Fourier*, apresentam o mesmo perfil de curva, e uma vez que os dois sinais são descritos no domínios das wavelets, é possível observar a alteração nos dois perfis dos níveis de coeficientes.

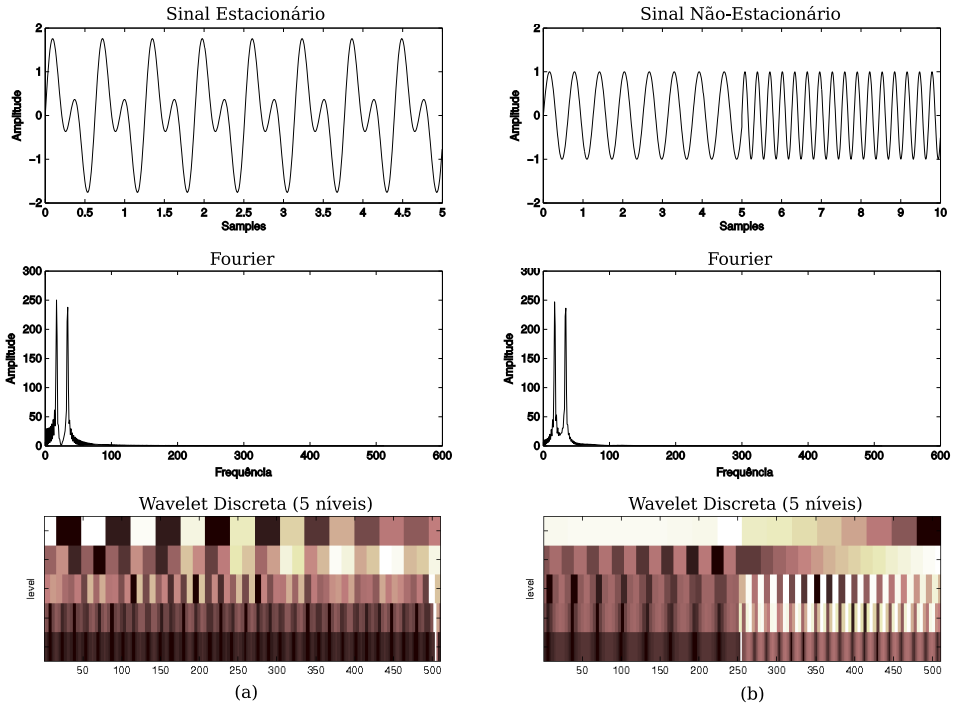
A forma mais simplificada de compreensão da aplicação das wavelets pode ser apresentada através de valores de uma série temporal (WALKER, 2008; DRONGELEN, 2007). Utilizando uma série de 8 valores consistindo de $56, 40, 8, 24, 48, 48, 40, 16$, conforme apresentado por Jensen e Cour-Harbo (2001) na Tabela 1.

Tabela 1 – Valores de uma série temporal com os cálculos de média e diferença (em negrito).

56	40	8	24	48	48	40	16
48	16	48	28	**8**	**-8**	**0**	**12**
32	38	**16**	**10**	**8**	**-8**	**0**	**12**
35	**-3**	**16**	**10**	**8**	**-8**	**0**	**12**

Fonte: “Adaptado de” Jensen; Cour-Harbo, 2001.

Figura 15 – Comparação da Transformada de Wavelets em relação Transformadas de *Fourier*. (a) sinal estacionário e (b) sinal não-estacionário.



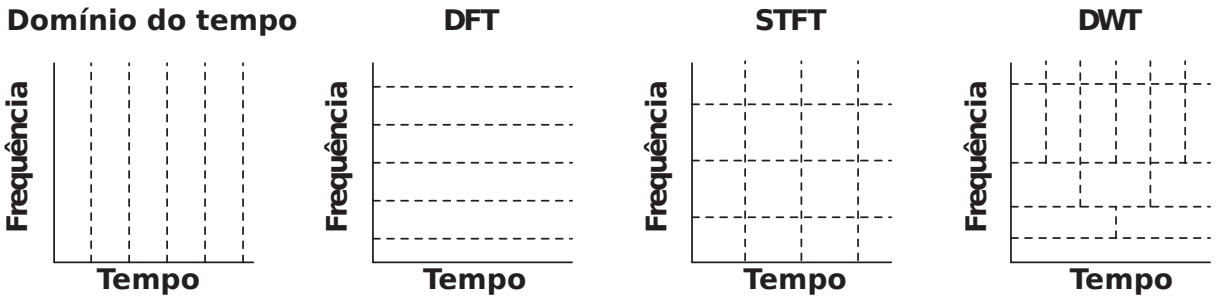
Fonte: "Adaptado de" Moraud, 2009.

Na primeira linha da Tabela 1 é apresentando o sinal, na segunda linha é apresentado nos quatro primeiros valores a média dos pares de valores de entrada, como por exemplo, $48 = (56+40)/2$, $16 = (8+24)/2$, $48 = (48+48)/2$ e $28 = (40+16)/2$. E na segunda linha, em negrito, os valores da diferença entre os pares $8 = (56-40)/2$, $-8 = (8-24)/2$, $0 = (48-48)/2$ e $12 = (40-16)/2$. A partir dessa etapa, na terceira linha, os valores das diferenças (negrito) são copiadas para a linha de baixo e o processo ocorre novamente com os valores das médias, sendo: $32 = (48+16)/2$, $38 = (48+28)/2$ e as diferenças $16 = (48-16)/2$, $10 = (48-28)/2$. Na quarta linha o processo de copiar os valores das diferenças (negrito) também é repetido e os valores médios também são calculados em $35 = (32+38)/2$, finalizando a decomposição em cada um dos níveis. O formato deste processo permite que o sinal seja revertido para o sinal original sem nenhuma perda de informação.

Com o exemplo descrito na Tabela 1 e na Figura 15 é possível identificar uma característica importante das wavelets sobre a capacidade de resolução entre as altas e baixas frequências ao longo do tempo, como apresentada em detalhe na Figura 16. As altas frequências possuem uma resolução maior ao longo do tempo através da subdivisão cada vez maior dos coeficientes gerados no processo de decomposição. Essa subdivisão faz com que as wavelets tratem da relação entre tempo e frequência (incerteza de *Heisemberg*) de uma forma natural; quanto maior a frequência, menor o tempo selecionado para a decomposição do sinal (ADDISON, 2017).

Cabe ressaltar, que uma alternativa para a aplicação de DFT, que pudesse ter multi-resolução para os diferentes níveis de frequências, seria a aplicação de *Short-Time Fourier Transform* (STFT). Onde, de forma análoga, as wavelets, diferentes janelas são formadas para obter-se a separação das frequências em função de janelas fixas de tempo, como representado na Figura 16 no domínio do tempo, DFT, STFT e DWT. Porém no caso da teoria de STFT, estas janelas de tempo tem a mesma resolução para todas as faixas de frequência ao longo do tempo, onde a teoria das wavelets se adapta de maneira a obter diferentes resoluções para as altas e baixas frequências (DICKTER; KIEFFABER, 2014).

Figura 16 – Comparação entre as resoluções de tempo e frequência.



Fonte: "Adaptado de" Dickter; Kieffaber, 2014.

A Transformada de wavelet de um sinal é definida como sendo a soma de todo o tempo do sinal $f(t)$ multiplicada por uma versão escalada e deslocada da função wavelet $\psi_{a,b}(x)$, descritos conforme a Equação 1:

$$\psi_{a,b}(x) = |a|^{-1/2}\psi\left(\frac{x-b}{a}\right), \quad a,b \in \mathbb{R}, a \neq 0. \tag{1}$$

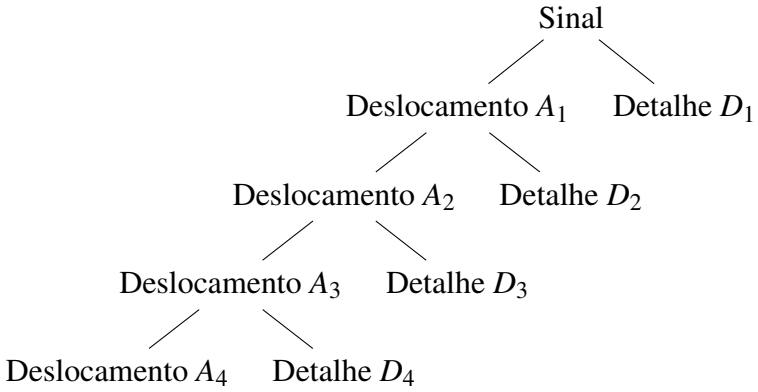
A função ψ é chamada de wavelet mãe e os valores para a e b são $a = 2^{-j}$, $b = k2^{-j}$ para $j,k \in \mathbb{Z}$. Onde, a é responsável pelo escalamento (*dilation* ou *scaling*) e b é a variável responsável pela translação (*translation*). O parâmetro de escalamento aumenta ou diminui a largura e tamanho da wavelet e o parâmetro de translação é responsável por percorrer o sinal ao longo do tempo (MORETIN, 1999). Dessa forma, um parâmetro pequeno de escalamento corresponde a um sinal reduzido da função da wavelet, que de maneira análoga, detecta no sinal as componentes de mudança rápida ou sinais de alta frequência. Por outro lado, as componentes de baixa frequência são destacadas do sinal com parâmetros de escala grande com uma função wavelet "esticada".

A Equação 1, define de forma geral o modo de operação das wavelets. Diferente da teoria da transformada de *Fourier*, que utiliza ondas senoidais como base, a teoria das wavelets permite o uso, e eventualmente novos desenvolvimentos, de uma variedade de wavelets mãe conforme a necessidade da aplicação (VALENS, 1999). As wavelets mãe, definem a função sobre o qual o sinal será decomposto, utilizando a analogia com a teoria de *Fourier*, no lugar das ondas senoidais.

$$\psi(x) = \begin{cases} +1 & 0 \leq x < 1/2, \\ -1 & 1/2 \leq x < 1, \\ 0 & \text{restante.} \end{cases} \tag{2}$$

A estrutura de formação das wavelets permite que tanto os sinais de alta frequência como os de baixa frequência possam ser aplicados de forma iterativa ao sinal. Dessa maneira, a decomposição pode ser aplicada em diversos níveis, aumentando a quantidade de detalhes do sinal e realçando as altas frequências. Na Tabela 1, o sinal é decomposto em 4 níveis conforme descrito na Figura 17.

Figura 17 – Decomposição de 4 níveis de Wavelet.



Fonte: Autor.

5 APRENDIZADO DE MÁQUINA

Todo o volume de dados utilizados e armazenados nesse tipo de sistema impõe uma necessidade de aplicar métodos que possam determinar, da melhor maneira possível, as respostas necessárias utilizando o maior volume possível de informações. Neste tipo de cenário, a aplicação de inteligência artificial (IA), como por exemplo, aprendizado de máquina, do inglês *machine learning* (ML) é uma alternativa que combina estatística e processamento computacional (RUSSELL; NORVIG, 2009; HASTIE; TIBSHIRANI; FRIEDMAN, 2009).

Existem dois formatos principais de ML, método supervisionado e não supervisionado. O método supervisionado, é comumente aplicado em conjuntos de dados categorizados, estes dados são utilizados para o treinamento de modelos matemáticos, que posteriormente, irão definir em que tipo de categoria um conjunto de dados de entrada se encaixa. O método não supervisionado, por sua vez, é aplicado em conjuntos de dados que não estão categorizados, sendo o objetivo principal, encontrar possíveis formas de agrupar o conjunto de dados, e compreender se existe algum modelo de distribuição do mesmo.

Do ponto de vista do tratamento das informações armazenadas em um banco de dados de imagens, algumas pesquisas mais recentes apresentam inovações relacionadas ao treinamento de um sistema de aprendizado de máquina utilizando dados sem classificação alguma para determinar aspectos de uma classe específica de imagens (LE et al., 2011). Um exemplo deste tipo de aplicação pode ser observado no trabalho de Crowley e Zisserman (2013), onde um conjunto de imagens obtidas através do repositório de vasos gregos clássicos, Beazley Archive [1], com o objetivo de buscar diferentes tipos de imagens nas pinturas da superfície dos vasos de cerâmica grega. Utilizando um conjunto de estágios de processamento que vai desde o ajuste da fotografia da cerâmica até a segregação das diferentes figuras que compõem a imagem para que as mesmas sejam inseridas no classificador que seleciona os vasos que possuem imagens correspondentes. Em um trabalho posterior, Crowley e Zisserman (2014), utilizaram um conjunto de imagens naturais e manualmente classificadas para pesquisar de forma distinta em um segundo conjunto de pinturas, imagens que fazem parte do mesmo agrupamento de categoria, como por exemplo, imagens que contenham figuras de “trem” ou “pássaro”, fazendo uso de recursos de visão computacional e aprendizado de máquina.

A aplicação de aprendizado de máquina também vem sendo utilizada para classificar e agrupar as informações, provenientes da interação entre um usuário e os sistemas de CBIR, para determinar e comparar as regiões de interesse e a forma como retornar com precisão as imagens pesquisadas (CHEN et al., 2006b; SCHULZE; FRISTER; SHAFAIT, 2013).

A partir do resultado de treinamento e teste de um classificador, são extraídos parâmetros para a comparação e avaliação da qualidade do mesmo. Estes fatores são calculados com o resultado obtidos de:

[1] `http://www.beazley.ox.ac.uk/pottery`

a) Verdadeiros positivos ou *True Positives* (TP): testes positivos que são classificados corretamente pelo classificador;
b) Falsos positivos ou *False Positives* (FP): testes negativos que são classificados erroneamente como positivos pelo classificador;
c) Verdadeiros negativos ou *True Negatives* (TN): testes negativos que são classificados corretamente pelo classificador;
d) Falsos negativos ou *False Negatives* (FN): testes positivos que são classificados erroneamente como negativos pelo classificador.

A partir dos resultados, são calculados os seguintes fatores de Precisão (PR), *Recall* (RC), Acurácia (ACC) e resultados de $F1$ (MURPHY, 2012), conforme apresentados respectivamente nas Equações 3, 4, 5 e 6.

$$PR = \frac{TP}{TP + FP} \tag{3}$$

$$RC = \frac{TP}{TP + FN} \tag{4}$$

$$ACC = \frac{TP + TN}{TP + TN + FP + FN} \tag{5}$$

$$F1 = \frac{PR \times RC}{PR + RC} \tag{6}$$

5.1 CLASSIFICADOR DE RIDGE

A aplicação de modelos para a modificação de coeficientes gerados após a transposição de um sinal para o domínio das wavelets, é amplamente utilizado. Um exemplo deste tipo de aplicação, pode ser observado no trabalho de Daubechies (1999) com modelos de compressão de imagens.

Hastie, Tibshirani e Friedman (2009) apresentam uma abordagem para suavização de wavelets baseada em regressão linear, para criar um limite para minimizar coeficientes pequenos no domínio wavelet, de forma a recompor um sinal "suavizado". Eles também descrevem um método de uso de regressão (*lasso*) que faz a redução dos coeficientes do sinal com um pacote de software chamado SURE (*Stein Unbiased Risk Estimation*) (DONOHO; JOHNSTONE, 1994), que define coeficientes para patamares mínimos, ou até mesmo para zero, no domínio das wavelets, e em seguida é revertido novamente para o domínio original do sinal com uma transformada inversa de wavelet.

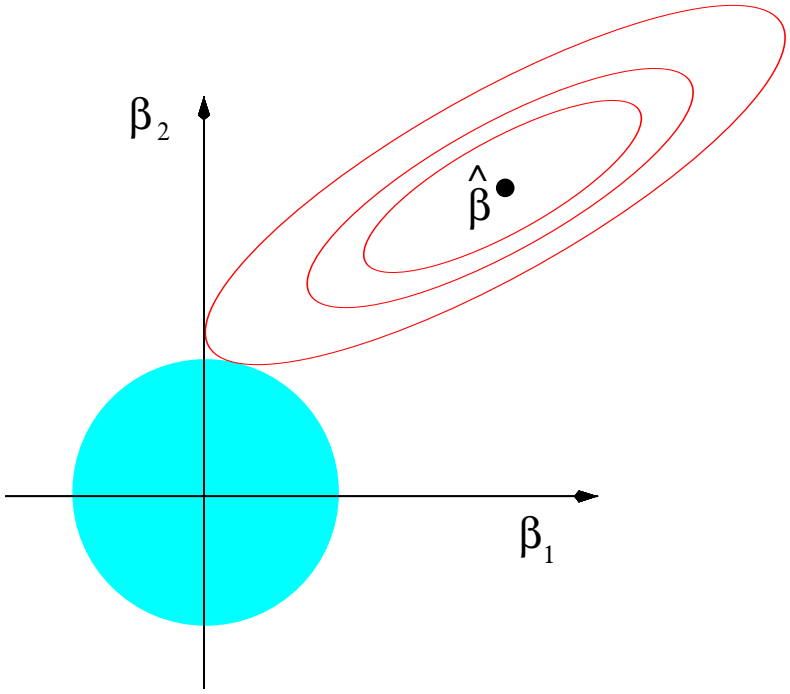
A Regressão de *Ridge*, descrita pela Equação 7, é uma técnica que trabalha a redução dos coeficientes β de uma equação linear, tendo como base a multiplicação destes coeficientes por um fator λ com o objetivo de minimizar o erro associado a mesma. Com esse processo de modificação das variáveis, a Regressão de *Ridge* reduz a multicolinearidade, que é uma situação

onde um forte relacionamento entre as variáveis independentes torna difícil a estimativa de efeitos destas mesmas variáveis sobre uma variável dependente, ou de forma mais precisa, a variância entre alguns coeficientes se tornam grandes.

$$\widehat{\beta}_{ridge} = (X'X + \lambda I)^{-1} X'y \tag{7}$$

A influência do $\widehat{\beta}_{ridge}$, pode ser interpretado através da Figura 18. A circunferência azul representa o termo de regularização l_2. O ponto $\widehat{\beta}$ representa o erro quadrático mínimo. O erro aumenta exponencialmente a medida que se aproxima do termo de regularização. Uma vez que os termos se cruzam, ambos serão minimizados, determinando o fator $\widehat{\beta}_{ridge}$.

Figura 18 – Estimativa dos coeficientes da regressão de *ridge*.



Fonte: “Adaptado de” Trevor Hastie; Robert Tibshirani; Wainwright, 2015.

Regressão de *Ridge* é conhecida pela resiliência ao sobre-ajuste (*overfitting*) de aprendizado de máquina, resultando em um único ponto de conversão (JAMES et al., 2013) e também tendo um bom desempenho contra vetores de entrada que tem tamanho ainda maior que o tamanho das amostras de entrada (vetores de alta dimensionalidade). Este grande vetor é processado por regressão de *ridge* para identificar uma quantidade de variáveis que contribuem para o sinal e variáveis que podem ser reduzidas (*ridge*) ou mesmo descartadas (*lasso*) (ZHAO; OGDEN; REISS, 2012).

No caso do classificador de *ridge* (PEDREGOSA et al., 2011), conforme a Equação 7, o fator α é configurado como parâmetro de entrada do treinamento do classificador para que seja controlado o resultado final da predição.

Sendo $w = (w_1, \ldots, w_p)$, um conjunto de coeficientes que multiplicam o vetor de entrada do classificador, w_0 o coeficiente linear que intercepta a função em y para $x = 0$, X representa o vetor de entrada das medidas que precisam ser classificadas, y representa a categoria (*label*) em que o vetor X é classificado e $\hat{y}(w,x)$ representa o resultado esperado após o treinamento do classificador, temos:

$$\hat{y}(w,x) = w_0 + w_1 x_1 + \cdots + w_p x_p \tag{8}$$

Para uma regressão através do método dos mínimos quadrados, temos:

$$\min_{w} ||Xw - y||_2{}^2 \tag{9}$$

E a regressão de *ridge* é determinada por:

$$\min_{w} ||Xw - y||_2{}^2 + \alpha ||w||_2{}^2 \tag{10}$$

A Equação 10, pode ser interpretada com uma regressão linear pelo método dos mínimos quadrados, apresentado na Equação 9, com um fator adicional $\alpha ||w||_2{}^2$ que impõe uma penalização através do conjunto de coeficientes.

Para controlar os resultados, vários níveis de α em um método de classificação de aprendizado de máquina de pesquisa de grade (*gridsearch*) podem ser aplicados para determinar o resultado máximo de classificação.

6 TRABALHOS CORRELATOS

Neste capítulo serão avaliados os principais artigos que formam a base da pesquisa deste trabalho, comparando um conjunto específico de textos que tratam da aplicação de EEG via RSVP em CBIR.

6.1 CLASSIFICAÇÃO DE CBIR VIA EEG

Este capítulo tem como objetivo mapear as metodologias utilizadas nos artigos Wang et al. (2009), Khosla et al. (2011), Ušćumlić, Chavarriaga e Millán (2011), Healy e Smeaton (2011), Ušćumlić, Chavarriaga e Millán (2013), Mohedano et al. (2014), Mohedano et al. (2015) e Healy, Gurrin e Smeaton (2016), avaliando a forma como os dados foram adquiridos em cada experimento apresentado pelos autores, quais as bases de dados de imagens utilizadas, as características e informações sobre os voluntários, que métodos foram utilizados para preprocessamento, processamento e métodos de classificação dos dados capturados.

A comparação entre esses artigos foi utilizada para a criação do protocolo de testes e definição dos parâmetros principais.

Com o objetivo de capturar os sinais cerebrais, todos os artigos pesquisados utilizaram o sistema internacional de posicionamento 10–20 (GERSON; PARRA; SAJDA, 2006) para distribuir os eletrodos sobre o couro cabeludo, na região do córtex cerebral.

O conjunto de artigos faz uso do método de Potencial Relacionado a um Evento (ERP) (SUTTON et al., 1965) também conhecido como Potencial Evocado, que corresponde a um sinal de resposta cerebral a um estímulo externo, no caso o visual.

6.2 HARDWARE DE AQUISIÇÃO DE EEG

Na maior parte dos artigos pesquisados as informações sobre o hardware de aquisição e o ambiente de captura dos sinais de EEG (ex.: Gaiola de Faraday) não são descritos e para aqueles que expõem estes detalhes a Tabela 2 é apresentada.

Cabe aqui observar que, a maior parte dos autores, além de ter utilizado a montagem 10–20, também utilizou equipamentos com uma quantidade considerável de eletrodos (em média mais do que 30 canais) e taxas de aquisição de dados da ordem de $1KHz$.

6.3 MÉTODO DE APRESENTAÇÃO DAS IMAGENS

O método utilizado em todos os trabalhos pesquisados é o método de apresentação rápida conhecido do inglês como RSVP.

Este método consiste em apresentar diferentes imagens, com uma determinada frequência, em um ambiente controlado para um voluntário. Desta forma, diferentes imagens são

Tabela 2 – Hardware utilizado.

Referências	**Hardware**
Wang et al. (2009)	64 canais *BioSemi ActiveTwo System* com uma montagem estendida do padrão 10–20 e uma taxa de aquisição de $2048Hz$
Khosla et al. (2011)	sem descrição
Ušćumlić, Chavarriaga e Millán (2011)	sem descrição
Healy e Smeaton (2011)	16 canais, *KT88-1016 EEG System* com a referência no mastoide esquerdo e utilizando o queixo como ponto de aterramento. Foram utilizados eletrodos de $Ag/AgCl$ com uma montagem no padrão 10–20. Os dados foram digitalizados com $100Hz$ e filtrados com uma banda de passagem de $0.1Hz$ até $20Hz$
Ušćumlić, Chavarriaga e Millán (2013)	64 canais, *BioSemi ActiveTwo System* com uma montagem estendida do padrão 10–20 e uma taxa de aquisição de $2048Hz$, calibrado via atividade de eletro-oculograma (EOG)
Mohedano et al. (2014)	31 canais, de BCI com uma taxa de aquisição de $1kHz$. Com os eletrodos em uma montagem no padrão 10–20 em uma sala isolada (Gaiola de *Faraday*)
Mohedano et al. (2015)	Amplificador actiCHamp de 32 canais com taxa de aquisição de $1kHz$ com os eletrodos em uma montagem no padrão 10–20 referenciados pela média de todos os canais. Os sinais foram filtrados com um filtro passa banda de $0.1Hz$ até $20Hz$
Healy, Gurrin e Smeaton (2016)	Amplificador actiCHamp de 32 canais com taxa de aquisição de $1kHz$ com os eletrodos em uma montagem no padrão 10–20 referenciados pela média de todos os canais. Os sinais foram filtrados com um filtro passa banda de $0.2Hz$ até $20Hz$. A impedância foi mantida abaixo de $5kOhm$ em canais. Os canais ruidosos foram removidos e interpolados. Não foi utilizada nenhuma estratégia de rejeição de teste. A Análise de Componentes Independentes (ICA) foi utilizada para remover a atividade relacionada aos movimentos oculares.

Fonte: Autor.

apresentadas para que o voluntário identifique as imagens alvo dentre uma série de imagens que não são alvo (MOHEDANO et al., 2014).

Através dos diferentes artigos pesquisados, os mesmos métodos são aplicados com pequenas variações conforme descrito na Tabela 3.

Em Mohedano et al. (2014), existe uma diferença na montagem do teste e no objetivo de detecção em relação aos outros artigos citados. Nesse experimento, uma única imagem é segmentada em pequenos recortes e este conjunto é apresentado para a detecção, misturado

Tabela 3 – Detalhes dos métodos de RSVP utilizados.

Referências	**Métodos de RSVP**
Wang et al. (2009)	$240px$ x $240px$ de dimensão das imagens apresentadas em $6Hz$, com um banco de dados de 62 objetos selecionados randomicamente particionados em dois conjuntos. Num total de 1000 imagens, apresentadas em 10 blocos de 100 imagens para cada
Khosla et al. (2011)	$256px$ x $256px$ de dimensão das imagens apresentadas em $10Hz$. As imagens são mostrados em 50 imagens por tela com 10 imagens por linha
Ušćumlić, Chavarriaga e Millán (2011)	220 imagens no total, apresentadas em $4Hz$, em sequências de 20 "imagens alvo"
Healy e Smeaton (2011)	"Imagens alvo" foram inseridas em blocos de 100 "imagens não alvo" apresentadas a $10Hz$. Total de 4.800 imagens com 60 "imagens alvo" distribuídas randomicamente
Ušćumlić, Chavarriaga e Millán (2013)	20 imagens por objeto (ex. "dog", "eagle" etc.) em 10 categorias. 10% de "imagens alvo". Os voluntários são posicionados a $60cm$ da tela com as imagens ocupando aproximadamente 6° x 4° do campo visual e todos foram orientados a contar silenciosamente as imagens de um objeto especificado
Mohedano et al. (2014)	22 imagens selecionadas com um objeto único e com pouca complexidade de fundo do objeto. Cada imagem é particionada em 192 cortes e os cortes, centralizados e ampliados, são apresentados em $5Hz$. Todos os voluntários foram orientados a contar silenciosamente as partes das imagens de um objeto especificado
Mohedano et al. (2015)	1000 imagens divididas em 5 blocos de 200 imagens, mantendo uma proporção de "imagens alvo" de 5% dentro de cada bloco e ordenadas aleatoriamente. Todos os voluntários foram orientados a contar e pressionar um botão toda vez que eles detectaram uma "imagens alvo" apresentadas em sessões de 5Hz e 10Hz
Healy, Gurrin e Smeaton (2016)	470 imagens escolhidas de aproximadamente 1800 imagens. As imagens foram apresentadas de forma randômica através do protocolo RSVP, em uma taxa de apresentação de $5Hz$, onde os participantes foram instruídos a observar imagens de seu interesse ou atenção.

Fonte: Autor.

com outras imagens que tem o objetivo de distrair o voluntário, que precisa detectar os recortes da imagem principal.

Em Khosla et al. (2011), os autores descrevem um possível problema na detecção do sinal de P300 para imagens que possuem diferentes cores, escalas de tamanho e texturas. Estas diferenças podem gerar um "alarme falso" baseado na surpresa gerada no voluntário. Por sua

vez, implementam um certo nível de processamento na base de dados de imagens, tratando da possibilidade de normalizar as cores, fundo e textura das imagens com o objetivo de reduzir os sinais de P300 com falsos positivos na captura do EEG. Inicialmente, é implementada uma conversão de cores do padrão RGB (*Red, Green, Blue*) para HSV (*Hue, Saturation, Value*), processo trivial em visão computacional para que se tenha uma mudança de dimensão de cores. Em seguida, é calculada a distância euclideana sobre os histogramas de cor e as imagens são ordenadas e agrupadas.

6.4 BANCOS DE DADOS DE IMAGEM

Com o objetivo de comparar os resultados obtidos em um ambiente de CBIR um banco de dados de imagens precisa ser selecionado. Atualmente, existem bancos de dados fornecidos por diferentes instituições, com o objetivo de comparar os diferentes métodos e sistemas. Os artigos pesquisados fazem uso de alguns desses bancos de dados conforme descrito na Tabela 4.

Tabela 4 – Banco de Dados de Imagens utilizado.

Referências	**Image Databases**
Wang et al. (2009)	3798 imagens do banco *Caltech 101* e 62 diferentes objetos da base *Caltech and Satellite imagery* com 1051 imagens de "Helipad"
Khosla et al. (2011)	sem descrição
Ušćumlić, Chavarriaga e Millán (2011)	1382 imagens de um banco de dados proprietário com anotações baseadas no modelo *Colored Pattern Appearance Model* (CPAM) e características de contorno (*Edge-based features*)
Healy e Smeaton (2011)	4800 imagens de 1000 objetos da base *Amsterdam Library of Object Images* (ALOI), com diferentes ângulos de câmera e condições de iluminação
Ušćumlić, Chavarriaga e Millán (2013)	*Caltech 101 dataset*
Mohedano et al. (2014)	*Berkeley Segmentation Dataset and Benchmark* (BSDB)
Mohedano et al. (2015)	1000 imagens selecionadas da base *TRECVid 2013*
Healy, Gurrin e Smeaton (2016)	*Lifelog* image dataset

Fonte: Autor.

6.5 SOFTWARE DE BCI

Nenhum dos artigos pesquisados expõe a plataforma de software de BCI utilizada para o desenvolvimento dos testes, aquisição dos dados e processamento dos sinais de EEG.

6.6 VOLUNTÁRIOS E CARACTERÍSTICAS

Todos os artigos utilizam dados primários coletados em sessões de aquisição de EEG via o protocolo RSVP através dos métodos descritos anteriormente. O número total de voluntários em cada experimento é relevante como forma de comparação do volume de informação coletada. Os detalhes de cada experimento estão descritos na Tabela 5.

Tabela 5 – Detalhes sobre os voluntários utilizados.

Referências	**Voluntários**
Wang et al. (2009)	4 alunos graduados ou não e funcionários da faculdade que não trabalham com análise de mídia, mas que estavam familiarizados com o trabalho de EEG
Khosla et al. (2011)	sem descrição
Ušćumlić, Chavarriaga e Millán (2011)	1 voluntário para a classificação do banco de imagens. A etapa de consulta no banco de dados foi desenvolvida a partir de dados sintéticos que emulam a decodificação de um sinal de EEG com uma taxa de erro de 30% e 60% respectivamente
Healy e Smeaton (2011)	8 pós-graduados e funcionários do campus, 5 masculinos e 3 femininos com uma idade média de 27.5 anos com um desvio padrão de 4.5 anos
Ušćumlić, Chavarriaga e Millán (2013)	15 voluntários com a visão normal ou corrigida para normal. Sem nenhum critério específico para o recrutamento dos voluntários
Mohedano et al. (2014)	5 voluntários com idade entre 21 e 32 anos
Mohedano et al. (2015)	8 voluntários entre 19 e 33 anos de idade, sendo: 6 mulheres e 2 homens. Dos oito usuários que participaram da tarefa, dois completaram o experimento com um RSVP a 10Hz e quatro em 5Hz. Os dois últimos usuários completaram a tarefa duas vezes, uma vez em 10Hz e uma vez em 5Hz
Healy, Gurrin e Smeaton (2016)	8 voluntários foram recrutados a partir do grupo de pesquisa e que trabalham no departamento de ciência da computação e tem compreensão do processo completo de marcação das imagens.

Fonte: Autor.

Ušćumlić, Chavarriaga e Millán (2013) efetuaram um levantamento de dados de dois voluntários através de uma sessão pública em um ambiente com muita interferência. Todos

os outros artigos, descritos de forma explícita ou não, foram conduzidos em uma ambiente de laboratório.

6.7 PRÉ-PROCESSAMENTO E PROCESSAMENTO DOS DADOS

Todos os artigos pesquisados, conforme descrito na Tabela 6, aplicam métodos de classificação baseados nos sinais de EEG, extraindo os padrões das imagens alvo e propagando esses padrões para outras imagens do banco. Nesta seção, serão descritos os métodos adotados em cada artigo.

Assim que os dados de EEG são capturados, os dados obtidos via RSVP precisam ser processados para se obter o ERP e assim gerar as características que serão propagadas para as imagens não classificadas.

Tabela 6 – Detalhes sobre os diferentes processamentos de EEG utilizados.

Referências	**Algorítimos de processamento de EEG**
Wang et al. (2009)	Método de detecção em tempo real do sinal P300 através de Análise de Discriminante Linear de Fisher (LDA) (GERSON; PARRA; SAJDA, 2006) utilizando uma janela de tempo de 100*ms*
Khosla et al. (2011)	sem descrição
Ušćumlić, Chavarriaga e Millán (2011)	Classificador Gaussiano, sem maiores detalhes
Healy e Smeaton (2011)	Máquina de Vetores de Suporte (SVM)
Ušćumlić, Chavarriaga e Millán (2013)	Método com duas etapas aplicando inicialmente análise de variáveis canônicas (CVA) e um classificador Gaussiano sobre as características mais discriminantes
Mohedano et al. (2014)	SVM como classificador com função *kernel* do tipo função de base radial (RBF) com um vetor de características de média zero e desvio padrão unitário.
Mohedano et al. (2015)	Rede Neural Convolucional (CNN) via software *Caffe* e treinada com o banco de imagens *ImageNet*
Healy, Gurrin e Smeaton (2016)	Classificador Bayesiano utilizando uma validação cruzada para se obter a melhor acurácia do modelo

Fonte: Autor.

Cabe uma observação em relação a variação de métodos de classificação aplicados aos sinais, o que demonstra uma intensa pesquisa por algorítimos que apresentam bons resultados ao tipo de teste imposto, tanto pelo protocolo RSVP quanto ao processo de classificação das respostas dos sinais de EEG obtidos através do estímulo visual utilizando imagens de CBIR.

6.8 ANÁLISE CRÍTICA

Todos os artigos pesquisados anteriormente fazem uso do sinal de potencial evocado P300 para efetuar uma correlação direta com as imagens que são alvo (ou não) e a partir daí um algoritmo de busca no banco de dados retorna um conjunto de imagens de sua preferência. A forma de avaliação na maior parte é feita através de uma comparação do resultado obtido com os dados de classificação prévia existente nos bancos de dados e disponíveis para pesquisa, como por exemplo, o banco de dados de imagens para pesquisa chamado PASCAL (WANG et al., 2009).

Acredita-se que a correlação direta pode ser estendida e granularizada para que uma melhor definição do objeto de busca no banco de dados possa ser executada. A correlação direta entre um sinal de potencial evocado e uma categoria de imagem define o início da pesquisa. Porém, uma revisão bibliográfica que trata, especificamente, da intensidade (seja de amplitude ou pela duração no tempo da faixa de resposta) de P300 em sinais de EEG descreve que existe uma correlação entre a intensidade dos sinais P300 com o grau de interesse ou atenção de um voluntário com o estímulo visual apresentado (DONCHIN; ARBEL, 2009; LUCK, 2014). Em nenhum dos artigos pesquisados este tipo de característica é extraída do conjunto de sinais de EEG.

Os autores Khosla et al. (2011) apresentam indícios de como tratar diferentes tipos de imagens que possam gerar sinais de P300; porém, no caso dessa pesquisa, descrevem este efeito como um problema que pode ser resolvido com o pré-processamento de cores das imagens, afim de evitar um potencial evocado gerado pelo choque de contraste visual. Entende-se que este tipo de ocorrência, se processado corretamente, pode ser reutilizado na camada de processamento da busca para reforçar as preferências do usuário e reclassificar de forma dinâmica utilizando sinais vindos do próprio voluntário, mas nenhum dos artigos pesquisados explora este tipo de cenário.

Na mesma linha, os sinais de potenciais de erro (*error potentials*), como por exemplo o sinal N1 (BUTTFIELD; FERREZ; MILLÁN, 2006), podem ser utilizados para qualificar os resultados de busca. A bibliografia selecionada, em nenhum momento, fez uso deste tipo de potencial evocado para a avaliação e qualificação dos resultados retornados pela consulta de dados.

Em Bellman et al. (2016), foi desenvolvido um experimento utilizando o equipamento Emotiv-EPOC onde é testada a hipótese de que pessoas conscientes e não conscientes sobre o tipo de imagem que precisa ser detectado são capazes de executar a tarefa de criar marcadores via EEG. O artigo faz uma análise negativa sobre o uso de equipamentos de mercado e conclui que não existe diferenças entre os dois tipos de cenários (conscientes e não conscientes). Porém, existe uma parte interessante sobre o processamento de diferentes tipos de potencial evocado, além do P300 e a aplicação da métrica de classificação $F1$, que combina a precisão e o *recall*.

A maior parte dos artigos não faz distinção entre os voluntários, exceto em Mohedano et al. (2015) e Bellman et al. (2016). Em Mohedano et al. (2015) existe distinção clara entre especialistas e não especialistas e em Bellman et al. (2016) existe uma distinção entre pessoas que estão conscientizadas ou não sobre a tarefa de marcar imagens de interesse.

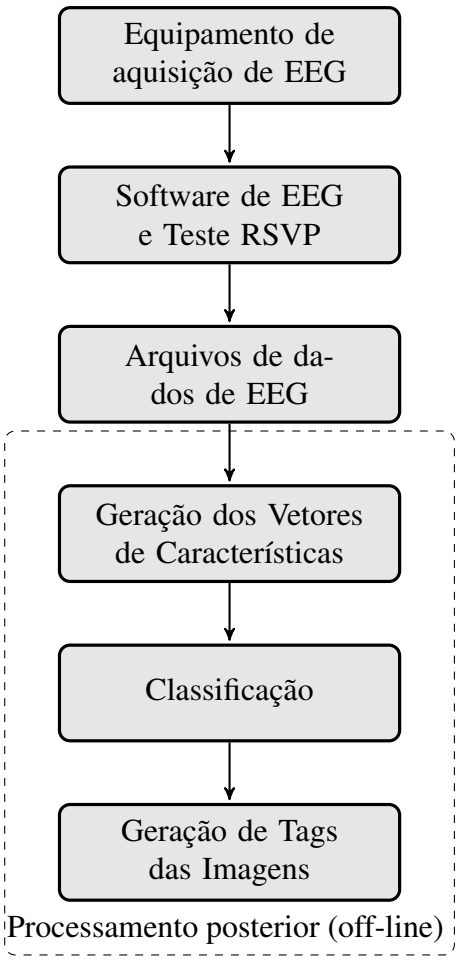
Em Lees et al. (2018), é apresentada uma discussão sobre o uso de ERP junto com RSVP para o desenvolvimento de aplicações de BCI e o artigo também apresenta um *survey* sobre os diferentes artigos nesta área. Um ponto interessantes desse trabalho é a apresentação de uma proposta de agrupamento dos diferentes tipos de aplicações nas seguintes categorias: *surveillance*, *data categorization*, *RSVP speller*, *face recognition* and *medical image analysis*. Seguindo esta categorização, os artigos apresentados, na primeira parte desse capítulo, fazem parte do segmento de *data categorization*. Os artigos comparados nesse capítulo (WANG et al., 2009; KHOSLA et al., 2011; UŠĆUMLIĆ; CHAVARRIAGA; MILLÁN, 2011; HEALY; SMEATON, 2011; UŠĆUMLIĆ; CHAVARRIAGA; MILLÁN, 2013; MOHEDANO et al., 2014; MOHEDANO et al., 2015; HEALY; GURRIN; SMEATON, 2016) fazem parte da segmentação apresentada por Lees et al. (2018).

7 MATERIAL E MÉTODOS

Neste capítulo será descrito o processo em que um voluntário é apresentado para um cenário de RSVP, com um banco de dados de imagens, e como os sinais de EEG são capturados e processados com o objetivo de segregar o grupo de imagens.

Na Figura 19, é possível observar as etapas do modelo de processamento dos sinais de EEG adotado.

Figura 19 – Fluxo de aquisição de dados de EEG e método de classificação de imagens.



Fonte: Autor.

7.1 MATERIAL

7.1.1 Banco de dados de imagens

A base para a criação dos experimentos são as bases de dados de imagens e para este projeto foram utilizados dois conjuntos de imagens: Caltech 101 (FEI-FEI; FERGUS; PERONA, 2007) e Nausitoo (REZENDE; BECHELLI, 2015).

Para o primeiro experimento foi utilizada a base de dados Caltech 101 para efetuar os primeiros levantamentos de EEG com os voluntários e validar a metodologia. Foram selecionados 54 categorias genéricas e com, aproximadamente, a mesma quantidade de imagens para que a sua apresentação fosse distribuída de maneira uniforme, resultando num total de 3436 imagens para a montagem dos testes (BECHELLI, 2016b).

As categorias selecionadas foram: *butterfly, camera, cannon, crab, crayfish, crocodile, cup, dalmatian, dolphin, elephant, emu, flamingo, ibis, inline_skate, joshua_tree, kangaroo, lamp, llama, lobster, lotus, mandolin, mayfly, menorah, metronome, minaret, nautilus, octopus, okapi, pagoda, panda, pigeon, pizza, platypus, pyramid, revolver, rhino, rooster, saxophone, schooner, sea_horse, snoopy, soccer_ball, stapler, starfish, stegosaurus, stop_sign, strawberry, sunflower, tick, trilobite, umbrella, watch, water_lilly, wild_cat, windsor_chair, wrench e yin_yang*. Alguns exemplos de imagens são apresentados na Figura 4.

No segundo experimento foi utilizado um compêndio de imagens que possuem o mesmo contexto e um grupo de voluntários especialistas formados em arqueologia e que trabalharam com o desenvolvimento do banco de dados Nausitoo. Para isso foi utilizado um conjunto de imagens extraídas do banco de dados de imagens Nausitoo. Deste banco foram selecionadas 33 categorias de imagens num total de 5121 imagens.

As categorias selecionadas foram: *ânfora, castelo, fosso, moeda, teatro, coluna, base de coluna, fragmentos, mosaico, templo, bloco, corte, kouros, terracota, bronze, coruja, mapa, oliveira, tumba, cabeça, estádio, maquete, painel, vaso, caminho, estátua, máscara, planta, vista, canal, fonte, modelo e reconstituição*. Alguns exemplos de figuras são apresentados na Figura 5.

7.1.2 Aquisição de EEG

Para a aquisição dos dados de EEG foi utilizado o dispositivo iCelera iBlue 52 (RODRIGUES; SILVERIO, 2014) que possui 20 eletrodos de EEG posicionados segundo o padrão internacional 10–20, 2 eletrodos de referência nas posições *A1 e A2*, uma taxa de aquisição de 512 medidas por segundo e uma faixa de operação em frequência que varia de $0{,}3Hz$ até $50Hz$. O equipamento possui um filtro *anti-alias* de $160Hz$ na entrada de cada canal. Do ponto de vista do posicionamento dos eletrodos, o dispositivo possui eletrodos para a aplicação com gel

úmido, nas posições fixas: *C3, C4, Cz, F3, F4, F7, F8, Fp1, Fp2, Fz, O1, O2, Oz, P3, P4, Pz, T3, T4, T5, T6*, conforme os eletrodos preenchidos na cor cinza na Figura 10 no item (c).

7.2 VOLUNTÁRIOS

Foram selecionados 11 voluntários, que fazem parte da comunidade acadêmica da universidade (Fundação Educacional Inaciana (FEI) e MAE), para que fosse aplicado o teste de RSVP. Destes, 5 (cinco) foram considerados especialistas no banco de dados Nausitoo, pois trabalham na área de arqueologia, e 6 (seis) considerados não especialistas.

A Tabela 7 apresenta as características de cada voluntário. A idade dos voluntários intervalo de 20 à 44 anos com média de 32. Para os especialistas, foram efetuadas 2 sequências de testes, a primeira utilizando as imagens do banco de dados Caltech 101 e uma segunda utilizando as imagens do banco de dados Nausitoo. Para o grupo de não especialistas os testes foram efetuados apenas com o banco de dados Caltech 101.

Tabela 7 – Detalhes sobre os voluntários.

Voluntário	DB	Categoria	Especialista	Idade	Gênero
Voluntário01	caltech101	elephant	0	20	m
Voluntário02	caltech101	elephant	0	21	m
Voluntário03	caltech101	elephant	0	25	m
Voluntário04	caltech101	elephant	0	40	m
Voluntário05	caltech101	elephant	0	34	m
Voluntário06	caltech101	elephant	0	26	f
Voluntário07	caltech101	elephant	1	26	f
Voluntário07	nausitoo	moeda	1	26	f
Voluntário08	caltech101	elephant	1	44	f
Voluntário08	nausitoo	moeda	1	44	f
Voluntário09	caltech101	elephant	1	36	f
Voluntário09	nausitoo	moeda	1	36	f
Voluntário10	caltech101	elephant	1	33	f
Voluntário10	nausitoo	moeda	1	33	f
Voluntário11	caltech101	elephant	1	34	m
Voluntário11	nausitoo	moeda	1	34	m

Fonte: Autor.

7.3 MÉTODO

7.3.1 Procedimento experimental

Em todos os artigos pesquisados, o protocolo RSVP tem a necessidade de um sincronismo de tempo entre o momento em que as imagens são apresentadas e a captura dos sinais de EEG. Outros requisitos, apresentados em todos os artigos no Capítulo 6, são o pré-processamento das imagens para que todas tenham as mesmas dimensões de altura e largura e a necessidade da criação de um conjunto de sessões para que o voluntário possa ter um intervalo de descanso entre a apresentação dos estímulos visuais para minimizar ou excluir a questão da fadiga visual.

Partindo desse conjunto de requisitos, foi desenvolvido um processo de software automatizado para gerar todas as etapas necessárias para a construção do experimento.

Este processo é constituído pelas seguintes etapas:

a) Download das bases de dados de imagens e anotações, extração das imagens, armazenamento e criação do índice de imagens para cada banco de imagens (Caltech 101 e Nausitoo);
b) Captura dos dados do voluntário (nome, idade, gênero etc.) e criação da área onde serão armazenados os dados do experimento e arquivos relacionados com as sessões de RSVP;
c) Geração do experimento junto ao software que efetua a coordenação entre a apresentação do experimento e a aquisição sincronizada dos dados.

As bases de dados de imagens, na maioria das vezes, são compostas por um conjunto de imagens separadas em pastas e compactadas em um único arquivo e em alguns casos possuem arquivos de anotações adicionais para fornecer informações adicionais sobre agrupamento, categorias e contornos, entre outros.

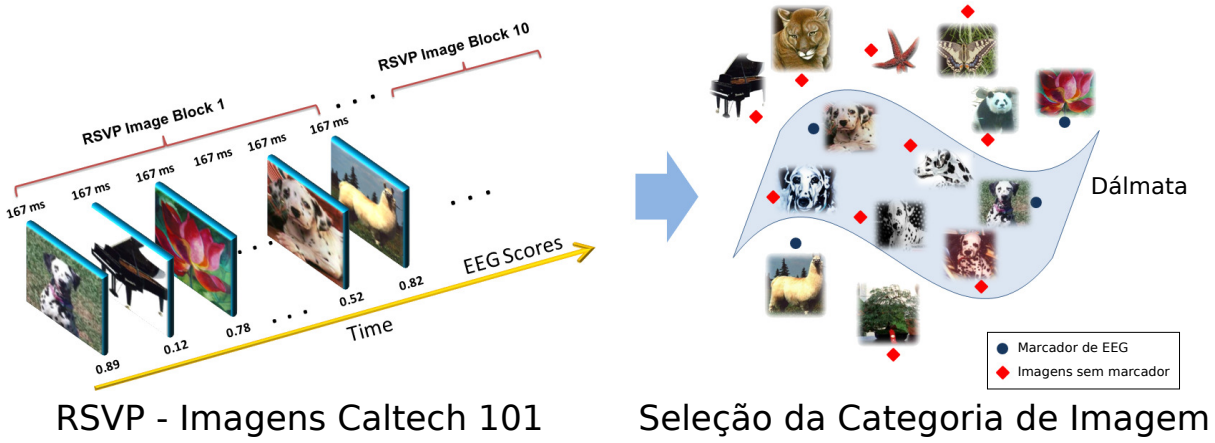
A partir desse arquivo compactado, com o objetivo de normalizar as duas bases de imagens para o experimento proposto, o sistema de software desenvolvido gera um arquivo com os dados essenciais do caminho de cada imagem e sua categoria.

Para cada voluntário, que participa do levantamento de dados, é gerado um diretório e para cada experimento são coletadas informações básicas de cada pessoa. Estes dados são armazenados para uso posterior em um diretório único, junto com todos os arquivos do experimento específico do voluntário. O sistema de software permite a confecção de um conjunto único de imagens que serão apresentadas para cada sessão, baseadas na categoria escolhida com imagens que serão os alvos.

O protocolo RSVP, conforme apresentado na Figura 20, foi desenvolvido para uma resolução de tela de $1024 \times 768px$. O experimento é iniciado com um tempo de intervalo de $5s$ onde é apresentado uma tela preta. Após esta etapa inicial, as imagens são apresentadas de forma sequencial, sobre um fundo preto, com uma taxa de apresentação de $4Hz$, equivalente a

4 imagens por segundo. Os testes são efetuados em 4 sessões, de aproximadamente 4 minutos, com sequências diferentes de imagens, num total de 60 imagens alvo e um total de 600 à 800 imagens não–alvo. A variação entre imagens não–alvo ocorre porque os testes são montados com uma aleatoriedade entre as imagens alvo utilizando um intervalo mínimo de 6 até um intervalo máximo de 20 imagens não–alvo entre as imagens alvo. Após cada uma das sessões, é oferecido um intervalo de 5 minutos para descanso. Foi solicitado aos voluntários que fizessem uma contagem silenciosa das imagens alvo.

Figura 20 – Método de RSVP.



Fonte: "Adaptado de" Wang et al., 2009.

Para cada sessão do experimento é necessário efetuar um ajuste de tamanho das imagens de $300px$ de altura por $300px$ de largura. Dessa forma, foi desenvolvido um processo que faz um ajuste no tamanho das imagens, aumentando (ou diminuindo) as dimensões originais, e recortando cada imagem a partir do centro para atingir as medidas especificadas.

Os processos descritos, nos parágrafos anteriores, podem ser efetuados diversas vezes para se obter o número total de sessões necessárias para o levantamento de dados. Após a geração da sequência de apresentação das imagens, para cada uma das sessões, é gerada uma camada de integração responsável por sincronizar a camada de apresentação do experimento, desenvolvida em Python (ROSSUM, 1995), e a interface de aquisição do hardware de captura de sinais de EEG. Para o desenvolvimento dos testes foi utilizada uma interface, também em Python, com o equipamento iCelera (BECHELLI, 2016b).

7.3.2 Processamento dos dados

Após a aquisição, uma etapa de pré-processamento foi aplicada ao sinal. Inicialmente, o sinal foi ajustado pela média dos dois eletrodos de referência nas posições *A1 e A2*. Em seguida, cada evento de exposição de imagem foi separado em um conjunto de janelas variando de $-0.2s$ à $+0.8s$, a partir do instante onde ocorre a apresentação de cada imagem. Foram aplicados uma

sequência de filtros, como o filtro de faixa de rejeição em $60Hz$ para remover possíveis ruídos provenientes da rede elétrica, um filtro passa banda de $0.3Hz$ à $40Hz$ removendo os níveis de corrente contínua, em todas as janelas o sinal foi subtraído do valor médio total (DUFFY; IYER; SURWILLO, 1989) e um filtro de amplitude de $140\mu V$ pico a pico também foi aplicado para cada janela com o objetivo de rejeitar todas as janelas de dados que possuem algum tipo de ruído, como por exemplo, movimentação dos eletrodos ou eventuais ruídos provenientes da contração muscular da face, preparando o sinal para a etapa de processamento dos dados.

7.3.3 Avaliação dos sinais de P300

Após a aquisição dos dados e o pré-processamento, uma primeira avaliação dos eventos foi desenvolvida utilizando o método descrito por Luck (2014), onde é aplicado em cada um dos eletrodos, o cálculo da média de todos os eventos marcados como alvo e comparado com a média de todos os eventos marcados como não-alvo. Em posse dessas duas curvas, foi feita a comparação visual das duas curvas para avaliar a existência da uma deflexão positiva próxima de $300ms$, conforme a Figura 12.

7.3.4 Classificação individual das imagens alvo

Uma outra forma de se classificar testes de RSVP é utilizando os sinais individuais de cada uma das janelas de tempo. Este método é apresentado principalmente por Blankertz et al. (2011), este método é conhecido como classificação individual dos sinais, do inglês (*Single Trial Detection*), onde será utilizado o método das wavelets em conjunto com um classificador de ML.

Para isso, os dados são processados via DWT. Foram utilizados no experimento um conjunto de 106 famílias e ordens de wavelets. As famílias utilizadas no experimento foram: Biortogonal (bior), Coiflets (coif), Daubechies (db), Aproximação discreta da wavelet de Meyer (dmey), Haar (haar), Biortogonal Reversa (rbio), Symlets (sym) em diferentes ordens (WASILEWSKI, 2006). Para cada janela de dados, foram utilizados os níveis máximos de decomposição disponíveis para a passagem dos dados para o domínio das wavelets.

Os vetores resultantes da etapa anterior, já no domínio das *wavelets*, foram aplicados a um classificador de aprendizado de máquina pelo classificador de *Ridge*. Onde, para uma melhor análise, afim de evitar mínimos locais, os dados de treinamento e teste foram embaralhados e dobrados (*shuffling and folding*), pelo menos cinco vezes (20% para teste para 80% treinamento) durante o treinamento de aprendizado de máquina para que fosse efetuada a validação cruzada e a pontuação média foi usada para obter os melhores parâmetros. Em cada etapa deste processo foi aplicado um processo para identificar o melhor parâmetro de α que resultasse no resultado mais alto de classificação de acurácia (*gridsearch*) (PEDREGOSA et al., 2011). Uma vez encontrado o melhor resultado de α, o classificador é executado novamente para se obter os

valores de *verdadeiros positivos, falsos positivos, verdadeiros negativos e falsos negativos* da matriz de confusão e cálculo matemático de fatores como $F1$, descrito na Equação 6, para posterior análise. O fator $F1$ será utilizando como referência para todas as analises, pois equilibra a PR e o RC, oferecendo uma qualidade maior para os resultados obtidos em bases de dados desbalanceadas, como é o caso dos dados adquiridos.

7.3.5 Seleção da família de wavelets

Uma abordagem estatística foi proposta para selecionar a melhor família de wavelets e usar essa informação para comparar especialistas e não especialistas. Uma taxa de classificação mínima de $F1$ de 0,7 (70%) foi definida para excluir resultados de baixa qualidade de classificação de aprendizado de máquina. Também foi realizado um teste de Análise de Variância (ANOVA), com todas as famílias de wavelets contra a hipótese nula. Se o teste ANOVA fornecer separação estatística entre os grupos, será conduzida uma *análise post hoc* para comparar as famílias individuais entre si usando o método de teste de diferenças significativas de Tukey (*Honest Significant Difference* (HSD)) (TUKEY, 1949) e avaliar quais famílias têm resultados que podem ser rejeitadas. Caso, após as etapas de exclusão descritas anteriormente, não seja suficiente para definir uma única família, o critério final será a maior quantidade de ordens.

7.3.6 Comparação entre os grupos de especialistas e não especialistas

Após a seleção da família de wavelets e ordem, propõe-se a aplicação do teste de independência *qui-quadrado* (χ^2) sobre os resultados positivos dos eventos processados pela matriz de confusão de aprendizado de máquina comparando grupos de especialistas e não especialistas, e em seguida utilizando o mesmo teste para comparar a relação dos especialistas com as bases de dados de imagens.

O objetivo é avaliar duas hipóteses em função dos resultados dados como positivos pela matriz de confusão do classificador (verdadeiros positivos e falsos negativos): a hipótese nula H_0, assume que não há associação entre os grupos de especialistas e não especialistas e os resultados positivos (independência entre variáveis); a hipótese H_1 assume que existe uma associação entre as variáveis (sem independência entre variáveis). Para esta análise, foi aplicado o Teste Qui-Quadrado da Independência (χ^2) sobre os resultados positivos dos eventos.

7.3.7 Comparação entre as bases de dados para o grupo de especialistas

A mesma avaliação foi desenvolvida para o grupo de especialistas em relação ao teste efetuado com as duas bases de dados de imagens: Caltech 101 e Nausitoo. Mas nesse caso, a hipótese nula H_0, assume que não há associação entre resultados positivos (independência entre variáveis) para os especialistas sendo expostos às duas bases, Caltech 101 e Nausitoo; e

a hipótese H_1, assume que existe uma associação entre as variáveis (sem independência entre variáveis).

7.3.8 Teste do classificador

Para confirmar a possibilidade de generalização do treinamento do classificador de *ridge*, foi solicitado que um dos especialistas (Voluntário10) fizesse uma aquisição adicional de EEG em cada uma das bases de dados (Caltech 101 e Nausitoo). Dessa forma, o treinamento do classificador foi efetuado conforme descrito na Tabela 8 e o teste de generalização do classificador foi efetuado com a aquisição adicional.

7.4 COMITÊ DE ÉTICA EM PESQUISA CIENTÍFICA

Todo o protocolo experimental deste projeto foi aprovado pelo Comitê de Ética em Pesquisa Científica da Universidade Metodista de São Paulo (UMESP) em 23 de Março de 2016, com identificação CAAE [1]: 52925116.0.0000.5508, sob designação do Sistema Nacional de Ética em Pesquisa (SISNEP) através da Plataforma Brasil.

Uma cópia do protocolo de aprovação é apresentada no Anexo A e uma cópia do termo de consentimento livre e esclarecido, que foi assinado pelos voluntários, também é apresentada no Apêndice B.

[1] http://aplicacao.saude.gov.br/plataformabrasil/

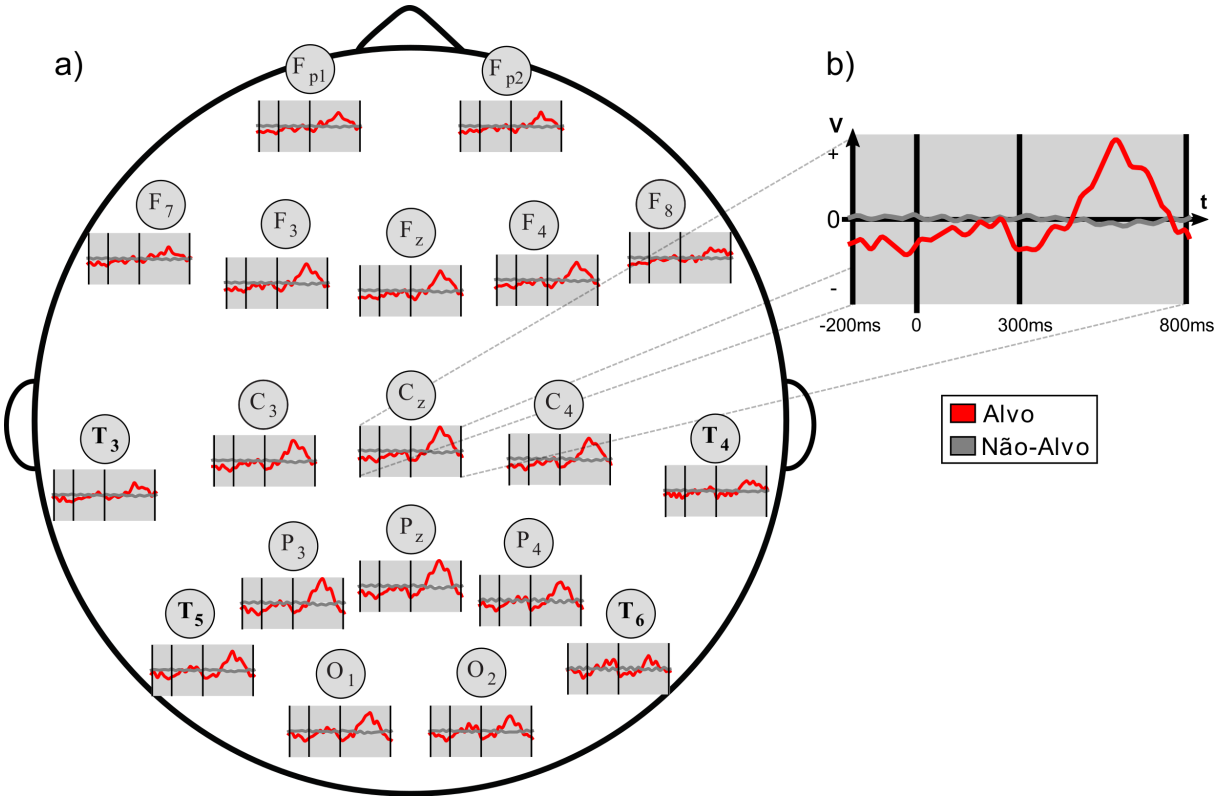
8 RESULTADOS

8.1 AVALIAÇÃO DOS SINAIS DE P300

Foram geradas as curvas de média de todos os sinais adquiridos, conforme apresentado na Figura 21. Cada uma das curvas representa a média dos sinais para cada eletrodo correspondente às posições descritas na Figura 10 para a montagem de 20 eletrodos (eletrodos marcados em cinza). As duas linhas verticais representam as marcações de tempo. Para a geração das curvas da Figura 21, foi aplicado um filtro passa banda entre 0.3 e $20Hz$ para facilitar a visualização dos sinais de ERP, no restante do projeto foram utilizados os padrões apresentados na Seção 7.3.2.

Utilizando como exemplo a Figura 21, nos eletrodos *C3, Cz, C4, P3, Pz, P4, O1 e O2*, é possível identificar uma variação entre as duas curvas. O pico positivo é o marcador que identifica o potencial evocado P300 para as imagens que são "alvo", conforme a apresentado na Figura 12.

Figura 21 – Média de sinal potencial evocado P300 em cada eletrodo para o Voluntário05 com a base Caltech 101: a) Visão de todos os eletrodos e b) Detalhe de um único eletrodo com as marcações de tempo a partir da apresentação da imagem.

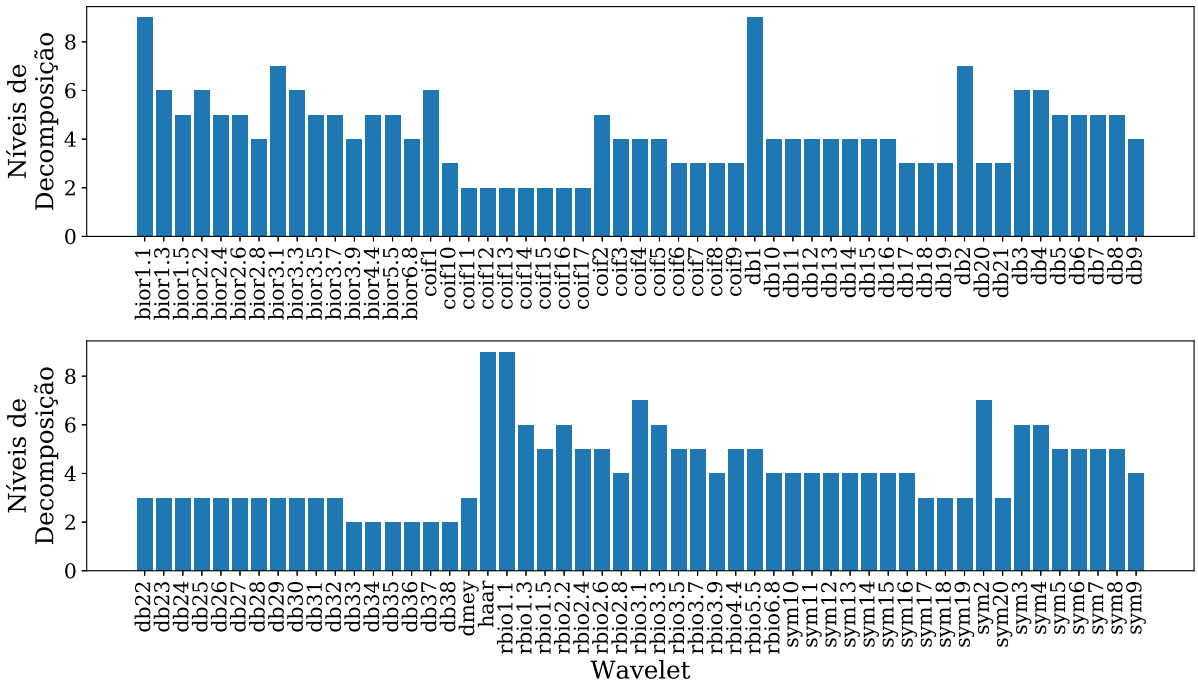


Fonte: Autor.

8.2 CLASSIFICAÇÃO INDIVIDUAL DAS IMAGENS ALVO

Para a etapa de extração de características foram utilizados no experimento um conjunto de 106 famílias e ordens de wavelets (WASILEWSKI, 2006). Os sinais foram decompostos através do método do nível máximo possível de decomposição em função das janelas de sinais de EEG. Na Figura 22 são apresentados o resultado do cálculo que determina quais são os níveis máximos de decomposição que serão aplicados para cada uma das wavelets utilizadas.

Figura 22 – Níveis de decomposição adotados para cada wavelet.



Fonte: Autor.

A partir dos resultados obtidos no cálculo dos níveis de decomposição, todos os sinais foram processados para as 106 wavelets disponíveis, gerando um conjunto de características para a entrada no classificador. Na Tabela 8 é possível observar em detalhe os vetores gerados para a inserção no classificador. Onde "Voluntários" representa cada um dos voluntários, "Características" é a família de wavelet utilizada, "Vetor" é a quantidade de levantamentos ou janelas de dados, "Dim" é a dimensão de cada um dos levantamentos, "Alvo" é a quantidade de alvos, "Não-Alvo" é a quantidade de não-alvos e "%" é o percentual entre alvos e não-alvos.

8.2.1 Seleção da família de wavelets

Os vetores foram aplicados ao classificador de *ridge* com o objetivo de identificar a qualidade dos resultados obtidos e eliminar as famílias de wavelets que obtivessem uma baixa qualidade de resultados para processar o restante dos dados. No momento em que foi necessário

Tabela 8 – Dimensão dos vetores por característica.

Voluntário	Característica	Vetores	Dim	Alvo	Não-Alvo	%
Voluntário01	DB32	3092	11917	200	2892	6.91
Voluntário02	DB32	3309	14721	229	3080	7.43
Voluntário03	DB32	2882	14721	199	2683	7.41
Voluntário04	DB32	3464	14721	232	3232	7.17
Voluntário05	DB32	3582	14721	238	3344	7.11
Voluntário06	DB32	3577	11917	237	3340	7.09
Voluntário07	DB32	3494	12618	223	3271	6.81
Voluntário07	DB32	2880	14721	188	2692	6.98
Voluntário08	DB32	3116	14721	206	2910	7.07
Voluntário08	DB32	3170	14721	209	2961	7.05
Voluntário09	DB32	2436	14721	165	2271	7.26
Voluntário09	DB32	2562	14721	171	2391	7.15
Voluntário10	DB32	1926	14721	105	1821	5.76
Voluntário10	DB32	2523	14721	115	2408	4.77
Voluntário11	DB32	2680	14721	173	2507	6.90
Voluntário11	DB32	1980	14020	135	1845	7.31

Fonte: Autor.

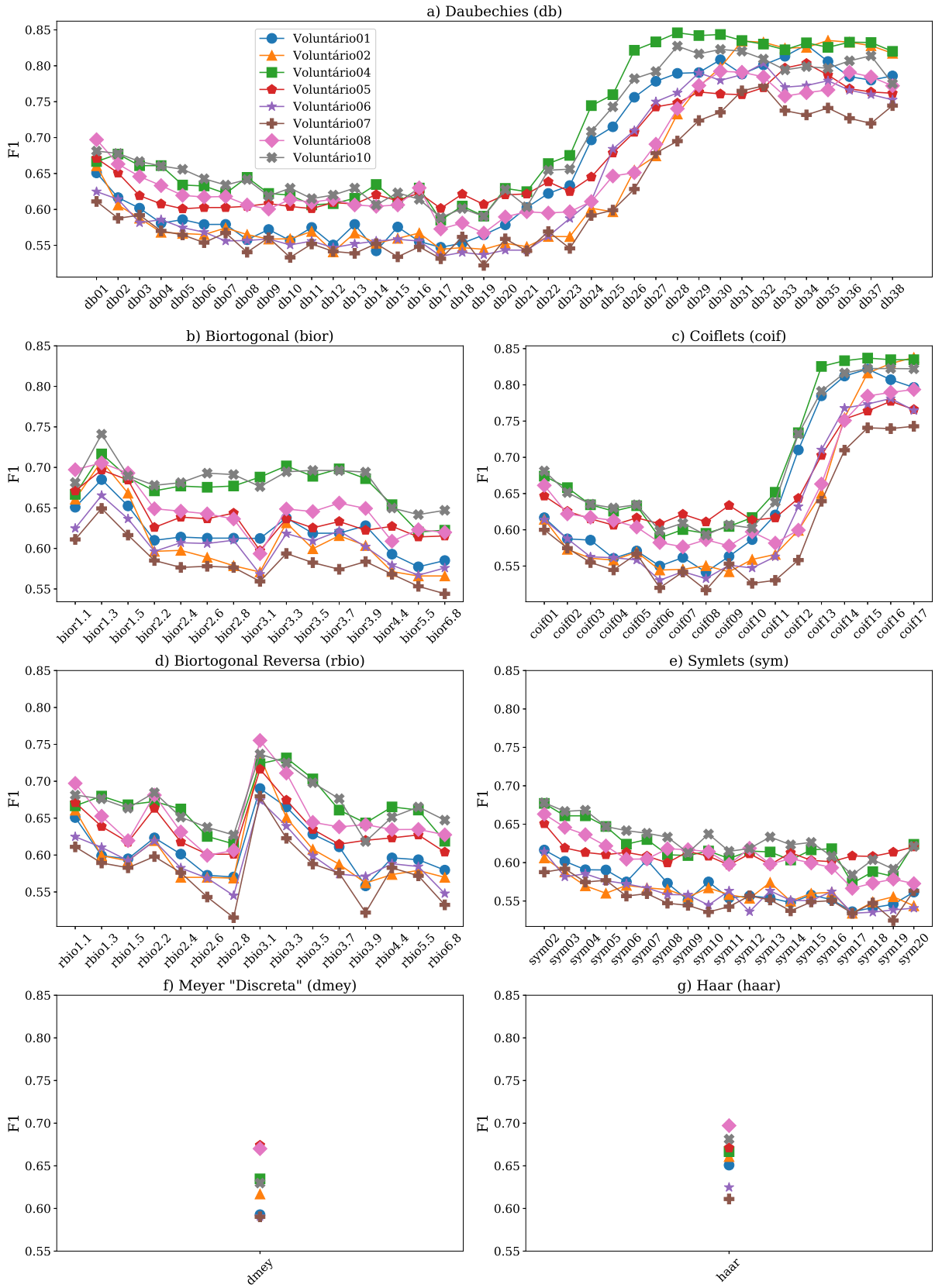
avaliar as famílias de wavelets, havia apenas um subconjunto de aquisições efetuadas com os voluntários, e foi decidido seguir adiante com esta etapa, afim de definir a melhor família de wavelets para que fosse utilizada no restante do trabalho.

Os resultados são apresentados na Figura 23, onde é apresentada a pontuação total $F1$ para cada família e ordem de wavelet processada para cada voluntário. As famílias estão agrupadas para uma melhor comparação entre si. Neste momento, é possível verificar um padrão visual semelhante nos resultados de classificação para todos os voluntários e também é possível verificar que algumas famílias apresentam melhores resultados frente ao conjunto de dados testados.

Uma visão mais detalhada é apresentada na Figura 24, onde é possível avaliar os resultados de $F1$ agrupado por famílias de wavelets. A linha horizontal marca famílias de wavelets com taxa de classificação abaixo de 0.7 (70%). Também conclui a avaliação estatística da Figura 23.

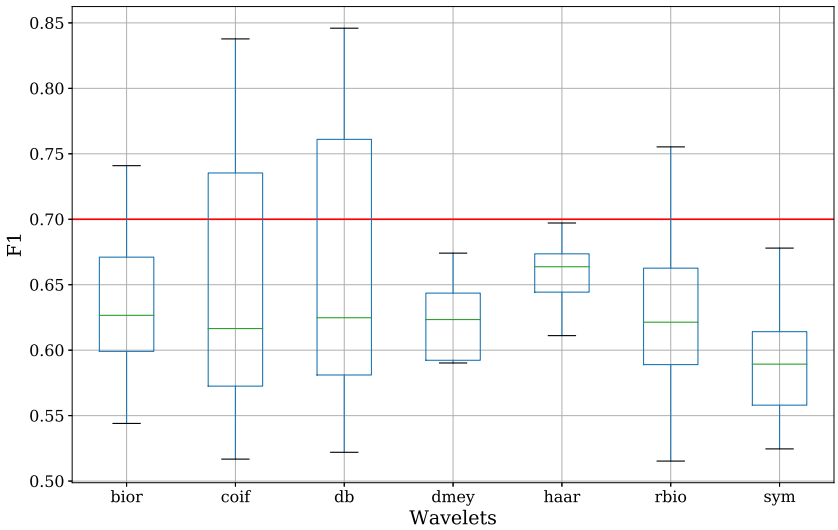
A pontuação ANOVA (F_{ANOVA}) obtida foi 17.025 e *p-value* de $1.264e^{-18}$. Desde que os dados sejam testados pela ANOVA e demonstrem que há diferença significativa entre os grupos selecionados, o próximo passo é executar um teste de diferença mínima significativa de Tukey

Figura 23 – Valores de $F1$, por voluntário, agrupados por família de wavelet.



Fonte: Autor.

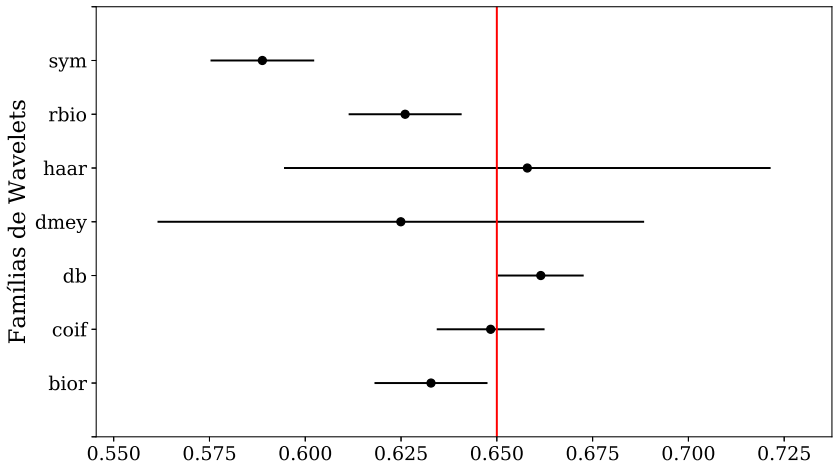
Figura 24 – Boxplot do valor de $F1$ agrupado por família de wavelet.



Fonte: Autor.

(HSD) para uma análise *post-hoc*. A Figura 25, representa a comparação de resultados de cada família e a linha vertical define um valor limite de 0,65 na análise da diferença média para as famílias que apresentam os melhores resultados. As famílias que têm os resultados abaixo da linha de corte definida, podem ser excluídas, para que se possa selecionar os grupos mais significativos.

Figura 25 – Teste de multicomparação de Tukey entre todas as famílias de wavelets.



Fonte: Autor.

Para esta etapa do experimento, com o objetivo de definir a melhor família de wavelets a ser usada posteriormente, foi feita uma comparação dos resultados extraídos da Figura 24 e

da Figura 25. As famílias que não obtiverem taxas superiores a 70% de $F1$ em qualquer teste foram excluídas, e as que não passarem no teste de Tukey HSD também. Outra consideração, foi o fato de que nem todas as famílias disponíveis neste teste têm a mesma quantidade de derivações, sendo este, um terceiro fator considerado, já que a quantidade maior de derivações pode fornecer mais dados estatísticos para a seleção da família.

Foram utilizadas neste experimento, um total de 7 (sete) famílias de wavelets: Biortogonal (BIOR), Coiflets (COIF), Daubechies (DB), Aproximação discreta da wavelet de Meyer (DMEY), Haar (HAAR), Biortogonal Reversa (RBIO) e Symlets (SYM) em diferentes ordens. Avaliando os resultados da Figura 24, as famílias DMEY, HAAR, e SYM podem ser excluídas pela taxa de classificação $F1$ abaixo de 70%. Na Figura 25, as famílias SYM, RBIO e BIOR, foram excluídas pelo teste de Tukey HSD. Das 7 (sete) famílias apenas COIF e DB, apresentaram os melhores resultados no critério adotado. Para definir qual família dessas duas opções, foi adotada a família de wavelets que possui a maior quantidade de ordens disponíveis. A família DB, que possui 38 ordens, foi escolhida contra COIF que possui apenas 17 ordens (WASILEWSKI, 2006). Da família DB, o melhor resultado foi obtido com Daubechies de ordem 32 (DB32), com precisão de 97% e resultado de $F1$ de 80% de média para todos os voluntários.

8.2.2 Comparação entre os grupos de especialistas e não especialistas

A partir da avaliação das famílias de wavelets, foi selecionado para este teste a wavelet DB32 para comparar os dois grupos de especialistas e não especialistas. Na Figura 26, é possível observar os resultados obtidos através do classificador de *ridge*. Para os levantamentos, a Área abaixo da curva ROC (AUC) foi de $0.999 \pm 1.092x10^{-4}$. O detalhe dos dados é apresentado na Tabela 9.

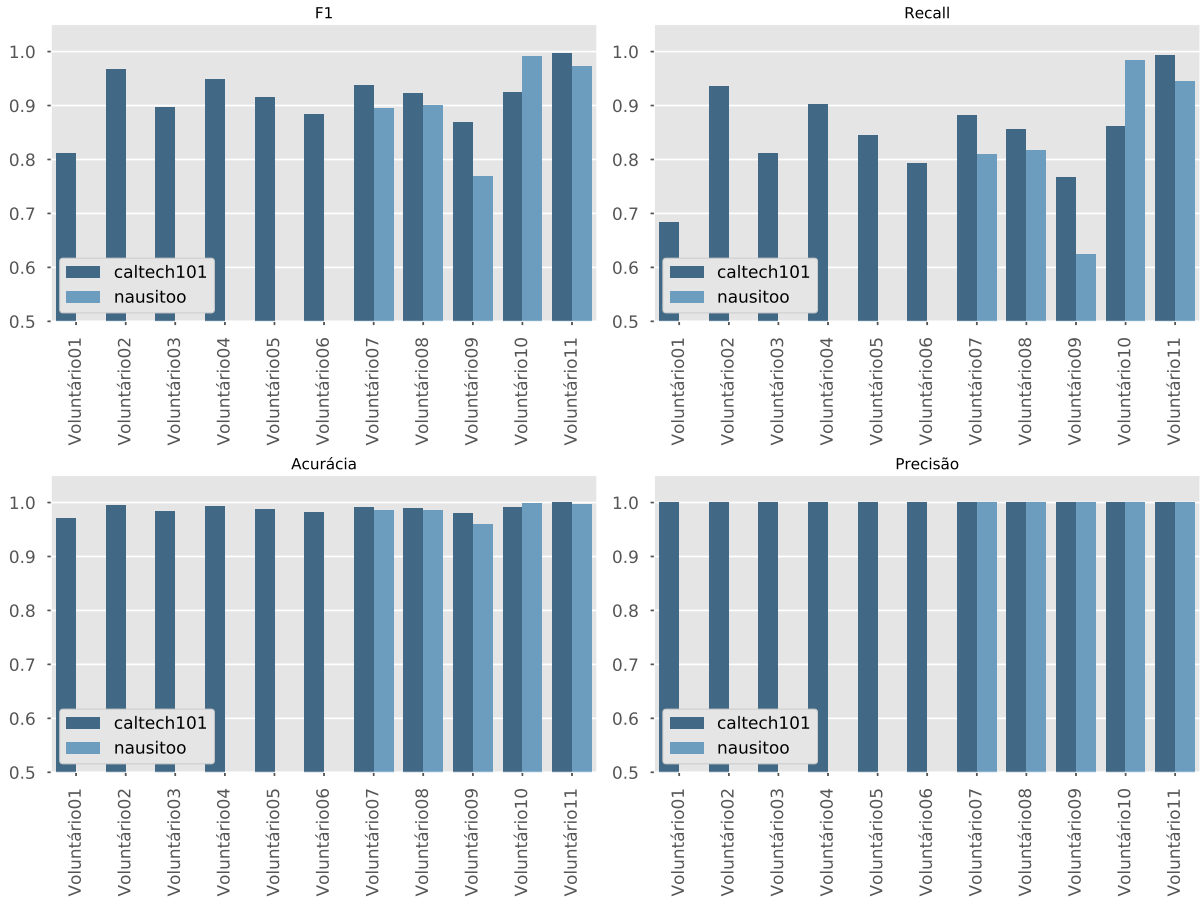
Adotando os resultados da classificação wavelet DB32 para comparar os dois grupos de especialistas e não especialistas, os valores de verdadeiros positivos para especialistas e não especialistas são respectivamente, 837 e 1335, e os valores de falsos negativos são respectivamente, 133 e 286.

Aplicando o teste de independência *qui-quadrado* com 5% de significância estatística e com 1 grau de liberdade, o resultado é um χ^2 de 6.634 com um *p-value* de 0.010, o χ^2_c (valor crítico de *qui-quadrado*) é 3.841. Com os resultados estatísticos, rejeita-se a hipótese H_0, considerando-se portanto a existência de associação entre os resultados positivos dos especialistas e os grupos de especialistas (H_1).

8.2.3 Comparação entre as bases de dados para o grupo de especialistas

Da mesma forma, adotando os resultados da classificação wavelet DB32 para comparar as duas bases de dados de imagens, os valores de verdadeiros positivos para as bases Caltech 101

Figura 26 – Comparação dos resultados do classificador para a família DB32 para todos os voluntários.



Fonte: Autor.

e Nausitoo são respectivamente, 837 e 853, e os valores de falsos negativos são respectivamente, 133 e 205.

Aplicando o teste de independência *qui-quadrado* com 5% de significância estatística e com 1 grau de liberdade, o resultado é um χ^2 de 11.288 com um *p-value* de $7.801x10^{-4}$, o χ^2_c (valor crítico de *qui-quadrado*) é 3.841. Com os resultados estatísticos, rejeita-se a hipótese H_0, considerando-se portanto a existência de associação entre os resultados positivos dos especialistas para os dois bancos de dados de imagens adotados (H_1).

8.2.4 Teste de generalização do classificador

Para este teste, foi obtido o resultados de $F1$ de 74% (AUC de 0.61) a base de dados Caltech 101 e 88% (AUC de 0.69) com a base de dados Nausitoo.

Tabela 9 – Tabela de comparação dos resultados do classificador para a família DB32 para todos os voluntários.

Voluntário	DB	TP	FP	TN	FN	PR	RC	ACC	F1
Voluntario01	caltech101	200	0	2799	93	1.0	0.682	0.969	0.811
Voluntario02	caltech101	229	0	3064	16	1.0	0.934	0.995	0.966
Voluntario03	caltech101	199	0	2637	46	1.0	0.812	0.984	0.896
Voluntario04	caltech101	232	0	3207	25	1.0	0.902	0.992	0.948
Voluntario05	caltech101	238	0	3300	44	1.0	0.843	0.987	0.915
Voluntario06	caltech101	237	0	3278	62	1.0	0.792	0.982	0.884
Voluntario07	caltech101	223	0	3241	30	1.0	0.881	0.991	0.936
Voluntario07	nausitoo	188	0	2648	44	1.0	0.810	0.984	0.895
Voluntario08	caltech101	209	0	2926	35	1.0	0.856	0.988	0.922
Voluntario08	nausitoo	206	0	2864	46	1.0	0.817	0.985	0.899
Voluntario09	caltech101	165	0	2221	50	1.0	0.767	0.979	0.868
Voluntario09	nausitoo	171	0	2288	103	1.0	0.624	0.959	0.768
Voluntario10	caltech101	105	0	1804	17	1.0	0.860	0.991	0.925
Voluntario10	nausitoo	115	0	2406	2	1.0	0.982	0.999	0.991
Voluntario11	caltech101	135	0	1844	1	1.0	0.992	0.999	0.996
Voluntario11	nausitoo	173	0	2497	10	1.0	0.945	0.996	0.971

Fonte: Autor.

9 DISCUSSÃO

O uso das wavelets, como base para o desenvolvimento da pesquisa, se dá através de, conforme descrito por Luck (2014), eventos de ERP que ocorrem próximos, mas com uma certa variação em relação aos tempos utilizados para marcar os eventos, como por exemplo os $300ms$ para os eventos de P300. A temporalidade dos evento de P300, em conjunto com a disposição espacial dos eletrodos no topo da cabeça e a possibilidade de decompor o sinal via a aplicação da transformada de wavelet, geram um vetor resultante que é utilizado para criar um dicionário de eventos de P300 ao longo de todo o trabalho desenvolvido.

O uso de sinais de EEG, para a aplicação em CBIR, apresenta um imenso desafio, tanto do ponto de vista de aplicação, quanto das diversas formas de desenvolvimento apresentadas pelos diversos artigos consultados nesse trabalho. O método proposto por Donchin, Ritter e McCallum (1978), para o desenvolvimento das pesquisas em potencial evocado, percorreu um grande caminho até a aplicação descrita por Wang et al. (2009) para a classificação de bancos de imagens. Mesmo em 2007, Liu et al. (2007) não menciona a aplicação das técnicas de BCI como forma alternativa de redução das descontinuidades semânticas através de ERP, conforme apresentado na Figura 3. O uso de modelos de aprendizado de máquina, aliados com sinais cerebrais, apresentam caminho novo para a redução, ou quem sabe, num futuro, a eliminação destas descontinuidades.

Após as aquisições desenvolvidas nos experimentos, foi possível comparar os resultados apresentados na Seção 6.1. Todas as pesquisas apresentam métodos desenvolvidos com ERP, extraídos dos sinais de EEG. O método de RSVP, também utilizado em todos os artigos, serviram de base para o desenvolvimento do método de aquisição. Utilizando uma frequência de apresentação de $4Hz$, na média dos experimentos, são apresentadas, aproximadamente, 4000 imagens num intervalo médio de $20min$. (sem os intervalos de pausa entre os experimentos). Dessa forma, pode-se inferir que o volume de classificação destas imagens, se comparado com o formato manual utilizado pelo LABECA, apresenta um desempenho interessante para o desenvolvimento de uma aplicação com essa linha de pesquisa.

Comparando o hardware utilizado nesse experimento em relação aos equipamentos descritos na Seção 6.2, pode-se destacar que apesar do número menor de eletrodos, na média, e mesmo com uma taxa de aquisição significantemente menor, foi possível desenvolver a pesquisa. A partir do desenvolvimento da integração de pesquisa com o dispositivo iCelera iBlue 52 (RODRIGUES; SILVERIO, 2014), foram sugeridas, ao fornecedor do equipamento, modificações de projeto para a melhoria da integração como, por exemplo, o uso de protocolos mais atuais e uso de marcadores de tempo nos pacotes de medidas de EEG, para que fosse possível a identificação de efeitos de atraso ou avanço nos pacotes de medidas que pudessem melhorar as taxas de classificação.

De todos os experimentos apresentados na Seção 6.3, foi desenvolvido nesse trabalho um experimento mais complexo, tanto no volume de variações das imagens, quanto no volume

total das imagens. O uso e comparação de múltiplos bancos de dados foi apenas proposto por (WANG et al., 2009) e mesmo assim, não aplicado aos diferentes grupos, níveis de percepção sobre o conteúdo das imagens (especialistas e não especialistas). Khosla et al. (2011), descreveu a possibilidade de problemas relacionados aos diferentes níveis de cor e variações de brilho nas imagens. Este problema não foi estudado nos experimentos.

Na Seção 6.6, o desenvolvimento da pesquisa apresentou um número total de voluntários acima da média dos experimentos apresentados, exceto o desenvolvimento feito por (UŠĆUMLIĆ; CHAVARRIAGA; MILLÁN, 2013) que utilizou 15 voluntários para a pesquisa. Porém, nenhuma das pesquisas desenvolveu o volume de aquisições feitas neste trabalho, uma vez que os especialistas efetuaram duas aquisições com as duas bases de dados de imagens (Caltech 101 e Nausitoo) e um dos voluntários desenvolveu ainda um ciclo de aquisição adicional para os dois testes de generalização do classificador, descritos na Seção 8.2.4.

Todos os resultados estão de acordo com a descrição feita por Luck (2014), sobre a aquisição dos indicadores de ERP, conforme o exemplo, apresentado na Figura 21. A média da tensão dos eventos alvo, comparada com a média de tensão de eventos não–alvo, apresentam as diferenças características dos sinais de ERP.

Não foi possível identificar um padrão claro de aplicação de processamento e pré-processamento entre os artigos utilizados. Mas o uso de redução de dimensionalidade, aliado ao uso de regressores lineares, dentro dos artigos apresentados na Seção 6.7, parece uma tendência. Apenas Mohedano et al. (2015) apresenta a aplicação de redes neurais profundas, o que pode caracterizar uma alteração de tendência, por ser um dos artigos mais recentes. Nesse sentido, este trabalho apresenta um caminho alternativo, com a aplicação das wavelets e a adoção de um método que reduz os valores dos índices dos coeficientes, conforme a regressão de *ridge*, amplamente utilizada em wavelets.

As transformadas de wavelet, aplicadas neste estudo, não foram observadas em Lees et al. (2018), apesar de ser um método de geração de características, aliado a uma regressão de aprendizado de máquina, com níveis superiores de precisão e $F1$. Este nível de resultado obtido com wavelet "daubechies" de ordem "32", com precisão média de 97% (e pontuação média de $F1$ de 80%), comparado aos resultados em Lees et al. (2018), que varia de 60% à 91%, apresenta um algoritmo alternativo com resultados de classificação mais altos. Deve-se notar que em Lees et al. (2018), um único estudo (WEIDEN; KHOSLA; KEEGAN, 2012) apresentou um nível de precisão de 97%, mas é baseado em clipes de imagens em movimento e pode ser considerado um tipo diferente de teste de RSVP.

Cabe ressaltar que, resultados de aprendizado de máquina, em aplicações desbalanceadas onde o número de imagens alvo é muito menor que o total de imagens não–alvo, devem ser relatados com um maior volume de detalhes, uma vez que, caso o classificador erre todas as imagens alvo e classifique corretamente todas as imagens não–alvo, o resultado final apresenta um índice de precisão consideravelmente alto. Nas pesquisas observadas no Capítulo 6, poucas são as informações e detalhes sobre o desbalanceamento e a forma como isso ocorre.

Nesses casos, uma orientação seria, não apenas reportar os resultados de acertos positivos, precisão ou curvas ROC, mas também informações sobre os outros dados da matriz de confusão ou métricas que consideram todos os seus campos. Lees et al. (2018) apresenta uma tabela de comparação de vários trabalhos sobre RSVP e a aplicação de aprendizado de máquina, similar a Bechelli (2016a), com vários algoritmos e variações diferentes. Os únicos campos apresentados são ROC e precisão (LEES et al., 2018) e estes, claramente, não estão bem documentados para comparar os melhores métodos. As informações sobre os dados balanceados ou desbalanceados devem ser um requisito, como uma camada adicional de comparação, aos trabalhos apresentados e também considerados para determinar a aplicação de métricas adicionais. Curvas ROC e AUC são uma medidas totalmente fora de escala para comparar conjuntos de dados altamente desbalanceados e não atingem o nível de detalhe necessário.

Um exemplo do que foi tratado no parágrafo anterior pode ser observado na Seção 8.2.2. Na Figura 26 é apresentado um conjunto completo de curvas estatística dos resultados do classificador ($F1$, recall, acurácia e precisão), onde o resultado de AUC apresenta uma taxa de $0.999 \pm 1.092x10^{-4}$ e mesmo assim é possível avaliar os resultados incorretos do classificador.

A aplicação das 106 famílias de wavelets, desenvolvida neste trabalho, propõe um estudo mais extensivo sobre os sinais de EEG e os possíveis desdobramentos com a geração de características via wavelets. Os resultados encontrados apresentam altas taxas de classificação, utilizando como referência os resultados avaliados em Lees et al. (2018) para a classificação individual dos eventos de ERP. Os resultados estatísticos obtidos em todas as etapas, mostrados na Seção 8.2.1, apresentam bons resultados para a seleção da família de wavelets.

Na Seção 8.2.2, o resultado apresentado na Figura 26, demostra que existem resultados muito próximos de acurácia e precisão para todos os voluntários, independente do grupo e que a diferença ocorre no *recall*, que influencia diretamente os resultados do score de $F1$.

O teste de χ^2 positivo, permite considerar a existência de associação entre os resultados positivos dos especialistas e não especialistas (H_1), e demonstra que os acertos entre os grupos tem forte relação com a especialização do voluntário, o que demostra um resultado esperado no início dos testes, uma vez que o grupo de especialistas apresenta familiaridade com o uso e processamento de imagens.

Na Seção 8.2.3, os resultados positivos do teste de χ^2 também permitem inferir uma associação entre os especialistas e os diferentes bancos de dados testados. No caso, o banco de imagens Nausitoo apresentou resultados inferiores aos resultados dos especialistas com os testes no banco Caltech 101, sendo que, apenas o "Voluntário10" obteve um resultado superior com a base Nausitoo. O que é contrário ao resultado esperado, uma vez que as imagens do banco Nausitoo, apesar de uma maior complexidade de informação, deveriam obter resultados superiores, por representarem as imagens de trabalho dos voluntários.

O teste de generalização do classificador, onde ocorre o treinamento da ML com os dados obtidos em um teste e testado com os dados de um segundo teste, apresenta taxas de acerto e erro em um cenário real de aplicação dos conceitos desenvolvidos no trabalho. Cabe ressaltar

que, mesmo que executado apenas com um dos voluntários especialistas, o teste apresenta uma etapa adicional em relação aos trabalhos apresentados no Capítulo 6 e com resultados significativos, quando comparado com os resultados obtidos com o treinamento e teste desenvolvido com os mesmos dados de aquisição de EEG.

10 CONCLUSÃO

O trabalho desenvolvido, contribuiu com outras pesquisas através da integração do equipamento com software de pesquisa em biosinais e com uma nova abordagem para o tratamento de sinais de EEG aplicados a classificação de bancos de dados de imagens, através da proposta de um processo que parte do teste de aquisição dos sinais de EEG, com a aplicação destes sinais em um sistema de ML e ainda apresenta uma proposta de modelo estatístico para a avaliação dos resultados.

Levando em consideração o trabalho desenvolvido, pode-se observar a aplicação de inteligência artificial aliada aos sinais cerebrais para a classificação de diferentes bancos de dados de imagens com uma taxa de sucesso acima dos resultados obtidos em outros trabalhos.

Os resultados foram tratados através de uma modelagem estatística que confirmam a análise desenvolvida. O uso da família de wavelets, se mostrou adequado com o tratamento estatístico e com a iniciativa de se manter grandes conjuntos de dados para a análise dos diferentes níveis de coeficientes e decomposição do sinal.

Através dos experimentos desenvolvidos, foi possível traçar uma relação entre o comportamento dos sinais cerebrais e os resultados apresentados em diferentes conjuntos de voluntários (especialistas e não especialistas) com uma alta relevância estatística.

Os experimentos comprovaram a existência de uma relação entre os sinais cerebrais e os resultados apresentados para os diferentes bancos de dados de imagens para o grupo de especialistas.

Foi possível desenvolver um modelo de teste e processamento de sinais EEG para complementar os métodos utilizados em sistemas de recuperação de imagens baseada em conteúdo para a área de arqueologia visual.

Através das informações capturadas via comportamento dos sinais cerebrais seria possível melhorar o desempenho (trabalhando descontinuidades) de um sistema de busca de imagens

Foi possível propor um modelo de processamento dos sinais cerebrais para a aplicações em arqueologia visual e o modelo apresenta resultados que permitem avaliar o desempenho dos especialistas através da relação dos dois bancos de dados de imagens pesquisados.

Outro ponto de destaque, é a possibilidade de aumentar o desempenho do processamento e classificação de imagens, quando comparado com a classificação feita manualmente pelos pesquisadores.

11 TRABALHOS FUTUROS

Uma possível continuação para esta pesquisa, seria:

a) Análise de modelos de classificação (*ranking*) de imagens através dos vetores resultantes da regressão de *ridge*;
b) Aplicação do processamento proposto para outros tipos de biosinais, como por exemplo, eletromiograma e eletrocardiograma;
c) Utilização de outras formas de minimização de coeficientes, como por exemplo, o uso do método de *lasso*.
d) O estudo de um modelo de processamento que utilize as categorias das diversas imagens dos bancos de dados de imagens como marcadores para a classificação (*Multi-label classification*).
e) Aplicação de testes estatísticos paramétricos para a análise dos resultados.
f) Aplicação de um modelo baseado em wavelets em que seja possível utilizar a transformada inversa de wavelet e avaliar as curvas resultantes.

REFERÊNCIAS

ADDISON, Paul S. **The illustrated wavelet transform handbook: introductory theory and applications in science, engineering, medicine and finance**. 2nd. Boca Raton, FL: CRC press, 2017. p. 44–51.

ALDROVANDI, Cibele Elisa Viegas; KORMIKIARI, Maria Cristina Nicolau; HIRATA, Elaine Farias Veloso. **Estudos sobre o espaço na Antiguidade**. São Paulo: EDUSP: FAPESP, 2011. p. 217–222. Disponível em: <http://www.edusp.com.br/detlivro.asp?id=413124>. Acesso em: 1 jul. 2018.

ALLISON, Brendan Z. et al. Towards Practical Brain-Computer Interfaces: Bridging the Gap from Research to Real-World Applications. Springer Publishing Company, Incorporated, ago. 2012. Disponível em: <http://dl.acm.org/citation.cfm?id=2424025>. Acesso em: 1 jul. 2018.

BECHELLI, R. P. Content based Image Retrieval Databases Classification with Brain Event Related Potential. In: PROCEEDINGS of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (DOCTORAL CONSORTIUM 2016) - BIOSIGNALS. Roma: SCITEPRESS – Science e Technology Publications, Lda., 2016. p. 3–8.

_____. **Software iCelera - iBlue52 Python Integration (0.2.0)**. São Paulo: BKL Press, 2016. Disponível em: <https://bitbucket.org/rodrigoprior/icelera_iblue52_integration/>. Acesso em: 1 jul. 2018.

BELARDI, Aldo Artur; CARDOSO, José Roberto; SARTORI, Carlos Antonio França. Application of Haarwavelets in the method of moments to solve electrostatic problems. **COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering**, v. 23, n. 3, p. 606–612, 2004. Disponível em: <https://doi.org/10.1108/03321640410540511>. Acesso em: 1 jul. 2018.

_____. Wavelet's application in electrostatic and their computing aspects. **COMPEL - The International Journal for Computation and Mathematics in Electrical and Electronic Engineering**, v. 23, n. 4, p. 923–931, 2004. Disponível em: <https://doi.org/10.1108/03321640410553355>. Acesso em: 1 jul. 2018.

BELARDI, Aldo Artur et al. Electrical Rotating Equipment Failure Detection Using Wavelet Based Current Signature. In: v. 1.

BELLMAN, Christopher et al. Challenges in the effectiveness of image tagging using consumer-grade brain-computer interfaces. In: LECTURE Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Switzerland: Springer International Publishing, 2016. v. 9769, p. 55–64.

BLANKERTZ, Benjamin et al. Single-trial analysis and classification of ERP components–a tutorial. **NeuroImage**, v. 56, n. 2, p. 814–825, 2011.

BRUNNER, C et al. BCI Software Platforms. In: BIOLOGICAL e MEDICAL PHYSICS. TOWARDS Practical Brain-Computer Interfaces. France: Biological e Medical Physics, 2013. p. 303–331.

BUTTFIELD, Anna; FERREZ, Pierre W.; MILLÁN, José Del R. Towards a robust BCI: Error potentials and online learning. In: 2. IEEE Transactions on Neural Systems and Rehabilitation Engineering. Switzerland: Springer International Publishing, 2006. v. 14, p. 164–168.

CAMPOS, Haroldo de. **Ilíada de Homero. v. I e II.** v. 1. São Paulo: Arx, 2003. p. 77.

CAPATI, F.; BECHELLI, R. P.; CASTRO, Maria Claudia F. Hybrid SSVEP/P300 BCI Keyboard Controlled by Visual Evoked Potential. In: PROCEEDINGS of the 9th International Joint Conference on Biomedical Engineering Systems and Technologies (BIOSTEC 2016) - Volume 4: BIOSIGNALS. Roma: SCITEPRESS – Science e Technology Publications, Lda., 2016. p. 214–218.

CHEN, Shu–ching et al. A Dynamic User Concept Pattern Learning Framework for Content-Based Image Retrieval. **Systems, Man, and Cybernetics, Part C: Applications and Reviews, IEEE Transactions on**, v. 36, n. 6, p. 772–783, nov. 2006.

CHEN, Shu–ching et al. Framework for Content-Based Image Retrieval. v. 36, n. 6, p. 772–783, 2006.

COX, Ingemar J et al. The Bayesian image retrieval system, PicHunter: Theory, implementation, and psychophysical experiments. **IEEE TRANSACTIONS ON IMAGE PROCESSING**, v. 9, n. 1, p. 20–37, 2000.

CROWLEY, E. J.; ZISSERMAN, A. Of Gods and Goats: Weakly Supervised Learning of Figurative Art. In: BRITISH Machine Vision Conference. England: Oxford Publishing, 2013. Disponível em: <http://www.robots.ox.ac.uk/~vgg/publications/2013/Crowley13/>. Acesso em: 1 jul. 2018.

_____. The State of the Art: Object Retrieval in Paintings using Discriminative Regions. In: BRITISH Machine Vision Conference. England: Oxford Publishing, 2014.

DAUBECHIES, Ingrid. **Ten Lectures on Wavelets**. Philadelphia: SIAM, 1999. cap. 1, p. 1–16.

DELORME, Arnaud et al. EEGLAB, SIFT, NFT, BCILAB, and ERICA: New tools for advanced EEG processing. **Computational Intelligence and Neuroscience**, v. 2011, 2011. Disponível em: <http://dx.doi.org/10.1155/2011/130714>. Acesso em: 1 jul. 2018.

DESERNO, Thomas M; ANTANI, Sameer; LONG, Rodney. Ontology of gaps in content-based image retrieval. **Journal of digital imaging**, v. 22, n. 2, p. 202–15, abr. 2009. Disponível em: <http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3043678/>. Acesso em: 1 jul. 2018.

DICKTER, Cheryl L.; KIEFFABER, Paul D. **EEG methods for the psychological sciences**. London: SAGE Publications, 2014. p. 9–24.

DONCHIN, Emanuel; ARBEL, Yael. P300 based brain computer interfaces: A progress report. In: LECTURE Notes in Computer Science (including subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics). Berlin: Springer-Verlag, 2009. 5638 LNAI, p. 724–731.

DONCHIN, Emanuel; RITTER, W.; MCCALLUM, C. **Cognitive psychophysiology: The endogenous components of the ERP**. Berlin: Springer-Verlag, 1978. p. 349–441. Disponível em: <http://scholarcommons.usf.edu/psy_facpub/224>. Acesso em: 1 jul. 2018.

DONOHO, David L; JOHNSTONE, Jain M. Ideal spatial adaptation by wavelet shrinkage. **Biometrika**, Oxford University Press, v. 81, n. 3, p. 425–455, 1994.

DRONGELEN, Wim Van. **Signal Processing for Neuroscientists: Introduction to the Analysis of Physiological Signals**. Chicago: Elsevier, 2007. p. 1–33.

DUCHOWSKI, Andrew T. A breadth-first survey of eye-tracking applications. **Behavior Research Methods, Instruments**, v. 34, n. 4, p. 455–470, 2002. Disponível em: <http://link.springer.com/article/10.3758/BF03195475>. Acesso em: 1 jul. 2018.

DUFFY, Frank H.; IYER, Vasudeva G.; SURWILLO, Walter W. **Clinical Electroencephalography and Topographic Brain Mapping**. New York: Springer New York, 1989. p. 68–76. Disponível em: <https://doi.org/10.1007/978-1-4613-8826-5>. Acesso em: 1 jul. 2018.

EAKINS, John P. Towards intelligent image retrieval. **Pattern Recognition**, v. 35, p. 3–14, 2002.

EAKINS, John P.; GRAHAM, Margaret. Content-based Image Retrieval. **October**, v. 3, p. 1–59, 1999. Disponível em: <http://www.inf.fu-berlin.de/lehre/WS00/webIS/reader/WebIR/imageRetrievalOverview.pdf>. Acesso em: 1 jul. 2018.

FARWELL, L; DONCHIN, Emanuel. Talking Off the Top of Your Head. **electroencephalography and clinical Neurophysiology**, v. 70, n. 6, p. 510–523, 1988.

FEI-FEI, Li; FERGUS, Rob; PERONA, Pietro. Learning generative visual models from few training examples: An incremental Bayesian approach tested on 101 object categories. **Computer Vision and Image Understanding**, v. 106, n. 1, p. 59–70, 2007.

FINGELKURTS, Andrew A.; FINGELKURTS, Alexander A.; NEVES, Carlos F.H. **Natural world physical, brain operational, and mind phenomenal space-time**. v. 7. Finland: Elsevier, 2010. p. 195–249.

FLORENZANO, Maria Beatriz Borba. **3 o . RELATÓRIO CIENTÍFICO - Projeto temático: A Cidade e seu território na Grécia Antiga. Organização do espaço e sociedade**. São Paulo, 2009.

_____. **4 o . RELATÓRIO CIENTÍFICO - Projeto temático: A Cidade e seu território na Grécia Antiga. Organização do espaço e sociedade**. São Paulo, 2010.

_____. **CNPQ - Projeto de Pesquisa: A cidade na Grécia antiga: desenhando a sociedade no terreno (03/14 - 02/17)**. São Paulo, 2014.

FU, Hong; CHI, Zheru; FENG, Dagan. Attention-driven image interpretation with application to image retrieval. **Pattern Recognition**, v. 39, n. 9, p. 1604–1621, 2006.

GERSON, Adam D.; PARRA, Lucas C.; SAJDA, Paul. Cortically coupled computer vision for rapid image search. In: 2. IEEE Transactions on Neural Systems and Rehabilitation Engineering. New York: IEEE, 2006. v. 14, p. 174–179.

GRAMFORT, Alexandre et al. MEG and EEG data analysis with MNE-Python. **Frontiers in Neuroscience**, 7 DEC, 2013.

GRAPS, Amara. Introduction to wavelets. **IEEE Computational Science & Engineering**, v. 2, n. 2, p. 50–61, 1995. Disponível em: <https://www.eecis.udel.edu/~amer/CISC651/IEEEwavelet.pdf>. Acesso em: 1 jul. 2018.

GUDIVADA, V.N.; RAGHAVAN, V.V. Content based image retrieval systems. **Computer**, v. 28, n. 9, p. 18–22, 1995.

HASTIE, T; TIBSHIRANI, R; FRIEDMAN, J. **The elements of statistical learning**. California: Springer, 2009. p. 1–8. Disponível em: <http://link.springer.com/content/pdf/10.1007/978-0-387-84858-7.pdf>. Acesso em: 1 jul. 2018.

HASTIE, Trevor; TIBSHIRANI, Robert; WAINWRIGHT, Martin. **Statistical learning with sparsity: the lasso and generalizations**. California: CRC press, 2015.

HEALY, Graham; GURRIN, Cathal; SMEATON, Alan F. Informed Perspectives on Human Annotation Using Neural Signals. In: TIAN, Qi et al. (Ed.). **MultiMedia Modeling**. Cham: Springer International Publishing, 2016. v. 9517. (Lecture Notes in Computer Science). p. 315–327. Disponível em: <http://link.springer.com/10.1007/978-3-319-27674-8>. Acesso em: 1 jul. 2018.

HEALY, Graham; SMEATON, Alan F. Optimising the number of channels in EEG-augmented image search. **Proceedings of the 25th BCS Conference**, p. 1–6, 2011. Disponível em: <http://dl.acm.org/citation.cfm?id=2305346>. Acesso em: 1 jul. 2018.

JAMES, Gareth et al. **An Introduction to Statistical Learning**. California: Springer, 2013. v. 112, p. 215–219.

JENSEN, Arne; COUR-HARBO, Anders la. **Ripples in mathematics: the discrete wavelet transform**. Berlin: Springer Science & Business Media, 2001. p. 9–56.

KATO, Toshikazu. Database architecture for content-based image retrieval. Edição: Albert A. Jamberdino e Carlton W. Niblack. **Proc. SPIE/IS&T 1992**, International Society for Optics e Photonics, v. 1662, p. 112–123, abr. 1992. Disponível em: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=987357>. Acesso em: 1 jul. 2018.

KHANDPUR, R S. **Biomedical Instrumentation: Technology and Applications**. New Delhi: Tata McGraw-Hill Publishing Company Limited, 2003.

KHOSLA, Deepak et al. Optimal Detection of Objects in Images and Videos Using Electroencephalography (EEG) Deepak. International Society for Optics e Photonics, Malibu, p. 80501c–80501c–11, maio 2011. Disponível em: <http://proceedings.spiedigitallibrary.org/proceeding.aspx?articleid=1270822>. Acesso em: 1 jul. 2018.

KLEM, George H. et al. The ten twenty electrode system of the international federation. **Electroencephalography and clinical neurophysiology**, v. 10, n. 2, p. 370–375, 1958. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/10590970>. Acesso em: 1 jul. 2018.

LE, Quoc V. et al. Building high-level features using large scale unsupervised learning, dez. 2011. arXiv: 1112.6209. Disponível em: <http://arxiv.org/abs/1112.6209>. Acesso em: 1 jul. 2018.

LEES, Stephanie et al. A review of rapid serial visual presentation-based brain-computer interfaces. **Journal of Neural Engineering**, v. 15, n. 2, p. 021001, 2018. Disponível em: <http://iopscience.iop.org/article/10.1088/1741-2552/aa9817>. Acesso em: 1 jul. 2018.

LI, Rui et al. Human-centric approaches to image understanding and retrieval. In: 2010 Western New York Image Processing Workshop. 2010 Western N, US: IEEE, nov. 2010. p. 62–65.

LIANG, Zhen et al. Content-based image retrieval using a combination of visual features and eye tracking data. **Symposium on Eye-Tracking.**, v. 1, n. 212, p. 41–44, 2010. Disponível em: <http://dl.acm.org/citation.cfm?id=1743675>. Acesso em: 1 jul. 2018.

LIU, Ying et al. A survey of content-based image retrieval with high-level semantics. **Pattern Recognition**, v. 40, p. 262–282, 2007.

LOWE, D.G. Object recognition from local scale-invariant features. **Proceedings of the Seventh IEEE International Conference on Computer Vision**, v. 2, 1999. Disponível em: <https://www.cs.ubc.ca/~lowe/papers/iccv99.pdf>.

LUCK, Steven J. **An Introduction to the Event-Related Potential Technique**. Edição: A Bradford Book. Second Edi. California: MIT Press, 2014. p. 2–100. Disponível em: <https://mitpress.mit.edu/books/introduction-event-related-potential-technique>. Acesso em: 1 jul. 2018.

MAIORANA, F. A medical content based image retrieval system with eye tracking relevance feedback. In: PROCEEDINGS of the 26th IEEE International Symposium on Computer-Based Medical Systems. Porto, Portugal: IEEE, jun. 2013. p. 297–302.

MEYER, Yves. **Algorithms and applications**. Philadelphia: SIAM, 1993. cap. 2, p. 30.

MOHEDANO, Eva et al. Exploring EEG for Object Detection and Retrieval. In: PROCEEDINGS of the 5th ACM on International Conference on Multimedia Retrieval. New York, NY, USA: ACM, 2015. (ICMR '15), p. 591–594. Disponível em: <http://doi.acm.org/10.1145/2671188.2749368>. Acesso em: 1 jul. 2018.

MOHEDANO, Eva et al. Object Segmentation in Images using EEG Signals. In: PROCEEDINGS of the ACM International Conference on Multimedia - MM '14. New York, New York, USA: ACM Press, ago. 2014. p. 417–426. arXiv: 1408.4363. Disponível em: <http://arxiv.org/abs/1408.4363>. Acesso em: 1 jul. 2018.

MORAUD, Eduardo Martin. Wavelet networks. **A report**, 2009.

MORETIN, PA. **Ondas e Ondaletas: da Análise de Fouriera Análise de Ondaletas**. São Paulo: Edusp, 1999. p. 10–31.

MORIMOTO, Carlos H.; MIMICA, Marcio R.M. Eye gaze tracking techniques for interactive applications. **Computer Vision and Image Understanding**, v. 98, n. 1, p. 4–24, abr. 2005. Disponível em: <http://linkinghub.elsevier.com/retrieve/pii/S1077314204001109>. Acesso em: 1 jul. 2018.

MÜLLER, Henning et al. A review of content-based image retrieval systems in medical applications-clinical benefits and future directions. **International journal of medical informatics**, v. 73, n. 1, p. 1–23, fev. 2004. Disponível em: <http://www.ncbi.nlm.nih.gov/pubmed/15036075>. Acesso em: 1 jul. 2018.

MURPHY, KP. Machine Learning: a Probabilistic Perspective. **MIT Press, Cambridge, MA**, 2012.

NATIONAL CANCER INSTITUTE. **Annotated lobes of the cerebrum**. Reuse of Graphics and Text Policy: https://www.cancer.gov/policies/copyright-reuse. 2013. Disponível em: <https://commons.wikimedia.org/wiki/File:Cerebrum_lobes.svg#file>. Acesso em: 4 jan. 2018.

PEDREGOSA, F. et al. Scikit-learn: Machine Learning in Python. **Journal of Machine Learning Research**, v. 12, p. 2825–2830, 2011.

PINKER, Steven. Como a mente funciona (Tradução Laura Teixeira Motta). **São Paulo: Companhia das Letras**, 1998.

RENARD, Yann et al. **OpenViBE: An Open-Source Software Platform to Design, Test, and Use Brain–Computer Interfaces in Real and Virtual Environments**. v. 19. France: INRIA, 2010. p. 35–53.

REZENDE, R. H.; BECHELLI, R. P. Acervos digitais e a socialização do conhecimento arqueológico: o caso do Labeca. In: XX Congresso da SBEC Público e Privado na Antiguidade. Mariana: SBEC, 2015.

RODRIGUES, Rafael; SILVERIO, Amaral Andrea Paula. **Manual de Utilização Software iBlue 52**. São Paulo: iCelera, 2014. Rev.: 1.3 12/14.

ROSSUM, Guido van. Python tutorial, May 1995. **CWI Report CS-R9526**, CS-R9526, p. 1–65, 1995. Disponível em: <http://oai.cwi.nl/oai/asset/5007/05007D.pdf>. Acesso em: 1 jul. 2018.

RUI, Yong; HUANG, Thomas S. Relevance feedback: a power tool for interactive content-based image retrieval. **Circuits and Systems**, v. 8, n. 5, 1998.

RUSSELL, Stuart; NORVIG, Peter. **Artificial Intelligence: A Modern Approach**. 3rd. Upper Saddle River, NJ, USA: Prentice Hall Press, 2009. Disponível em: <http://dl.acm.org/citation.cfm?id=1671238>. Acesso em: 1 jul. 2018.

SANEI, Saeid; CHAMBERS, Jonathon. A. **EEG Signal Processing**. Chichester: John Wiley & Sons Ltd, 2007. v. 1, p. 1–31. Disponível em: <https://onlinelibrary.wiley.com/doi/book/10.1002/9780470511923>. Acesso em: 1 jul. 2018.

SCHULZE, Christian; FRISTER, Robby; SHAFAIT, Faisal. Eye-tracker based part-image selection for image retrieval. In: 2013 IEEE International Conference on Image Processing. Melbourne, VIC, Australia: IEEE, set. 2013. p. 4392–4396.

SMEULDERS, AWM; WORRING, Marcel. Content-based image retrieval at the end of the early years. **Pattern Analysis**, v. 22, n. 12, p. 1349–1380, 2000.

SMITH, JR; CHANG, SF. VisualSEEk: a fully automated content-based image query system. **Proceedings of the fourth ACM international**, 1997. Disponível em: <http://dl.acm.org/citation.cfm?id=244151>. Acesso em: 1 jul. 2018.

SUTTON, S et al. Evoked-potential correlates of stimulus uncertainty. **Science (New York, N.Y.)**, v. 150, n. 700, p. 1187–1188, 1965.

TORRES, Ricardo da Silva; FALCÃO, Alexandre Xavier. Content-Based Image Retrieval: Theory and Applications. In: 2. v. 13, p. 161–185. Disponível em: <http://www.ic.unicamp.br/~rtorres/mo805R_13s1/2006-Torres2006RITA.pdf>. Acesso em: 1 jul. 2018.

TUKEY, John W. Comparing Individual Means in the Analysis of Variance. **Biometrics**, [Wiley, International Biometric Society], v. 5, n. 2, p. 99–114, 1949. Disponível em: <http://www.jstor.org/stable/3001913>. Acesso em: 1 jul. 2018.

UŠĆUMLIĆ, Marija; CHAVARRIAGA, Ricardo; MILLÁN, José Del R. An iterative framework for EEG-based image search: robust retrieval with weak classifiers. Edição: Michael J. Proulx. **PloS one**, Public Library of Science, v. 8, n. 8, e72018, jan. 2013. Disponível em: <http://journals.plos.org/plosone/article?id=10.1371/journal.pone.0072018>. Acesso em: 1 jul. 2018.

_____. On coupling Computer Vision and Single trial detection of ERP under RSVP. **International Journal of Bioelectromagnetism**, v. 13, n. 3, p. 133–135, 2011.

VALENS, C. **A Really Friendly Guide to Wavelets**. California: UNM Computer Science, 1999.

VAN FLEET, Patrick J. **Discrete wavelet transformations: An elementary approach with applications**. New Jersey: John Wiley & Sons, 2011. p. 1–5.

VIDAL, J J. Toward direct brain-computer communication. en. **Annual review of biophysics and bioengineering**, Annual Reviews 4139 El Camino Way, P.O. Box 10139, Palo Alto, CA 94303-0139, USA, v. 2, p. 157–80, jan. 1973. Disponível em: <http://www.annualreviews.org/doi/abs/10.1146/annurev.bb.02.060173.001105?journalCode=biophys.1>. Acesso em: 1 jul. 2018.

WALKER, James S. **A primer on wavelets and their scientific applications**. London: CRC press, 2008. p. 16–46.

WANG, Jun et al. Brain State Decoding for Rapid Image Retrieval. In: PROCEEDINGS of the seventeen ACM international conference on Multimedia. California: ACM, 2009. p. 945–954. Disponível em: <https://dl.acm.org/citation.cfm?id=1631463>. Acesso em: 1 jul. 2018.

WASILEWSKI, Filip. **PyWavelets – Wavelet Transforms in Python**. 2006. Disponível em: <https://github.com/PyWavelets/pywt>. Acesso em: 1 jul. 2018.

WEIDEN, Matthew; KHOSLA, Deepak; KEEGAN, Matthew. Electroencephalographic Detection of Visual Saliency of Motion Towards a Practical Brain-computer Interface for Video Analysis. In: PROCEEDINGS of the 14th ACM International Conference on Multimodal Interaction. Santa Monica, California, USA: ACM, 2012. (ICMI '12), p. 601–606. Disponível em: <http://doi.acm.org/10.1145/2388676.2388800>. Acesso em: 1 jul. 2018.

ZHAO, Yihong; OGDEN, R. Todd; REISS, Philip T. Wavelet-Based LASSO in Functional Linear Regression. **Journal of Computational and Graphical Statistics**, Taylor & Francis, v. 21, n. 3, p. 600–617, 2012. PMID: 23794794. eprint: https://doi.org/10.1080/10618600.2012.679241. Disponível em: <https://doi.org/10.1080/10618600.2012.679241>. Acesso em: 1 jul. 2018.

APÊNDICE A – ICELERA IBLUE52 PYTHON INTEGRATION

Durante o desenvolvimento do projeto, ocorreu a possibilidade de aquisição de um equipamento de EEG mais robusto para o desenvolvimento da pesquisa. A empresa iCelera, fornecedora do equipamento iBlue52, ofereceu a possibilidade de que efetuar a integração direta com o equipamento, sem a necessidade de utilização do software médico de aquisição e processamento dos dados comercializado junto ao equipamento.

O equipamento foi adquirido através do "Edital CAPES Pró-Equipamentos IES Comunitárias (051/2012), Subprojeto – EEG em Interface Cérebro-Computador", coordenado pelo Prof. Dr. Plinio Thomaz Aquino Junior.

A integração foi desenvolvida com a linguagem de programação Python e foi disponibilizada para que outros pesquisadores pudessem utilizar a integração em seus estudos e aquisições de eletroencefalograma, eletromiograma e eletrocardiograma além das interfaces de sinal DC e botões disponíveis no equipamento.

Como parte do processamento da conversão analógico-digital do equipamento é feito na camada de software de integração, foi efetuado uma calibração do equipamento junto ao fornecedor iCelera e encontra-se no Anexo B.

O software encontra-se online [1] junto com a documentação e também foi desenvolvido um estudo de caso de integração para que fosse utilizado em treinamentos, palestras e aulas. Este tutorial foi apresentado no "Curso de Verão - 2017" da FEI e também encontra-se online [2].

[1] `https://bitbucket.org/rodrigoprior/icelera_iblue52_integration/`
[2] `https://bitbucket.org/rodrigoprior/2017-fei-summercamp/`

APÊNDICE B – TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

CENTRO UNIVERSITÁRIO DA FEI

TERMO DE CONSENTIMENTO LIVRE E ESCLARECIDO

Você está sendo convidado a participar de um estudo denominado: "Arquitetura adaptativa para plataforma de processamento de acervos museológicos", cujo objetivo é o desenvolvimento de uma Interface Cérebro Computador (ICC) que constitui de um sistema de Identificação de Imagens baseado na análise de sinais de Eletroencefalograma (EEG).

I - DADOS DE IDENTIFICAÇÃO DO VOLUNTÁRIO DA PESQUISA

NOME DO VOLUNTÁRIO: _____

DOCUMENTO DE IDENTIDADE Nº:_____

SEXO: M F

DATA NASCIMENTO (dd/mm/aaaa): _____

ENDEREÇO:_____ Nº:_____ Comp:_____

BAIRRO:_____ CIDADE:_____

CEP:_____ TELEFONE: DDD (_____) _____

II - DADOS SOBRE A PESQUISA CIENTÍFICA

1. TÍTULO DO PROTOCOLO DE PESQUISA:

Arquitetura adaptativa para plataforma de processamento de acervos museológicos

2. PESQUISADOR:

ALUNO DE DOUTORADO (RG): Rodrigo Prior Bechelli (RG. 208871007)

CARGO/FUNÇÃO: PESQUISADOR

ORIENTADOR DA PESQUISA: Profa. Dra. Maria Claudia Ferrari de Castro (CPF 104.951.588-95)

CARGO/NÍVEL/DEPTO. : Profa. Tempo Integral, Adjunto I, Depto de Engenharia Elétrica

LABORATÓRIO: Laboratório de Engenharia Biomédica

CO-ORIENTADOR: Prof. Dr.. Plinio Thomáz Aquino Júnior (CPF 162.153.428-66)

CARGO/NÍVEL/DEPTO. : Prof. Tempo Integral, Adjunto I, Depto de Ciência da Computação

LABORATÓRIO: Laboratório de Usabilidade do 3iGROUP (Interaction and Intelligent Interfaces Group)

3. AVALIAÇÃO DO RISCO DA PESQUISA:

Tendo em vista a pré-seleção dos voluntários, considera-se que os riscos desse estudo são mínimos não havendo nenhuma evidência específica de que o participante sofrerá algum dano como consequência imediata ou tardia do estudo, devido à baixa frequência de apresentação das imagens (menor do que 10Hz). Contudo, em função do esforço visual constante, realizado durante o experimento, poderá haver um desconforto devido a um leve cansaço nos olhos. Afim de minimizar este cansaço, durante o experimento são oferecidos vários períodos de descanso.

De qualquer forma, em casos de sensibilidade visual não diagnosticada ou não informada anteriormente, podendo acarretar na ocorrência de uma convulsão, a observação de padrões

alterados de EEG, durante o procedimento, resultará na sua imediata interrupção e no encaminhamento do voluntário ao corpo clínico ambulatorial da universidade.

4.DURAÇÃO DA PESQUISA:

Prazo da pesquisa é de 20 meses.

III - REGISTRO DAS EXPLICAÇÕES DO PESQUISADOR AO VOLUNTÁRIO:

1. Justificativa e os objetivos da pesquisa:

O estudo em que você está sendo convidado a participar tem como objetivo desenvolver uma **Interface Cérebro Computador (ICC) que constitui de um sistema de Identificação e reconhecimento de Imagens baseado na análise de sinais de Eletroencefalograma (EEG). Ou seja, propor um sistema para ampliar a forma de indexar e extrair as informações armazenadas em arquivos que não contem textos escritos utilizando sinais cerebrais produzidos durante a visualização dessas imagens.**

2. Procedimentos que serão utilizados e propósitos, incluindo a identificação dos procedimentos que são experimentais:

O procedimento de teste será realizado em quatro sessões em um único dia com duração total de 1 hora para a realização de todo o experimento.

Na primeira sessão, você será esclarecido sobre o protocolo experimental e colocado em uma cadeira confortável sendo instruído a olhar fixamente no monitor de fundo preto. Após 4 segundos uma sequência aleatória de imagens será apresentada com frequência de 4Hz (4 imagens por segundo). Entre as imagens serão inseridas um subconjunto de imagens alvo para que você possa identificá-las. A categoria da imagem alvo será apresentada no início da sessão. Após a primeira sessão será oferecido um intervalo de 5 minutos para descanso.

O procedimento para a segunda, terceira e quarta sessões é igual ao da primeira, sendo oferecido o mesmo tempo de descanso de 5 minutos entre as sessões, porém será alterado apenas a sequência das imagens, totalizando para cada sessão aproximadamente 10 minutos de teste.

A distinção entre cada uma das sessões é a aleatoriedade das posições entre as imagens que são alvo e a inserção de novas imagens que não são alvo. Todas as imagens possuem o mesmo tamanho e são apresentadas no centro de uma tela preta com resolução adequada.

3. Desconfortos e riscos esperados:

A sua participação nesta pesquisa é voluntária e a avaliação oferece risco e desconforto mínimos. Em função do esforço visual constante, realizado durante o experimento, poderá haver um desconforto

devido a um leve cansaço nos olhos. Contudo, serão oferecidos vários períodos de descanso, afim de minimizar este cansaço. Caso você tenha uma sensibilidade visual não diagnosticada ou não informada anteriormente, podendo acarretar na ocorrência de uma convulsão, em caso de observarmos padrões alterados de EEG, durante o procedimento mesmo será imediatamente interrompido e você será encaminhado ao corpo clínico ambulatorial da universidade.

4. Benefícios que poderão ser obtidos:
A sua participação não lhe trará nenhum benefício direto, mas permitirá um melhor conhecimento a respeito da função cerebral relacionada a estímulos visuais (potencial evocado) e do desenvolvimento de sistemas de reconhecimento e classificação de padrões e de controle de Interfaces Humano Máquina.

IV - ESCLARECIMENTOS DADOS PELO PESQUISADOR SOBRE GARANTIAS DO VOLUNTÁRIO DA PESQUISA CONSIGNANDO:

1. Acesso, a qualquer tempo, às informações sobre procedimentos, riscos e benefícios relacionados à pesquisa, inclusive para dirimir eventuais dúvidas: **Informo que o Sr(a) tem a garantia de acesso, em qualquer etapa do estudo, a qualquer esclarecimento de eventuais dúvidas, bem como o acesso aos resultados parciais e totais da pesquisa.**

2. Liberdade de retirar seu consentimento a qualquer momento e de deixar de participar do estudo: **Também é garantida a liberdade da retirada de seu consentimento, a qualquer momento, deixando de participar do estudo.**

3. Salvaguarda da confidencialidade, sigilo e privacidade: **A equipe se compromete a utilizar os dados coletados somente para pesquisa e os resultados serão veiculados através de artigos científicos em revistas especializadas e/ou em encontros científicos e congressos, sem nunca tornar possível a sua identificação, garantindo o sigilo de sua participação.**

4. Disponibilidade de assistência, por eventuais danos à saúde, decorrentes da pesquisa: **O presente estudo oferece risco mínimo à saúde, e os pesquisadores se disponibilizam a dar assistência por eventuais intercorrências.**

5. Viabilidade de indenização por eventuais danos à saúde decorrentes da pesquisa: **Sim, através de recursos próprios dos pesquisadores.**

V. INFORMAÇÕES DE NOMES, ENDEREÇOS E TELEFONES DOS RESPONSÁVEIS PELO ACOMPANHAMENTO DA PESQUISA, PARA CONTATO EM CASO DE INTERCORRÊNCIAS CLÍNICAS E REAÇÕES ADVERSAS.

ALUNO DE DOUTORADO (RG): Rodrigo Prior Bechelli (RG. 208871007)

CARGO/FUNÇÃO: PESQUISADOR

ORIENTADOR DA PESQUISA: Profa. Dra. Maria Claudia Ferrari de Castro (CPF 104.951.588-95)

CARGO/NÍVEL/DEPTO. : Profa. Tempo Integral, Adjunto I, Depto de Engenharia Elétrica

LABORATÓRIO: Laboratório de Engenharia Biomédica

TEL.: (11) 99186-3413

CO-ORIENTADOR: Prof. Dr.. Plinio Thomáz Aquino Júnior (CPF 162.153.428-66)

CARGO/NÍVEL/DEPTO : Prof. Tempo Integral, Adjunto I, Depto de Ciência da Computação

LABORATÓRIO: Laboratório de Usabilidade do 3iGROUP (Interaction and Intelligent Interfaces Group)

TEL.: (11) 996513607

LABORATÓRIO: Laboratório de Usabilidade do 3iGROUP (Interaction and Intelligent Interfaces Group)

Se você tiver alguma consideração ou dúvida sobre a ética da pesquisa, entre em contato com o Comitê de Ética em Pesquisa (CEP-UMESP) – Rua do Sacramento, 232 – Ed. Capa sala 321- Telefone: 4366-5814 – E-mail: cometica@metodista.br."

VI - CONSENTIMENTO PÓS-ESCLARECIDO

Declaro que, após convenientemente esclarecido pelo pesquisador e ter entendido o que me foi explicado, consinto em participar do presente Protocolo de Pesquisa. O presente documento é emitido em duas vias, sendo uma do pesquisador e outro do participante.

São Bernardo do Campo, de de 201 .

_____ _____

Voluntário da pesquisa

Maria Claudia Ferrari de Castro
CPF. 192.472.678-54

_____

Plinio Thomáz Aquino Júnior
CPF 162.153.428-66

ANEXO A – RELATÓRIO DE APROVAÇÃO DO COMITE DE ÉTICA

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: Arquitetura adaptativa para plataforma de processamento de acervos museológicos.

Pesquisador: Maria Claudia Ferrari de Castro
Área Temática:
Versão: 2
CAAE: 52925116.0.0000.5508
Instituição Proponente: FUNDACAO EDUCACIONAL INACIANA PADRE SABOIA DE MEDEIROS
Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 1.463.920

Apresentação do Projeto:
Trata-se de um procedimento experimental para indexação de arquivos e identificação de categoria de imagens baseado na análise de Potencial Evocado Visual. Durante o procedimento de RSVP (Rapid Serial Visual Presentation), no qual diferentes imagens são apresentadas, o Eletroencefalograma é adquirido para posterior análise e identificação automática do P300, evocado durante a apresentação das imagens alvo. A aplicação de modelos matemáticos possibilitará portanto a detecção de padrões específicos no EEG possibilitando a distinção entre diferentes categorias de imagens.
Para a obtenção de dados que possibilitará o desenvolvimento desta interface, serão convidados 15 participantes, com idade entre 20 e 50 anos, não portadores de deficiência física, neurológica ou motora e sem histórico de sensibilidade à estímulos visuais.
Estes participantes serão submetidos a 4 sessões de exposição de imagens aleatórias numa frequência de 4 Hz, com pausa de 5 minutos entre cada sessão. Durante a exposição serão colhidos dados de eletroencefalografia.
Os dados eletroencefalográficos serão então tratados com o intuito de encontrar padrões para imagens alvos que estarão distribuídas de forma aleatória nas sessões de imagem.
Todo este projeto será executado nas dependências da Fundação Educacional Inaciana Padre

Sabóia de Medeiros- São Bernardo do Campo- SP.

Objetivo da Pesquisa:

O objetivo é ampliar a forma de indexar e extrair as informações armazenadas em arquivos que não contém textos escritos, utilizando para tal diferentes tipos de tecnologias para ampliar a compreensão de como tratamos o conhecimento obtido por outras formas que não a descrita por texto.

Avaliação dos Riscos e Benefícios:

Segundo os pesquisadores:

"Os riscos são mínimos não havendo nenhuma evidência específica de que o participante sofrerá algum dano como consequência imediata ou tardia do estudo estimulado por frequências menores do que 15Hz. Em função do esforço visual constante, realizado durante o experimento, poderá haver um desconforto devido a um leve cansaço nos olhos. Contudo, durante o experimento são oferecidos vários períodos de descanso, afim de minimizar este cansaço."

O relator considera que todos os cuidados estão sendo tomados para minimizar os riscos. Antes mesmo do procedimento os participantes da pesquisa também passarão por um processo de seleção, com o intuito de excluir aqueles que possuírem histórico de sensibilidade a estímulos visuais.

Quanto aos benefícios, segundo os mesmos pesquisadores este não virá de maneira direta para o participante, mas, será de grande valia para o desenvolvimento de novas interfaces Humano Máquina.

Comentários e Considerações sobre a Pesquisa:

Pesquisa relevante para o momento atual.

Considerações sobre os Termos de apresentação obrigatória:

Todos os termos de apresentação obrigatória foram incluídos.

Recomendações:

Não há.

Conclusões ou Pendências e Lista de Inadequações:

Após leitura e análise do projeto e exame criterioso de todos os itens que compõem os documentos do Protocolo de Pesquisa, incluindo os itens presentes no Roteiro de Checagem para o parecerista e o Termo de Consentimento Livre e Esclarecido nada foi constatado no processo que fira os princípios e normas da ética em pesquisa.

Considerações Finais a critério do CEP:

O CEP-UMESP considera o projeto de pesquisa APROVADO, lembrando que a condição de aprovação da pesquisa propriamente dita exige o que segue:

- Que sejam encaminhados ao CEP-UMESP relatórios anuais sobre o andamento da pesquisa (parciais e finais);
- Que sejam notificados ao CEP-UMESP eventos adversos que tenham ocorrido no curso da pesquisa e que sejam significativos do ponto de vista ético e metodológico;
- Que sejam notificadas eventuais emendas e modificações no protocolo de pesquisa.

Este parecer foi elaborado baseado nos documentos abaixo relacionados:

Tipo Documento	Arquivo	Postagem	Autor	Situação
Informações Básicas do Projeto	PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_644878.pdf	15/03/2016 17:09:16		Aceito
Outros	20160315_formulario_resposta_cepumesp.pdf	15/03/2016 17:05:54	Rodrigo Prior Bechelli	Aceito
Projeto Detalhado / Brochura Investigador	projeto_comite_etica_MODIFICADO_20160315.pdf	15/03/2016 17:03:22	Rodrigo Prior Bechelli	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TermoConsentimento_MODIFICADO_20160315.pdf	15/03/2016 17:02:51	Rodrigo Prior Bechelli	Aceito
Projeto Detalhado / Brochura Investigador	projeto_comite_etica.pdf	15/03/2016 11:28:54	Rodrigo Prior Bechelli	Aceito
TCLE / Termos de Assentimento / Justificativa de Ausência	TermoConsentimento.odt	15/03/2016 11:24:40	Rodrigo Prior Bechelli	Aceito
Folha de Rosto	folha_de_rosto.pdf	28/01/2016 18:47:45	Rodrigo Prior Bechelli	Aceito
Declaração de Pesquisadores	scan_decl_responsabilidade.pdf	25/01/2016 17:14:51	Rodrigo Prior Bechelli	Aceito

Situação do Parecer:

Aprovado

Necessita Apreciação da CONEP:

Não

Continuação do Parecer: 1.463.920

SAO BERNARDO DO CAMPO, 23 de Março de 2016

Assinado por:
Dario Paulo Barrera Rivera
(Coordenador)

Endereço: DO SACRAMENTO
Bairro: RUDGE RAMOS **CEP:** 09.640-000
UF: SP **Município:** SAO BERNARDO DO CAMPO
Telefone: (11)4366-5817 **E-mail:** cometica@metodista.br

ANEXO B – CERTIFICADO DE CALIBRAÇÃO ICELERA IBLUE52

RELATÓRIO DE CALIBRAÇÃO

Número: IC520034/2016

DADOS DO CLIENTE	
Cliente:	Fundação Ed. Inaciana Pe Saboia de Medeiros
Endereço	Rua Vergueiro, 165
CNPJ:	61.023.156/0001-82
Contato	

DADOS DO EQUIPAMENTO	
Equipamento:	Eletroencefalógrafo Digital portátil
Marca	Icelera
Modelo:	iBlue52
No de Série:	VGI52160154

DADOS DA CALIBRAÇÃO	
Data da Calibração:	14/dez/2016
Próxima Calibração:	14/dez/2017
Local da Calibração:	Icelera STO-SP
Padrões Utilizados:	Simulador de EEG, Nro de Série:33193
	Certificado do Padrão Nro: 4664/2016 Válido até 01/09/2017

Procedimento de Calibração:

1.A calibração é realizada para detectar uma tendência nas faixas que abrangem a escala total de trabalho do equipamento.
2.Os pontos de calibração são definidos para atender a área útil de utilização.
3.O ensaio é realizado por simulações de sinais elétricos, sendo seus valores verdadeiramente conhecidos e aprovados.
4.A dispersão sistemática existente na indicação nominal do equipamento é herdada pelo padrão de trabalho utilizado.
5.Os valores obtidos correspondem à medida de três medições, para o nível de confiança de 95%

Terminologia
V.V.C – Valor Verdadeiro Convencional
I.T.M – Incerteza Total da Medição
V.M.M – Valor Medido Médio (média de 3 medições)

Eng.º Rafael Rodrigues Amaral
Responsável Técnico
CREA-SP:5061256152

A reprodução poderá ser comente por completo, reproduções parciais devem ser aprovadas pela Icelera
Os resultados apresentados referem-se exclusivamente ao objeto ensaiado, não sendo extensivo a quaisquer lotes.

RESULTADOS DA CALIBRAÇÃO

V.V.C (uVpp):	50,000		
Canais	**V.M.M (uVpp)**	**Erro %**	**I.T.M. (uVpp)**
FP1	50,080	0,16	0,23048
FP2	50,039	0,08	0,18926
F7	50,695	1,39	0,85686
F3	50,572	1,14	0,73037
FZ	50,121	0,24	0,27177
F4	50,162	0,32	0,31312
F8	50,408	0,82	0,56266
T3	51,105	2,21	1,28285
C3	50,080	0,16	0,23048
CZ	50,572	1,14	0,73037
C4	50,654	1,31	0,81463
T4	50,244	0,49	0,39603
T5	50,654	1,31	0,81463
P3	50,654	1,31	0,81463
PZ	50,449	0,90	0,60449
P4	50,654	1,31	0,81463
T6	50,408	0,82	0,56266
O1	49,711	-0,58	-0,13809
OZ	50,285	0,57	0,43759
O2	50,121	0,24	0,27177

A reprodução poderá ser comente por completo, reproduções parciais devem ser aprovadas pela Icelera
Os resultados apresentados referem-se exclusivamente ao objeto ensaiado, não sendo extensivo a quaisquer lotes.

RESULTADOS DA CALIBRAÇÃO

V.V.C (uVpp):	100,000		
Canais	**V.M.M (uVpp)**	**Erro %**	**I.T.M. (uVpp)**
FP1	100,078	0,08	0,37852
FP2	99,053	-0,95	-0,64065
F7	99,422	-0,58	-0,27617
F3	99,094	-0,91	-0,60029
FZ	99,832	-0,17	0,13200
F4	99,996	0,00	0,29621
F8	99,258	-0,74	-0,43850
T3	100,119	0,12	0,41972
C3	98,930	-1,07	-0,76154
CZ	99,545	-0,45	-0,15408
C4	99,709	-0,29	0,00919
T4	99,176	-0,82	-0,51946
T5	99,709	-0,29	0,00919
P3	100,201	0,20	0,50223
PZ	100,037	0,04	0,33734
P4	99,545	-0,45	-0,15408
T6	99,627	-0,37	-0,07251
O1	99,299	-0,70	-0,39797
OZ	98,725	-1,28	-0,96235
O2	98,848	-1,15	-0,84197

A reprodução poderá ser comente por completo, reproduções parciais devem ser aprovadas pela Icelera
Os resultados apresentados referem-se exclusivamente ao objeto ensaiado, não sendo extensivo a quaisquer lotes.

RESULTADOS DA CALIBRAÇÃO

V.V.C (uVpp):	500,000		
Canais	V.M.M (uVpp)	Erro %	I.T.M. (uVpp)
FP1	504,328	0,87	5,87956
FP2	498,586	-0,28	0,08685
F7	502,523	0,50	4,04441
F3	500,227	0,05	1,72888
FZ	502,236	0,45	3,75381
F4	501,785	0,36	3,29783
F8	501,867	0,37	3,38068
T3	504,082	0,82	5,62868
C3	499,570	-0,09	1,07018
CZ	501,293	0,26	2,80133
C4	503,631	0,73	5,16937
T4	501,703	0,34	3,21501
T5	501,621	0,32	3,13222
P3	502,893	0,58	4,41952
PZ	500,021	0,00	1,52216
P4	500,719	0,14	2,22329
T6	500,514	0,10	2,01717
O1	500,268	0,05	1,77005
OZ	500,186	0,04	1,68773
O2	499,734	-0,05	1,23444

A reprodução poderá ser comente por completo, reproduções parciais devem ser aprovadas pela Icelera
Os resultados apresentados referem-se exclusivamente ao objeto ensaiado, não sendo extensivo a quaisquer lotes.

Certificado de Calibração

INTERMETRO

F-015 rev. A

Nº 4664/16

Folha 01/01

Contratante: *BALITEK INSTRUMENTOS E SERVICOS LTDA*
Solicitante: *ICELERA INDÚSTRIA E COMÉRCIO DE EQUIPAMENTOS MÉDICOS LTDA-ME*
Endereço: *Rua José Manoel de Almeida, 177 Cep: 06730-000 - Vargem Grande Paulista - SP*
Item Calibrado: *SIMULAR DE ECG* **Nº de Patrimônio:** ---
Marca: *NETECH* **Modelo:** *MINISIM-EEG* **Nº de Série:** *33193*
Acessórios Conjugados: *Não Possui.* **Nº de Identificação:** ---
OSC Nº: *4882/16* **Data da Calibração:** *31/08/16*

Condições Ambientais Aplicáveis à Calibração

Temperatura durante a calibração: *(24 ± 1)°C* **Umidade Relativa do Ar:** *(51 ± 15) %ur*

Metodologia de Calibração

"Vide nota 1"

Padrões Utilizados

Padrão de Trabalho	Certificado de Calibração	Validade do Padrão
Osciloscopio	*94XRH2*	*Dez/2017*

Resultados Obtidos

Amplitude com Freqüência de 2Hz

Unidade Medida: uV

Valor de Referência	Valor Medido	Erro	Incerteza	k
10,0	9,9	0,1	0,05	2,16
100,0	98,7	1,3	0,05	2,16
500,0	501,5	-1,5	0,05	2,16

Amplitude com Freqüência de 5Hz

Unidade Medida: uV

Valor de Referência	Valor Medido	Erro	Incerteza	k
10,0	10,4	-0,4	0,05	2,16
100,0	101,3	-1,3	0,05	2,16
500,0	498,2	1,8	0,05	2,16

Notas

1. A Calibração é realizada por comparação direta com o padrão de referência e a leitura do erro , indicada no instrumento em calibração.
2. A incerteza de medição é considerada a partir de uma incerteza padrão de medição multiplicada pelo fator de abrangência de k=2,0 ;que, para uma distribuição normal, corresponde a uma probabilidade de abrangência de aproximadamente 95%, determinada em conformidade com o procedimento NIT-DICLA-021

3.Os resultados apresentados no presente documento foram extraídos do certificado de Laboratório coligado e qualificado e referem-se exclusivamente ao item de serviço supra mencionado e nas condições de calibração referenciadas. O presente certificado somente pode ser reproduzido na sua forma e conteúdo integrais e sem alterações. Não pode ser utilizado para fins promocionais.

Data de Emissão: *01/09/16*

Eng. José Stankevicius
Signatário Autorizado
Este certificado foi assinado eletronicamente pelo Signatário Autorizado.