

ESCOLA POLITÉCNICA
PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA
MESTRE

MILTON MACHADO DE SOUZA JÚNIOR

**Reconhecimento facial e fNIRS para a detecção de traços de emoção
aplicados a jogos sérios**

Porto Alegre - RS, Brasil

2018

PÓS-GRADUAÇÃO - *STRICTO SENSU*



Pontifícia Universidade Católica
do Rio Grande do Sul

MILTON MACHADO DE SOUZA JÚNIOR

Reconhecimento facial e fNIRS para a detecção de traços de emoção aplicados a jogos sérios

Dissertação apresentada como requisito para a obtenção do grau de Mestre pelo Programa de Pós-Graduação em Engenharia Elétrica da Escola Politécnica da Pontifícia Universidade Católica do Rio Grande do Sul.

Área de Concentração: Engenharia Elétrica.

Linha de Pesquisa: Engenharia Biomédica.

Pontifícia Universidade Católica do Rio Grande do Sul

Escola Politécnica

Programa de Pós-Graduação em Engenharia Elétrica

Orientador: Prof. Dario F. Guimarães de Azevedo, Ph.D.

Ficha Catalográfica

J99r Júnior, Milton Machado de Souza

Reconhecimento facial e fNIRS para a detecção de traços de emoção aplicados a jogos sérios / Milton Machado de Souza Júnior . – 2019.

117 p.

Dissertação (Mestrado) – Programa de Pós-Graduação em Engenharia Elétrica, PUCRS.

Orientador: Prof. Dr. Dario Francisco Guimarães de Azevedo.

1. Jogos Sérios. 2. Reconhecimento Facial. 3. fNIRS. 4. Emoções. 5. Classificadores. I. Azevedo, Dario Francisco Guimarães de. II. Título.

Elaborada pelo Sistema de Geração Automática de Ficha Catalográfica da PUCRS
com os dados fornecidos pelo(a) autor(a).
Bibliotecária responsável: Salete Maria Sartori CRB-10/1363



Pontifícia Universidade Católica do Rio Grande do Sul

ESCOLA POLITÉCNICA

PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA - PPGE

Reconhecimento Facial e fNIRS para a Detecção de Traços de Emoção Aplicados a Jogos Sérios

CANDIDATO: MILTON MACHADO DE SOUZA JÚNIOR

Esta Dissertação de Mestrado foi julgada para obtenção do título de MESTRE EM ENGENHARIA ELÉTRICA e aprovada em sua forma final pelo Programa de Pós-Graduação em Engenharia Elétrica da Pontifícia Universidade Católica do Rio Grande do Sul.

DR. DARIO F. GUIMARÃES DE AZEVEDO - ORIENTADOR

BANCA EXAMINADORA

DR. AUGUSTO BUCHWEITZ - PÓS-GRADUAÇÃO EM MEDICINA E CIÊNCIAS DA SAÚDE - PUCRS

DRA. MARLISE ARAUJO DOS SANTOS - PROGRAMA DE PÓS-GRADUAÇÃO EM ENGENHARIA ELÉTRICA - PUCRS

PUCRS

MILTON MACHADO DE SOUZA JÚNIOR

Reconhecimento facial e fNIRS para a detecção de traços de emoção aplicados a jogos sérios

A todos os meus professores, que lutam por uma educação digna e de qualidade.

Agradecimentos

A presente dissertação de mestrado não poderia chegar a bom porto sem o essencial apoio de várias pessoas.

Em primeiro lugar, não posso deixar de agradecer ao meu orientador, Professor Dr. Dario F. Guimarães de Azevedo, por toda a ajuda orientação e apoio no decorrer do trabalho. Gostaria de agradecer também ao Professor Dr. Alexandre Rosa Franco por toda sua paciência e conhecimento com que sempre me ajudou neste trabalho. Muito obrigado por me terem corrigido quando necessário sem nunca me desmotivar.

Agradeço a HP por proporcionar o desenvolvimento desde mestrado com o apoio financeiro.

Desejo igualmente agradecer a todos os meus colegas do Mestrado em Engenharia Elétrica, especialmente a Me. Guilherme Fróes Silva e Ingrid Gradaschi Lamadrid e Me Henrique Cabral.

Agradeço aos demais funcionários da Escola Politécnica, que foram sempre prestáveis e exemplares em transmitir seus conhecimentos e auxílios em busca de nossa excelência.

Por último, e mais importante quero agradecer à minha família, namorada e amigos pelo apoio incondicional sem o qual nada seria, especialmente aos meus pais pelas revisões incansáveis ao longo da elaboração deste trabalho.

"Comecemos, pois até agora pouco ou nada fizemos"

(São Francisco de Assis)

Resumo

O desempenho no aprendizado de escolares está diretamente ligado ao bem-estar do mesmo, o que consolida a necessidade da criação de ferramentas capazes de observar traços de emoção durante o desenvolvimento de atividades específicas. Aliando jogos educacionais à observação de emoções, pode-se garantir que as atividades de aprendizado sejam mais prazerosas, melhorando a eficiência dos escolares.

Para observar traços de emoções, deve-se evocá-las atentando para que estas estejam em destaque para a metodologia utilizada para a observação. Nesta pesquisa, usam-se duas formas de observação dos traços de emoção: micro expressões faciais e resposta hemodinâmicas do córtex cerebral. Na busca de evidências de emoções, utilizam-se métodos e protocolos de evocação dos traços de emoções, onde são utilizadas imagens fortes, responsáveis por provocar reações emocionais. Com a construção de um paradigma, pode-se observar a influência da sequência das imagens, usadas no exame, na variação de resposta das percepções emocionais.

Foram observadas emoções positivas, negativas e neutras, relacionando-as aos traços de expressão. O estudo se limita a diferenciar estas três emoções, que são possíveis de se distinguir a partir de análise fisiológica. Para tanto, utilizaram-se câmeras e sensores óticos, respectivamente, para reconhecer micro expressões da face e respostas hemodinâmicas corticais. Estas emoções foram classificadas a partir de um método de inteligência artificial cuja entrada são os traços emocionais. Este método cria um conjunto de padrões de respostas para cada uma das emoções, classificando a emoção atual do participante.

A possibilidade de gerar índices emocionais para uma atividade fim, amplia em muito o ensino de jogos sérios, mas também fortalece o uso deste método de reconhecimento em outras aplicações. A partir da união das duas metodologias de reconhecimentos fisiológicos, obteve-se um método robusto de classificação das emoções, com taxa de 77,2% de acerto.

Palavras-chaves: Jogos Sérios. Reconhecimento Facial. fNIRS. Emoções. Classificadores.

Abstract

This work aims to improve the recognition of emotional responses to different stimuli that provoke alterations into humans. Said alterations can be measured through several tools such as Magnetic Resonance Imaging, Functional Near Infrared Spectroscopy (fNIRS), Electroencephalography (EEG) and Facial Recognition. In this work, we used a combination of the fNIRS system, which measures the variation of hemoglobin oxygenation, and facial recognition tools.

The participants were stimulated with a sequence of images from the IAPS database, which are labeled with the weighted emotions they provoke. This allowed the training of a classifier that was capable of predicting the emotion that was experienced by the user during the activity performed. This article covers the recognition of positive, negative and neutral emotions using classifiers, created in Matlab, that are based on participant physiological responses. The combination of fNIRS, facial recognition and machine learning supported the creation of a predictor with 77.2% correct classification rate.

Key-words:Serious Games. Facial recognition. fNIRS. Emotions. Classifiers.

Lista de ilustrações

Figura 1 – Espectroscopia da luz	34
Figura 2 – Influência da luz no tecido	35
Figura 3 – Espectro de absorção de hemoglobina	35
Figura 4 – Função Bold (Deoxi e Oxi-hemoglobina)	36
Figura 5 – Espectro de absorção da hemoglobina	37
Figura 6 – NirsCap	37
Figura 7 – Sonda gerada pelos canais	38
Figura 8 – Sensores foto elétricos (optodos)	39
Figura 9 – NIRScout	39
Figura 10 – Circuito principal de controle e instrumentação	40
Figura 11 – Esquema de funcionamento do Equipamento	41
Figura 12 – Software NIRStar	41
Figura 13 – Mapa de criação para região motora	42
Figura 14 – Mapa de de visualização 3d	43
Figura 15 – Mapa de pontos de referência da face.	44
Figura 16 – Demonstração das expressões.	45
Figura 17 – Máquinas de vetores de suporte	49
Figura 18 – Árvore de decisões	51
Figura 19 – Fluxo de classificação de SVM	52
Figura 20 – Fluxo de classificação de um CNN	53
Figura 21 – Matriz de confusão do CNN	54
Figura 22 – Estímulos em blocos.	57
Figura 23 – Estímulos Independentes.	58
Figura 24 – Imagens similares as de evocação.	58
Figura 25 – Imagens da apresentação dos dados da IAPS.	59
Figura 26 – Gráfico de todos as amostras.	63
Figura 27 – Fluxo de processo para um estímulo.	64
Figura 28 – Sequência de perguntas.	64
Figura 29 – Mapa de disposições dos optodos.	66

Figura 30 – Matriz de canais.	67
Figura 31 – Definição dos períodos.	68
Figura 32 – Gráfico original e após remover o <i>Spike</i>	68
Figura 33 – Gráfico original e após passar o filtro.	69
Figura 34 – Índice de análise pelo NIRSlab	71
Figura 35 – Definindo os parâmetros	71
Figura 36 – SPM - level 2	72
Figura 37 – Contraste para comparações de fNIRS	73
Figura 38 – Respostas do dados fNIRS	74
Figura 39 – Fluxograma de aquisição do fNIRS	75
Figura 40 – Matriz erros do classificador floresta aleatória.	78
Figura 41 – Matriz erros do classificador floresta aleatória.	78
Figura 42 – Matriz erros do classificador SVM	79
Figura 43 – Fluxograma de aquisição do fNIRS.	81
Figura 44 – Resposta hemodinâmica corporal (BOLD).	83
Figura 45 – Fluxograma do sistema proposto de Reconhecimento	93
Figura 46 – Histograma das respostas para a Excitação	96
Figura 47 – Histograma das respostas de Valência	96
Figura 48 – Respostas de cada emoção do fNIRS	98
Figura 49 – Respostas entre emoções no fNIRS	99

Lista de tabelas

Tabela 1 – Tabela de resultados da matriz de confusão	53
Tabela 2 – Imagens selecionadas como positivas.	61
Tabela 3 – Imagens selecionadas como neutras.	62
Tabela 4 – Imagens selecionadas como negativas.	62
Tabela 5 – Médias para cada exames.	65
Tabela 6 – Conjunto de dados gerados por paciente.	70
Tabela 7 – Classificadores da CNN pelo <i>Matlab</i>	77
Tabela 8 – Classificadores da SVM pelo <i>Matlab</i>	78
Tabela 9 – Classificadores de ambos os métodos pelo <i>Matlab</i>	79
Tabela 10 –Conjunto de dados das câmeras para treino.	85
Tabela 11 –Conjunto de dados de fNIRS para treino.	86
Tabela 12 –Treino dos classificadores	87
Tabela 13 –Matriz de confusão do conjunto câmera.	91
Tabela 14 –Matriz de confusão do conjunto fNIRS.	91
Tabela 15 –Matriz de confusão do conjunto Ambos.	92
Tabela 16 –Médias para cada atividade.	97
Tabela 17 –Comparação entre método original e re-treinado.	100
Tabela 18 –Comparação entre método original e re-treinado.	101
Tabela 19 –Comparação entre os métodos originais e o re-treinado.	101
Tabela 20 –Respostas dos testes dos reconhecimentos faciais.	102
Tabela 21 –Respostas dos classificadores de similaridade entre sujeitos.	103
Tabela 22 –Respostas do classificador para os dados de treino.	104
Tabela 23 –Respostas do classificador treino em oito e testa em dois.	105
Tabela 24 –Respostas do classificador similaridade dos exames.	106
Tabela 25 –Respostas dos testes validação cruzada.	107
Tabela 26 –Matriz de confusão dos conjuntos.	107

Lista de abreviaturas e siglas

fNIRS	do inglês <i>Functional near-infrared spectroscopy</i> , que significa Espectroscopia Funcional de Próximo do Infravermelho.
CNN	do inglês <i>Convolutional Neural Networks</i> , que significa Rede Neural Convolutacional.
SVM	do inglês <i>Support vector machine</i> , que significa Máquina de Vetores de Suporte.
RF	do inglês <i>Randon Flores</i> , que significa Floresta aleatória.
EEG	Eletroencefalograma.
fMRI	do inglês <i>Functional magnetic resonance imaging</i> , que significa Imageamento por Ressonância Magnética Funcional.
BOLD	do inglês <i>Blood Oxygenation Level Dependent</i> , que significa Nível de Dependência de Oxigenação no Sangue.
PFC	Córtex pre-frontal.
IAPS	do inglês <i>International Affective Picture System</i> , que significa Sistema Internacional de Imagens Afetivas.
PCA	Análise de Componentes Principais.
FERET	do ingles <i>The Facial Recognition Technology</i> , que significa a Tecnologia de Reconhecimento Facial.
PET	Tomografia por Emissão de Pósitron.
HRF	do ingles <i>hemodynamic response functions</i> , que significa a função de resposta hemodinâmica.

Sumário

1	INTRODUÇÃO	27
1.1	Objetivos	29
1.2	Estrutura	29
2	FUNDAMENTOS TEÓRICOS	31
2.1	Jogos Sérios	31
2.1.1	Jogos adaptativos	32
2.1.2	Emoções como fonte de adaptação	33
2.2	Espectroscopia Funcional Próxima do Infravermelho (fNIRS)	34
2.2.1	Conceitos Físicos	34
2.2.2	Conceitos Fisiológicos	35
2.2.3	Os componentes do fNIRS	36
2.2.4	O Funcionamento do NIRS	40
2.3	Análise dos métodos de reconhecimento facial	43
2.3.1	Expressões faciais	45
2.3.2	Aprendizado de máquina	47
2.3.3	Extração da Face	48
2.3.4	Modelos de classificador de dados	49
2.3.5	Adaptação e uso do reconhecimento facial	51
2.3.5.1	Software SVM	51
2.3.5.2	Software CNN	53
2.4	Estado-da-arte	54
3	METODOLOGIA	57
3.1	Desenvolvimento do protocolo de evocação	57
3.1.1	Escolha do método de estímulos	58
3.1.2	Método de escolha das imagens de estímulo	59
3.1.3	Paradigma	63
3.2	Análise hemodinâmica usando fNIRS	65

3.2.1	Área de ativação de emoções corticais	65
3.2.2	Pre-processamento dos dados de fNIRS	67
3.2.3	Teste estatístico dos dados fNIRS	70
3.2.3.1	Teste SPM Nível 1	71
3.2.3.2	Teste SPM level 2	72
3.2.4	Normalização dos dados	74
3.3	Análise de emoções pelo reconhecimento facial	75
3.3.1	Gerando índices de confiabilidade	76
3.3.1.1	Avaliando os melhores classificadores	77
3.4	Voluntários e procedimentos	80
3.4.1	Descrição dos voluntários	80
3.4.2	Testes preliminares	80
3.4.3	Descrição do procedimento	80
3.4.4	Preparação do experimento usando fNIRS	81
3.4.5	Preparação do experimento usando webcam	82
3.5	Pre-processamento e adaptação de software de reconhecimento emocional	82
3.5.1	Pre-processamento - União das plataformas de dados	83
3.5.2	Classificador pelo Matlab	84
3.5.2.1	Ajuste dos conjuntos de dados para uso no Matlab	84
3.5.2.2	Conjunto de Classificadores 1 - Similaridade entre sujeitos.	88
3.5.2.3	Conjunto de Classificadores 2 - Teste de validação cruzada	89
3.5.2.4	Conjunto de Classificadores 3 - Similaridade entre exames	89
3.5.2.5	Conjunto de Classificadores 4 - integrando os exames e os sujeitos	90
3.6	Sistema de reconhecimento de sinais	92
4	RESULTADOS	95
4.1	Resultados do protocolo de evocações	95
4.2	Resultados do fNIRS	97

4.3	Resultado do reconhecimento facial	100
4.4	Resultados do software de reconhecimento	102
4.4.1	Conjunto de Classificadores 1 - Similaridade dos sujeitos. . .	102
4.4.2	Conjunto de Classificadores 2 - Teste de validação cruzada .	104
4.4.3	Conjunto de Classificadores 3 - Similaridade dos exames . .	105
4.4.4	Conjunto de Classificadores 4 - Treino em exames com todos os sujeitos	106
5	DISCUSSÃO	109
5.1	Desenvolvimento do protocolo de evocação	109
5.2	Análise hemodinâmica usando fNIRS	110
5.3	Análise de emoções pelo reconhecimento facial	111
5.4	Pré-processamento e adaptação de software de reconhecimento emo- cional	112
6	CONCLUSÃO	115
6.1	TRABALHOS FUTUROS	116
	Referências	117
	Anexos	121

1 INTRODUÇÃO

O cérebro é uma máquina de aprender. Estamos diante da possibilidade de proporcionar uma aprendizagem mais eficaz e prazerosa com base em nossos conhecimentos sobre o funcionamento cerebral. Qualquer pessoa é capaz de aprender, visto que a aprendizagem não se resume à infância e à adolescência: ela acontece durante toda a vida.

Atualmente está em amplo crescimento o desenvolvimento de novos jogos sérios (SUSI; JOHANNESSON; BACKLUND, 2007; BOYLE et al., 2016). Estes jogos usualmente contemplam jogos de computadores onde o usuário “aprende jogando”. Tendo como objetivo principal a educação, jogos sérios são atualmente aplicados em diversas áreas, incluindo educação básica, medicina e militar.

A eficácia no processo de aprendizado depende de muitos fatores, incluindo o bem-estar, foco e interesse do aluno no conteúdo a ele apresentado. Em métodos de ensino tradicionais, é possível o professor observar, e até mensurar, a atenção do aluno. Isto é possível através da observação das expressões corporais e faciais do aluno. No entanto, em jogos sérios, este tipo de observação, atualmente, não é realizado de forma efetiva.

A indústria está desenvolvendo jogos cada vez mais interativos, que visam aumentar o interesse dos alunos ao jogo. Porém, a medida mais utilizada para medir o engajamento do aluno em relação ao jogo e o processo de aprendizado está ligada ao score do aluno no jogo. Para manter a foco e o engajamento dos alunos, jogos atuais permitem aumentar ou reduzir sua dificuldade. Sendo assim, o aluno não desiste do jogo por ser difícil ou fácil demais.

Existem estudos recentes tentando mensurar o engajamento e emoções do aluno em relação ao jogo utilizando diversas medidas, como sinais cardiológicos (VACHIRATAMPORN et al., 2015), eletromiografia (EMG) para mensurar emoções pela face (HAZLETT, 2006), condutância da pele (KIVIKANGAS et al., 2011) e pela posição da cadeira (BIANCHI-BERTHOUBE et al., 2006). Neste úl-

timo estudo foi concluído que quanto mais a cadeira estava estática, mais focado estava o jogador. Com essas medidas é possível adaptar o jogo baseando-se nas emoções dos usuários e não somente no seu desempenho no jogo.

Esta dissertação apresenta outras medidas para quantificar o engajamento e as emoções dos alunos em jogos sérios por outras medidas fisiológicas. Mais especificamente avaliou-se a utilização de uma câmera de *laptop* (*webcam*) para reconhecer as emoções do jogador através da face e a *Functional Near InfraRed Spectroscopy* (fNIRS) para medir a atividade cerebral específica da resposta emocional.

A utilização da câmera para detectar emoções foi escolhida, pois em todos os *laptops* modernos há uma *webcam*, sendo assim possível aplicar o *software* o qual está sendo desenvolvido neste estudo em larga escala. Logo, será possível difundir a sua aplicação em jogos sérios. Os métodos de reconhecimento de emoções por câmeras já estão em ampla utilização, com as mais recentes aplicações utilizando técnicas avançadas de aprendizado de máquina, como por exemplo, o *Deep Learning* (KIM; LEE; PROVOST, 2013).

O fNIRS, por sua vez, não é um equipamento para ser utilizado em larga escala. Porém é uma medida indireta da atividade cerebral (sinal BOLD – do inglês *Blood Oxygen Level Dependent*, que significa Dependência do Nível de Oxigênio no Sangue). Este equipamento, por sua vez, será utilizado para corroborar com os resultados do sistema de reconhecimento de emoções.

Em virtude das dificuldades encontradas nos métodos de aprendizagem, hora por tempo, hora pelo ambiente desfavorável, o estudo propõe a utilização de uma ferramenta que seja capaz de adaptar-se as características cognitivas emocionais do aluno, tornando o aprendizado favorável. Essa ferramenta realiza um monitoramento emocional correlacionado à atividade que o aluno esta desenvolvendo, visando assim posteriormente adaptar a atividade para que a mesma esteja sempre com alta aceitação do usuário.

1.1 Objetivos

O objetivo principal desta dissertação é detectar os traços emocionais, através do processamento das imagens corticais a partir do fNIRS, enquanto monitora-se as expressões faciais.

Definimos os seguintes objetivos parciais que acompanham o progresso deste estudos:

1. Desenvolver protocolo para evocar emoções.
2. Definir e analisar os métodos de reconhecimento facial.
3. Analisar e definir as configurações utilizadas no fNIRS.
4. Mapear as áreas do cérebro responsáveis pelo reconhecimento de emoções a partir do fNIRS.
5. Desenvolver um método responsável por integrar os sinais de emoção gerados uma análise dos dados.

1.2 Estrutura

No capítulo seguinte, examinou-se as principais abordagens utilizadas em relação à detecção dos sinais de emoções, visando melhor situar o leitor. Além disso, construiu-se um breve histórico sobre reconhecimento de padrões de emoções e suas características para cada método específico utilizado. O capítulo 3 descreve e elucida, em detalhes, as etapas que compõem a análise, preparação e construção dos métodos de detecção de emoções desenvolvidos. Serão Apresentados resultados específicos de cada um dos métodos no capítulo 4. No capítulo 5, teceu-se uma discussão do projeto e os resultados obtidos. Por fim, considerações finais foram feitas no capítulo 6, no qual também foram apresentadas possibilidades de trabalhos futuros.

2 FUNDAMENTOS TEÓRICOS

A abordagem utilizada com maior frequência para a caracterização de emoções, baseia-se na resposta do paciente a uma série de perguntas específicas para montar o classificador. Os testes emocionais avaliam respostas subjetivas ou diretas, e este somatório demonstra a propensão emocional a partir destas respostas (IHHP, 2017).

Outra tendência é avaliar as propriedades de usuários em suas redes sociais e seu uso da internet (HEATMAP, 2017). Em aplicações de uso *web*, as técnicas aproveitam as várias formas de obtenção destes dados como, por exemplo, a ferramenta *Heat Maps* (ŠPAKOV; MINIOTAS, 2007), que cria uma matriz de intensidade onde o usuário mais mantém os olhos sobre a tela.

A detecção de padrões de emoções pode ser observada por três métodos: análises psicológicas, análises cerebrais e expressões faciais. Essas aplicações fazem uso de métodos diversos. Um exemplo são as análises cerebrais que utilizam imagens funcionais de ressonância magnética (FUSAR-POLI et al., 2009). Neste caso a modelagem de dados é complexa, demorada e extremamente cara, não sendo aplicável a jogos sérios. Cada estudo apresentado gera respostas características a cena que ele está integrado, derivadas do uso de estímulos diferentes.

2.1 Jogos Sérios

Jogos sérios, são ditos sérios, porque sua finalidade é a resolução de uma atividade séria, usualmente o ensino. São ferramentas construídas com os princípios de serem jogos interativos, assumindo o princípio de transmitir conteúdo educacional de modo envolvente e imersivo. Seu propósito estabelece a necessidade da combinação de cognição, comportamento e habilidades, buscando chegar ao propósito final do aprendizado.

"O que separa *serious games* do restante [dos jogos] é o foco em um resultado de aprendizado específico e intencional para alcançar mudanças de performance e comportamento sérios, mensuráveis e continuadas"(DERRYBERRY, 2008).

Com o aprimoramento e a ampliação de jogos sérios para fins educacionais há a necessidade da ferramenta se tornar atraente para o aprender. Algumas aplicações se concentram nas emoções como características vinculadas ao aprendizado. Em sua grande maioria, as novas formas de desenvolvimento de jogos sérios direcionam-se para pesquisas nas áreas de jogos adaptativos.

2.1.1 Jogos adaptativos

Estes jogos utilizam das abordagens micro e macro adaptativas. A abordagem macro adaptativa usa o contexto aptidão, tratamento e interação. Esta ligação torna como princípio de adaptação os modelos de estereótipos dos jogadores. Modelos estes são pré-definidos no desenvolvimento da atividade. Por outro lado a abordagem micro adaptativa fornece o contexto da autonomia dos elementos característicos do modelo dos jogadores, design, dificuldade e outras necessidades dos mesmos (MA; OIKONOMOU; JAIN, 2011).

Dentro dos elementos de adaptação estão três condições fundamentais recorrentes dos sistemas:

"(1) mudanças na mecânica e representação dos sistemas visíveis no comportamento dos sistemas para (2) melhor ajustar as necessidades e desejos de usuários para (3) otimizar a interação entre sistema e o usuário para alcançar uma certa qualidade."(MA; OIKONOMOU; JAIN, 2011)

Com a implementação destes critérios pode-se garantir uma abordagem focada ao jogador. Já que a mecânica do jogo responde diretamente aos estímulos gerados pelo jogador, com objetivo de aperfeiçoar o jogo. Quando a adaptação do

jogo gera uma modelagem do jogador, o método apresenta inúmeras variáveis como aptidão, conhecimento, habilidades, domínio, etc. E são essas características que os diferenciam entre os jogadores virtuais (MAGERKO et al., 2008; BRUSILOVSKY; KOBASA; NEJDL, 2007).

2.1.2 Emoções como fonte de adaptação

Para se chegar a adaptação impulsionada pela emoção, pode-se usar diversos elementos específicos dos jogos, interagindo não só na jogabilidade mas também em elementos de design visual, estrutura e história do jogo. Pode-se utilizar de forma concisa os traços de emoções para a construção de *feedback* de emoção. E estes *feedbacks* por sua vez serem trabalhados como resposta a jogos sérios, comprovando os achados de Cacioppo et al. (2000), que revela a diferenciação entre as emoções, sendo maior ativação hemodinâmica das emoções negativas comparadas a emoções positivas.

A modelagem baseada nas emoções é mais refinada que as demais, devido a homogeneidade das expressões. Quando consideramos as características do jogador, suas percepções são mais homogêneas, mesmo entre jogadores distintos. É perceptível também que o progresso, com o decorrer do jogo, é sensível a qualidade do mesmo (DÖRNER et al., 2016).

As emoções podem ser expressadas de diversas formas, pois não causam apenas alterações subjetivas, mas também atividade fisiológica e expressões faciais. O campo da fisiologia apresenta diversas áreas de atuação podendo obter medidas de quase todas as respostas emocionais do corpo. Os dados fisiológicos que podemos inferir estão ligados ao sistema nervoso somático, como frequência cardíaca, pressão e volume sanguíneo e ao sistema nervoso central, pela atividade cerebral. Esta pode ser observada pela oxigenação do sangue, através do *fNIRS*.

Há outras medidas fisiológicas como a atividade eletrodermal, chamada também de condutância da pele, que já é consolidada como boa medida fisiológica, mas que comparada ao potencial emocional, não apresenta parâmetros para uma resposta consistente (LARSEN et al., 2008).

É possível monitorar as emoções expressas pela face, inferindo-as de formas

diferentes, como com o uso de eletromiografia (*EMG*) ou com imagens visuais da face, método similar ao processado por humanos involuntariamente.

Portanto a escolha do melhor método de obtenção dos dados é de suma importância para a qualidade da resposta do sistema, como também a escolha das variáveis medidas nestes métodos.

2.2 Espectroscopia Funcional Próxima do Infravermelho (fNIRS)

A Espectroscopia Funcional Próxima do Infravermelho (*fNIRS*) é uma tecnologia de neuroimagem funcional em ascensão, tanto pelas questões de custo quanto por se tratar de um método não invasivo, seguro, prático e portátil. A variabilidade na construção favorece diferentes exames e garante mobilidade, o que se torna uma vantagem importante deste método. Como métodos de Ressonância Magnética Funcional (*fMRI*) e Eletroencefalograma (*EEG*), o *fNIRS* pode ser utilizado para monitorar respostas cerebrais a determinados estímulos.

2.2.1 Conceitos Físicos

O estudo da interação da radiação eletromagnética com todos os tipos de matéria é chamado espectroscopia. Avanços tecnológicos acarretaram uma melhoria de eficiência e na criação de diversos equipamentos de medidas ópticas, capazes de realizar inúmeras novas técnicas de espectroscopia em diversos materiais (SALA, 1996). Em sua maioria, estes equipamentos são utilizados como ferramenta para mapear e identificar estruturas moleculares como apresentado na Figura 1.

Figura 1 – Espectroscopia da luz



Fonte: (SALA, 1996)

Dentro destas estruturas moleculares foca-se na hemoglobina, devido a esta ser a transportadora de oxigênio em todos os animais vertebrados. Neste caso, a propagação ocorre através do tecido que é atravessado pela luz. Esta se comporta de diferentes formas dependendo do tipo ou espessura do tecido, como pode-se observar na Figura 2.

Figura 2 – Influência da luz no tecido



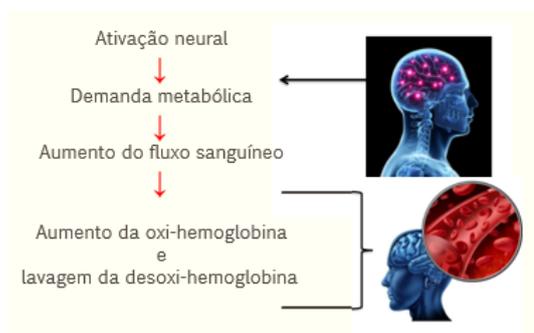
Fonte: <http://nirx.net/fnirs-and-nirx/>

Pode-se concluir que a luz vermelha é a que apresenta melhor incidência e rendimento em comparação aos outros espectros possíveis, visto que é a única que apresenta incidência no outro lado. No entanto, apresenta a limitação de não permitir a observação interna do tecido.

2.2.2 Conceitos Fisiológicos

É possível medir a variação de oxigênio no sangue (*BOLD*) através do *fMRI*, decorrente de atividades funcionais no processo corporal demonstrado na Figura 3.

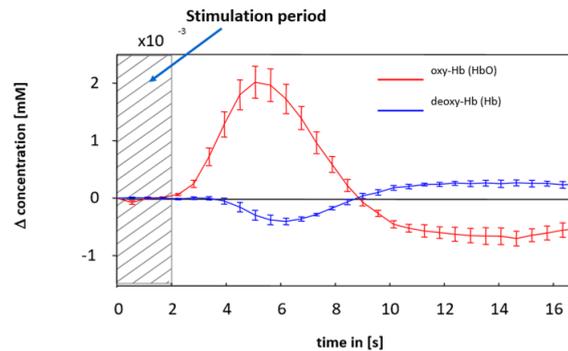
Figura 3 – Espectro de absorção de hemoglobina



Fonte: (SCHMITZ, 2016)

Com o uso do *fNIRS*, tem-se uma variação de oxigenação e desoxigenação cerebral. Percebe-se a variação hemodinâmica ligada a absorção na área cortical do cérebro, acarretada pelas alterações de concentração local de oxi e desoxi-hemoglobina a partir da atividade cerebral funcional. Este fenômeno fica mais compreensível no gráfico mostrado na Figura 4.

Figura 4 – Função Bold (Deoxi e Oxi-hemoglobina)



Fonte: <http://www.ncbi.nlm.nih.gov/pubmed/11131197>

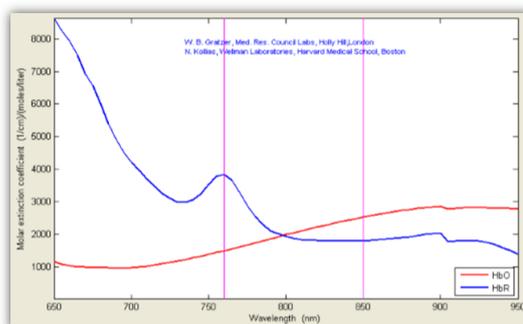
As mudanças hemodinâmicas apresentam influência na dispersão da luz, e esta influência gera o sinal de variação medido pelo equipamento, em forma de emissão e recepção de luz. O espectro de absorção de luz na área cortical para cada comprimento de onda é apresentado na Figura 5 (SCHMITZ, 2016). Com isto, o sistema *fNIRS* usa múltiplos comprimentos de onda para distinguir a oxi e a desoxi-hemoglobina, são esses 760nm e 850nm, indicados pela linha vertical, em rosa.

A escolha destes dois comprimentos de onda se concretiza pela análise mostrada na Figura 5, onde para a frequência de 760nm temos a desoxi-hemoglobina apresentada com coeficiente de intensidade maior. Já no caso de 850nm há uma inversão deste fator e a oxi-hemoglobina apresenta intensidade maior.

2.2.3 Os componentes do *fNIRS*

O equipamento é composto por periféricos essenciais para o funcionamento e a utilização do *fNIRS*. Sendo assim, faz-se uma descrição destes critérios. Para

Figura 5 – Espectro de absorção da hemoglobina



Fonte: (SCHMITZ, 2016)

cada periférico é mostrado sua construção, suas características e seus princípios de funcionamento.

1. NISRCap

O *NISRCap*, como apresentado na Figura 6, trata-se de uma touca em malha construída para diversas montagens. Este equipamento tem sua criação derivada da touca utilizada para realizar exames de *EEG*.

Figura 6 – NirsCap



Fonte: (TECHNOLOGIES., 2016a)

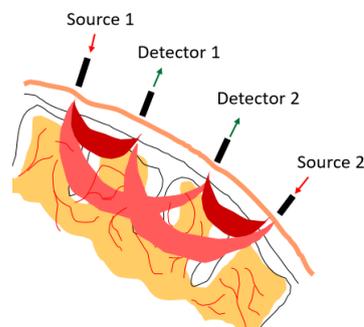
Junto a esta malha são posicionados, em pontos estratégicos, os locais para a colocação dos optodos. Isto favorece a repetibilidade de experimentos, como também uma maior certeza da posição dos dados adquiridos. Os locais de fixação dos optodos replicam a posição dos eletrodos no *EEG* padrão. As posições

dos optodos na cabeça são personalizáveis dependendo do experimento. Fontes e detectores podem ser colocados de acordo com qualquer layout desejado e os canais definidos de acordo (TECHNOLOGIES., 2016a).

2. Optodos

A montagem dos optodos gera um mapeamento espacial da atividade cerebral, realizando uma matriz de múltiplos pares fonte-detector que são colocados sobre a área de interesse, produzindo uma matriz de $n \times m$ canais, como apresentado na Figura 7.

Figura 7 – Sonda gerada pelos canais



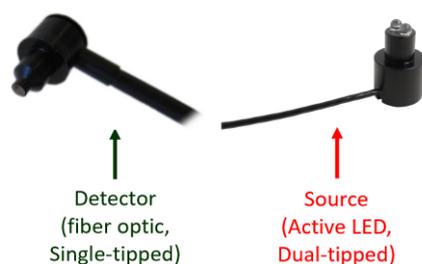
Fonte: (SCHMITZ, 2016)

O método de observar os sinais cerebrais é através dos optodos, os emissores e detectores de luz que realizam a irradiação e detecção da luz. Estes optodos, são diodos emissores de luz e receptores de luz, como apresentados na Figura 8. Parte da luz que incide sobre o meio é refletida pela superfície e parte penetra no meio. Esta penetração se dá tanto por absorção quanto por espalhamento, sendo transmitida através do meio.

3. NIRScout

O equipamento inclui a touca (NIRScap) para definir as estruturas de montagem, os sensores e emissores para captação do sinal, e, principalmente, o *NIRScout*, apresentado na Figura 9, a central de processamento dos dados, condicionamento de sinais e conjuntos de controles.

Figura 8 – Sensores foto elétricos (optodos)



Fonte: (SCHMITZ, 2016)

Figura 9 – NIRScout



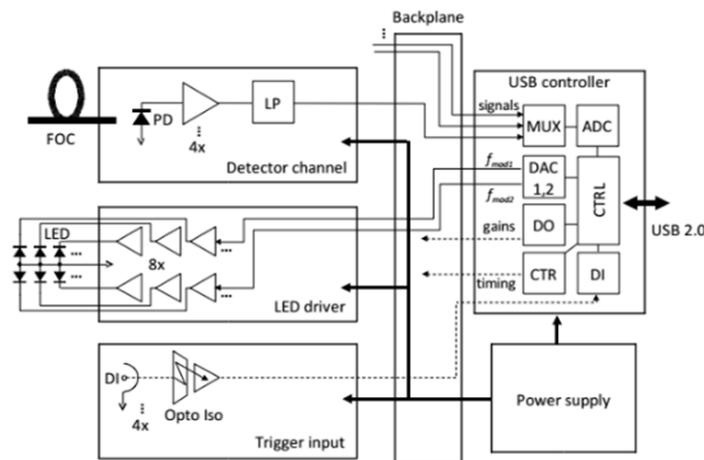
Fonte: (TECHNOLOGIES., 2016c)

O *NIRScout* realiza medições tomográficas difusas de ondas contínuas de comprimento duplo, perto do infravermelho, em estruturas de tecidos com altas taxas de amostragens. O sistema facilita a iluminação de vários locais de destino em varredura multiplexada no tempo e na frequência. A multiplexação do uso de emissores LED em dois comprimentos de onda distintos permite a discriminação dos dois estados de oxigenação da hemoglobina no tecido. Ambos os comprimentos de onda são emitidos simultaneamente, sendo distinguidos pela modulação e demodulação de cada frequência na faixa de kHz (TECHNOLOGIES., 2016c).

O instrumento emprega leitura paralela de múltiplos canais de detecção óp-

tica, cada um dos quais usa troca de ganho adaptativo para maximizar a faixa de medição dinâmica, operado através de uma interface gráfica de usuário em um computador pessoal. Apresenta junto a isto, entradas e saídas digitais em linhas, contadores, conversores analógico-digitais e conversores digital-analógicos, como apresentado na Figura 10.

Figura 10 – Circuito principal de controle e instrumentação



Fonte: (TECHNOLOGIES., 2016c)

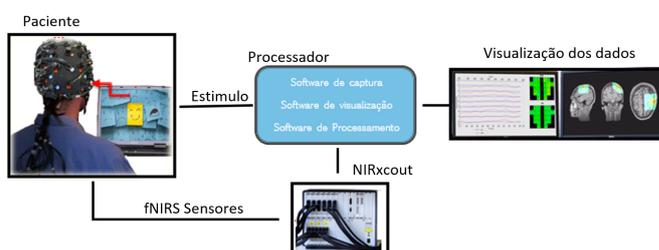
2.2.4 O Funcionamento do NIRS

A funcionalidade no uso do *fNIRS* está ligado a integração *hardware* e *software*, como visto na Figura 11, que permite a realização de exames e obtenção de dados em tempo real. Além disto, existem outros softwares de pós processamento dos dados permitindo análises estatísticas.

1. Obtenção dos dados

O sistema gera uma quantidade grande de dados em forma de matrizes de intensidade. Para capturar estes dados, é empregada uma plataforma de processamento de dados robusta e capaz de suprir e estruturar estes dados. O software *NIRStar*, é a ferramenta que capta estes dados, nos seus respectivos tempos, armazenando os n canais em forma de funções de intensidade para cada canal. Podemos ver, na Figura 12, a apresentação dos diversos canais e

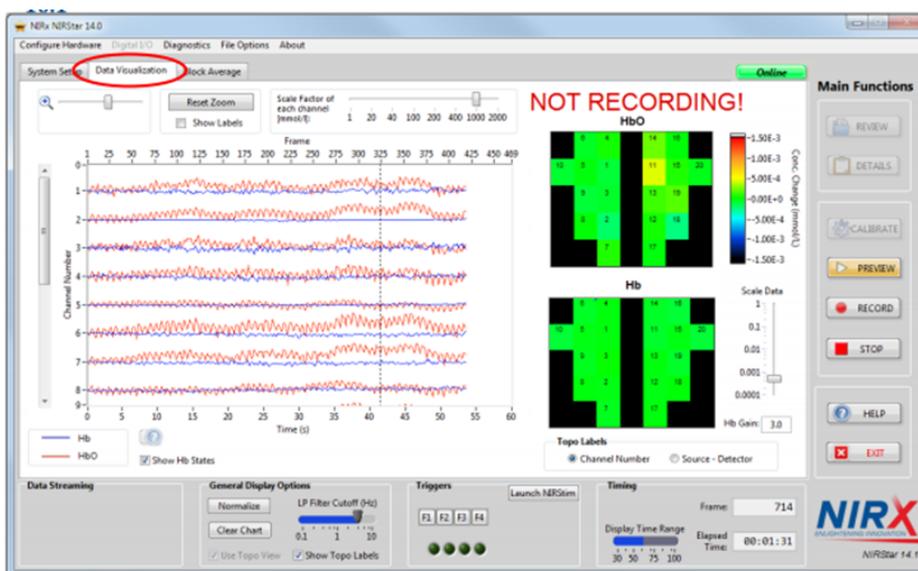
Figura 11 – Esquema de funcionamento do Equipamento



Fonte: O Autor (2018)

sua visualização de coloração diversificada que é relacionada às intensidades de ativação.

Figura 12 – Software NIRStar

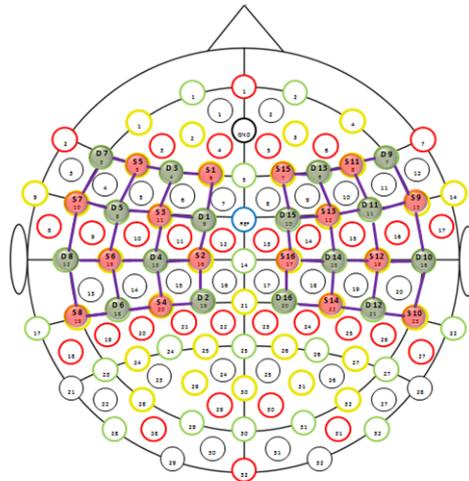


Fonte: O Autor (2018)

Para melhorar a compreensão das formas de obtenção e criação das imagens, descreve-se as atividades principais presentes neste processo.

Após a especificação do hardware, a estrutura necessita ser especificada entre qual a montagem utilizada e a posição dos canais nas áreas de interesse.

Figura 13 – Mapa de criação para região motora



Fonte: (TECHNOLOGIES., 2016a)

Neste item, é possível utilizar áreas pré determinadas, como a apresentada na Figura 13, ou criar uma nova área específica para o experimento.

O sistema de captura e análise funciona realizando leituras a partir da variação do sinal versus a linha de referência em repouso *baseline*. Para se obter a *baseline* e a calibração do sistema, todo o teste inicia com um momento em repouso, onde uma varredura realiza o mapeamento e ajuste dos ganhos para cada canal. Após realizadas as fases de configuração, o exame pode ser realizado.

Uma segunda funcionalidade é a visualização das ativações em tempo real onde o software utiliza uma interface visual gráfica construída em *Matlab*. Este procedimento auxilia a verificação do exame em tempo real para sua realização correta.

2. Processamento dos dados

Para o processamento dos dados, onde se converte os dados brutos em uma imagem estrutural e de fácil compreensão, utiliza-se o software *NIRSlab*.

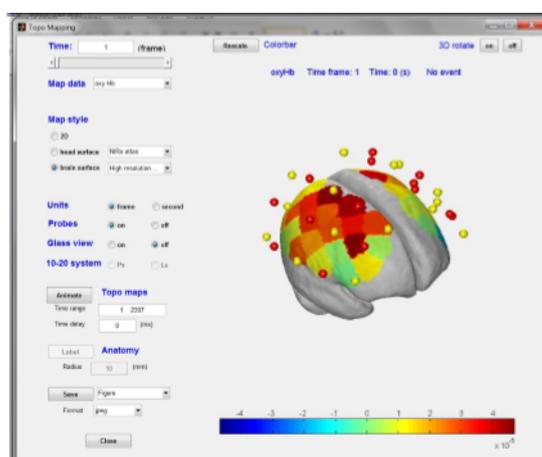
Como o pós processamento é complexo, este é subdividido em cinco etapas. Deve-se, na primeira etapa, verificar as marcações temporais realizadas pelo software, chamadas de *triggers*. A etapa dois é onde realiza-se o pré processa-

mento, onde os sinais são corrigidos. Estas correções retiram as alterações que não tenham sido realizadas pelas respostas hemodinâmicas, e aplicam filtros para remoção de ruídos.

A etapa três, define os parâmetros de condições hemodinâmicas e realiza-se a conversão dos dados em sinais de luz para sinais em critérios hemodinâmicos. Este procedimento é realizado no comando *Compute*.

Na etapa quatro, os dados são visualizados em séries hemodinâmicas, onde são exibidos em três dimensões, como apresentado na Figura 14, apresentando as ativações hemodinâmicas ao longo do procedimento.

Figura 14 – Mapa de de visualização 3d



Fonte: O Autor (2018)

A última etapa trata das questões estatísticas onde realiza-se as análises avaliando sujeito a sujeito, ou em grupos. Esta etapa está ligada à análise de respostas científicas.

2.3 Análise dos métodos de reconhecimento facial

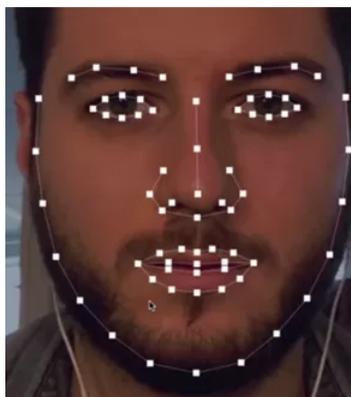
Há diversas maneiras de se analisar as expressões faciais, observando gestos, formato do rosto, da boca ou sobrancelhas. As micro expressões são uma análise

muito mais específica de pequenos traços como covas, rugas e o franger do rosto. Estas características foram apresentadas anteriormente.

Existem diversas ferramentas que trabalham as formas de reconhecimento facial, como por exemplo a *Emotion API - Microsoft*¹. Esta ferramenta apresenta uma confiabilidade sobre a emoção das pessoas presentes em uma imagem. Para este fenômeno usa um banco de dados web como comparativo. A *Google Cloud Vision*² utiliza o mesmo conceito do software base, utilizando o seus próprio bancos de dados.

Sistemas de reconhecimento de emoções baseados em visão computacional e inteligência artificial analisam uma sequência de imagens para prever as emoções expressadas pelo participante enquanto faz uma atividade. Os novos estudos se diferenciam principalmente por se tratarem de análises em tempo contínuo. Estes métodos favorecem muito em relação as demais formas, pois ampliam as aplicações destas ferramentas.

Figura 15 – Mapa de pontos de referência da face.



Fonte: <https://www.weet-magazine.nl/hoe-zie-jij-eruit-in-de-toekomst/>

Como pode-se observar na Figura 15, o formato da boca e dos olhos demonstram as regiões da face onde são marcados os pontos que delimitam a periferia da expressão. A partir destes pontos é possível realizar a análise de traços de emo-

¹ <https://docs.microsoft.com/pt-br/azure/cognitive-services/emotion/home>

² <https://cloud.google.com/vision/>

ção, estes expressam emoções distintas, e são responsáveis pela classificação dos padrões de emoções.

2.3.1 Expressões faciais

Existem modelos de micro expressões já estudados, em que se pode classificar as emoções referentes a uma análise facial do indivíduo, como apresentado por Fridlund, Ekman e Oster (1987).

Figura 16 – Demonstração das expressões.



Fonte: <http://www.thejournal.ie/lying-2308154-Sep2015/>

O estudo feito por (FRIDLUND; EKMAN; OSTER, 1987), demonstra a uniformidade das expressões faciais independente das culturas, analisados 10 po-

pulações distintas. Esta comprovação fortalece o uso de reconhecimento facial, usando as micro expressões apresentadas na Figura 16.

1. Expressões do Medo

2. Principais micro expressões faciais do medo:

a) As sobrancelhas são levantadas como na surpresa, mas elas buscam se encontrar no centro da face, como se elas levantassem para dentro.

b) Os olhos são abertos, sendo a pálpebra superior levantada e a inferior tensionada, a boca se abre em medo, mas os lábios são tensionados e podem recuar firmemente.

Todos esses elementos podem aparecer sozinhos, em dupla ou todos juntos para demonstrar medo ou sentimentos diferentes. O medo também pode ser seguido de qualquer outra emoção.

3. Principais micro expressões faciais da raiva:

a) As sobrancelhas são aproximadas, abaixadas e normalmente não aparecem rugas na testa (como no medo).

b) As pálpebras podem ou não estar tensionadas, e os olhos parecem mais focados e penetrantes.

c) A boca pode estar com os lábios pressionados um contra o outro e para cima, ou aberta de forma mais quadrada com os dentes a mostra (normalmente quando a pessoa raivosa está falando).

A raiva só é caracterizada se as três áreas são observadas. Se apenas uma ou duas áreas estiverem com expressão de raiva, a emoção pode não necessariamente ser de raiva.

4. Principais micro expressões faciais da Surpresa:

a) As sobrancelhas são levantadas para fora e curvadas.

b) Os olhos são abertos, sendo a pálpebra superior levantada e a inferior relaxada.

c) A boca se abre, mais precisamente a mandíbula cai, com os lábios e dentes separados.

Todos esses elementos podem encontra-se sozinhos ou todos juntos para demonstrar medo ou sentimentos diferentes. A surpresa sempre é seguida de qualquer outra emoção.

5. Principais micros expressões faciais que demostramos quando estamos com felicidade:
 - a) Os cantos da boca são puxados para trás e para cima – podendo ser de maneira leve ou mais severa.
 - b) Os lábios podem estar juntos ou separados, mostrando os dentes também de maneira leve ou mais severa.
 - c) Dobras aparecerem do entre a base nariz e a boca com o levantamento dos cantos da boca e a bochecha se eleva.
 - d) A pele abaixo da pálpebra inferior é levantada e podem ser formadas “rugas” na lateral dos olhos.
 - e) Os olhos também podem apresentar um certo brilho.
6. Principais micro expressões faciais do Desgosto:
 - a) O lábio superior da boca é levantado, causando uma mudança de aparência do nariz, muitas vezes com rugas na parte superior ou lateral.
 - b) O lábio inferior da boca pode também estar levemente a frente ou a baixo.
 - c) As bochechas são levantadas causando uma mudança de aparência das pálpebras inferiores. A sobrancelha pode ou não estar abaixada.

As ferramentas que convertem a informação do reconhecimento facial em uma resposta emocional, se utilizam de aprendizado de máquina para a construção de mecanismos inteligentes, capazes de classificar as emoções a partir das variações da face.

2.3.2 Aprendizado de máquina

O aprendizado de máquina é uma área da ciência da computação que transformou as formas de reconhecimento de padrões e teoria do aprendizado computacionais. O uso de inteligência artificial, neste contexto, torna o sistema mais dinâmico e automático, gerindo as previsões a partir dos dados e não mais seguindo

linhas inflexíveis de código. Bons exemplos estão ligados na sua maioria a previsões do tempo e da bolsa de valores (MICHALSKI; CARBONELL; MITCHELL, 2013).

Dentro desta área se enquadram diversas estruturas de processamento e estatística dos dados. Podendo realizar regressões, reduções, agrupamentos e classificações. Estruturas estas baseadas principalmente nestas quatro formas de aprendizado.

1. Aprendizado supervisionado
2. Aprendizado não supervisionado
3. Aprendizado semisupervisionado
4. Aprendizado por reforço

A principal linha é a aprendizagem profunda com arquitetura estruturada em redes neurais profundas. Sendo seu aprendizado supervisionado, semisupervisionado ou não supervisionado, aplicada a diversas áreas como reconhecimento de fala, de visão, tradução automática e bioinformática.

2.3.3 Extração da Face

O método de reconhecer emoções passa primeiramente pelo reconhecimento da face humana, usando PCA (análise de componentes principais). As PCA são técnicas estatísticas construídas com algoritmos de decisão, que utilizam redução de dimensionalidade em uma decomposição ótima de mínimos quadrados lineares, aplicados no conjunto de treino (KIRBY; SIROVICH, 1990). As PCA em reconhecimento de face podem ser divididas em três módulos. O primeiro módulo, como pre-processamento, é onde a imagem sofre quatro etapas, o filtro de compressão, normalização geométrica, máscara de *pixels* e, por último ajuste, a iluminação da imagem. No segundo módulo, a imagem é decomposta nos seus autovetores e autovalores. Já no terceiro módulo, classifica-se os parâmetros de variação dos autovetores e autovalores, que analisam as faces processadas, segundo sua semelhança (MOON; PHILLIPS, 2001).

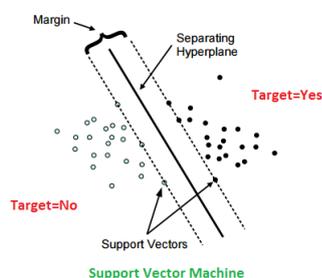
Cada PCA apresenta características específicas, causando resultados diferenciados, portanto, utiliza-se métodos de avaliação de desempenho (*FERET*), que identifica o desempenho de cada classificador. O *FERET* trata-se de um banco de dados, de imagens faciais padronizadas e estáticas. Portanto faz a avaliação e análise dos resultados, para as imagens do seu banco, permitindo medir o desempenho (PHILLIPS et al., 2000).

2.3.4 Modelos de classificador de dados

Foram analisados os melhores classificadores para o estudo, serão descritos os modelos máquina de vetores de suporte (SVM) e modelo Floresta aleatória, para melhor compreensão.

Uma máquina de vetores de suporte (*SVM do inglês Suport Vector Machine*) é um algoritmo de aprendizado de máquina que encontra um hiperplano ideal entre os pontos de dados mais próximos a cada classe, a fim de maximizar as distâncias entre o hiperplano e a classe, como apresentado na Figura 17.

Figura 17 – Máquinas de vetores de suporte



Fonte: <https://cran.r-project.org/web/packages/e1071/vignettes/svmdoc.pdf>

Uma SVM pode ser aplicada para mais de uma dimensão, para um maior conjunto de dados, utilizando funções de *Kernel*³. Como a SVM é um teste binário, para o caso de medir emoções variadas, a implementação necessita de um *kernel* de base radial (*RBF*). As emoções são separadas em cascata de SVMs, gerando múltiplas classes.

³ Kernel é uma função que retorna o produto escalar das dimensões.

Para dar início ao uso de uma SVM, usa-se os dados previamente processados. Seja, então, a matriz T (conjunto de dados de treinamento), construídos com $x_i \in X$ dados de entrada e suas respectivas classes $y_i \in Y$, $Y = [0, 1, 2]$.

Para separar os dados em espaços de entrada, utiliza-se a equação 2.1 do hiperplano, gerando um hiperplano para cada combinação das saídas.

$$f(\mathbf{x}) = \mathbf{w} \cdot \mathbf{x} + b = 0 \quad (2.1)$$

Na equação w e x representam o vetor normal do hiperplano e b representa as distâncias no hiperplano.

A partir desta equação se tem duas regiões: ($w\hat{u}x + b > 0$ e $w\hat{u}x + b < 0$). Quando destas regiões definidas, usa-se a função responsável pela obtenção da classificação, conforme demonstrado na Equação 2.2 (SMOLA et al., 2000).

$$g_1(\mathbf{x}) = \text{sgn}(f(\mathbf{x})) = \begin{cases} +1 & \text{se } \mathbf{w} \cdot \mathbf{x} + b > 0 \\ 0 & \text{se } \mathbf{w} \cdot \mathbf{x} + b < 0 \end{cases} \quad (2.2)$$

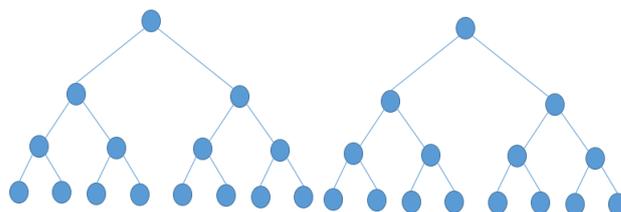
Com a função resultante da Equação 2.2, obtém-se hiperplanos equivalentes a multiplicação w e b por uma constante. Para encontrar o hiperplano canônico em relação ao conjunto de treino T , onde os exemplos mais próximos ao hiperplano ($w\hat{u}x + b = 0$) satisfazem a Equação 2.3.

$$|\mathbf{w} \cdot \mathbf{x}_i + b| = 1 \quad (2.3)$$

Para obter a distância entre os hiperplanos apresentados na Equação 2.2, onde: $X1 : w\hat{u}x + b = 1$ e $X2 : w\hat{u}x + b = 0$ são pontos dos hiperplanos, determina-se a d (distância entre os pontos).

Floresta aleatória (*RF do inglês Random Forest*) é um algoritmo de aprendizado de máquina que utiliza classificação e regressão em conjunto. Neste algoritmo é utilizada a técnica de árvore de decisão apresentada na Figura 18.

Figura 18 – Árvore de decisões



Fonte: O Autor (2018)

Para a construção de uma floresta aleatória são utilizadas diversas árvores de decisão. O termo aleatório deve-se a esta variação entre as árvores. Chega-se ao resultado avaliando todas as decisões, sendo assim um processo democrático e em conjunto (LIAW; WIENER et al., 2002). Para avaliar as emoções, tem-se as saídas do algoritmo como as emoções que apresentarem melhores resultados nas respostas das árvores.

Existem combinações de ideias arquitetônicas importantes que reforçam a robustez da rede. Como os pesos que são compartilhados entre a rede, os campos receptivos locais e a ocorrência de subamostragem espacial ou temporal, quando necessário (LECUN; BENGIO et al., 1995).

2.3.5 Adaptação e uso do reconhecimento facial

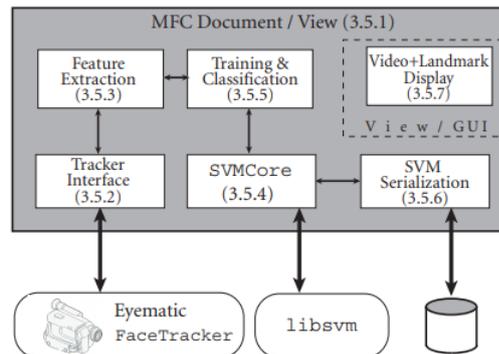
Para a criação das respostas emocionais a partir da face utilizou-se dois métodos gerados a partir de classificadores.

2.3.5.1 Software SVM

O software usado como referência para a aquisição de emoções por SVM apresenta uma abordagem similar ao software utilizado por Michel (2003). A implementação usa a biblioteca de classificadores *LIBSVM*, orientada a objeto, para que tenha uma forma incremental de trabalho. Esta forma de uso permite o re-treino como também o uso em diferentes aplicações. No caso deste estudo analisou-se o classificador para um único sujeito e para grupo de sujeitos.

Para melhor compreensão da estrutura do classificador usou-se o fluxo de dados apresentado na Figura 19.

Figura 19 – Fluxo de classificação de SVM



Fonte: (MICHEL, 2003)

Como pode-se observar, existe o pré-processamento em que a SVM gera um rastreamento dos marcos faciais em sequências e são estes os dados que são cruzados na *SVM RBF Kernel*⁴. Com o treinamento da classificação criou-se o *SVMcore*, que é o arquivo onde estão armazenados os parâmetros, os dados treinados e os resultados. Tendo este *SVMcore*, entra-se na etapa do preditor dos dados futuros que neste fluxograma é chamado de *SVM Serialization*. Para este classificador obteve-se, segundo Michel (2003), uma acurácia acima de 62% para todos as emoções, principalmente para a emoção de felicidade que atingiu 91% de acurácia, como apresentado pela Tabela 1.

⁴ Função de base radial (RBF).

Tabela 1 – Tabela de resultados da matriz de confusão

	Anger	Disgust	Fear	Joy	Sorrow	Surprise	Overall
Anger	8	0	2	0	2	0	66.7%
Disgust	3	9	2	0	0	0	64.3%
Fear	2	0	8	0	2	0	66.7%
Joy	0	1	0	11	0	0	91.7%
Sorrow	0	2	4	0	10	0	62.5%
Surprise	2	0	0	0	0	10	83.3%
Total accuracy: 71.8%							

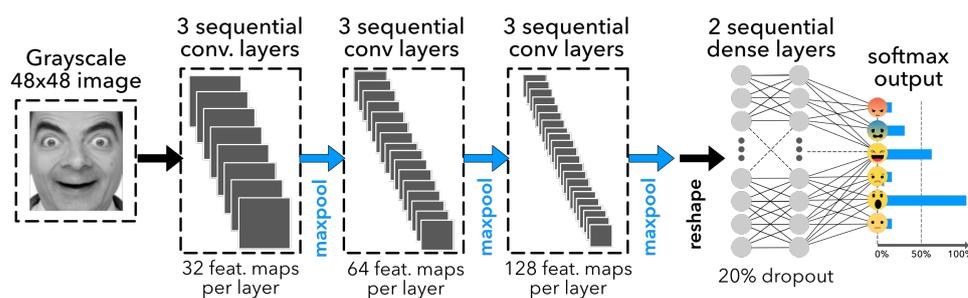
Fonte: (MICHEL, 2003)

Como demonstrado na Tabela 1, este classificador apresentou uma acurácia total de 71,8% para a classificação das 6 emoções. Foi observada uma maior dificuldade em prever a emoção triste.

2.3.5.2 Software CNN

Na plataforma CNN, cada camada da convolução é formada por um número específico de filtros (HO, 2016). Estes conjuntos de filtros (*kernel*) deslizam pela imagem, formando os mapas de características. Usa-se uma sequência de camadas de convolução e uma camada de rede neural densa, para encontrar a resposta emocional. A Figura 20 demonstra estas etapas do processamento da rede.

Figura 20 – Fluxo de classificação de um CNN

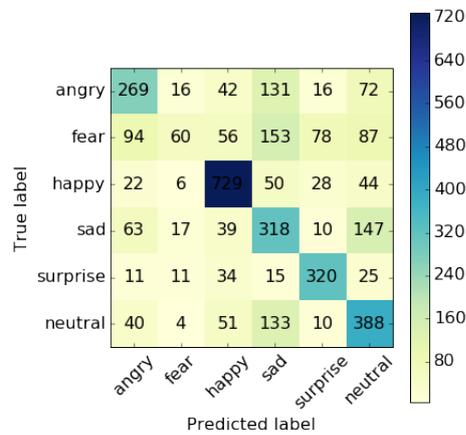


Fonte: (HO, 2016)

Como visto, o software apresenta uma boa estimativa dos resultados emo-

cionais. A matriz de confusão, apresentada na Figura 21 abrange estes resultados.

Figura 21 – Matriz de confusão do CNN



Fonte: (HO, 2016)

Analisando a matriz de confusão da rede, observa-se que as emoções feliz, surpreso e neutro obtiveram bom grau de acurácia. Já para as demais categorias, o grau fica um pouco abaixo. Encontrando-se uma exatidão global de cerca de 58% para o classificador.

2.4 Estado-da-arte

Há estudos atuais que utilizam comportamento facial, questionários e bases visuais de reconhecimento facial para ambientes de aprendizagem, como em Harley, Bouchet e Azevedo (2013). Este estudo analisou a ocorrência de variação nas expressões faciais de alunos enquanto aprendiam com um tutor. O estudo foi capaz de analisar variações de expressões positivas e neutras, obtendo concordância entre os métodos de 75,6%.

Wiklund et al. (2015), mais recentemente, utiliza o software de reconhecimento para analisar as respostas ao jogos de educação e de entretenimento. Neste estudo foram utilizados dois métodos em paralelo para inferir dados de emoções, com EMG e com resposta visual, para comparar os métodos entre si. O estudo

observa que jogos de educação apresentam um aumento significativo nas ativações de emoções negativas. Pode-se usar este teste para avaliar a qualidade do jogo educacional.

O estudo de [Busso et al. \(2004\)](#), que analisa a resposta facial e auditiva, revela que o sistema baseado na expressão facial tem melhor desempenho do que o sistema baseado apenas na resposta auditiva para o reconhecimento de emoções. Outro ponto importante deste estudo demonstra que a união de dois métodos causa um aumento no desempenho e na robustez dos sistemas de reconhecimento de emoção. Neste estudo, chegou-se a índices de emoções próximos a 84% para as emoções mensuradas.

Em relação ao uso de fNIRS, [Whalen et al. \(1998\)](#) comprova, com imagens geradas por exames de Ressonância Magnética Funcional, que há atividades em diferentes regiões do cérebro suscetíveis à resposta de estímulos emocionais. Mostrando que os locais com maior intensidade, nas áreas de interesse, podem ser segmentados como uma região, criando a abordagem para estimar a resposta emocional de cada indivíduo a um estímulo específico.

Há estudos onde as diferenças significativas na frequência cardíaca classificam o jogo adaptado ([TIJS; BROKKEN; IJSSELSTEIJN, 2008](#)). Neste exemplo, o estudo analisou a resposta emocional relacionada a velocidade do jogo. O estudo observou as emoções por cinco diferentes análises fisiológicas: pulsação, Eletromiografia Facial, respiração, condutância da pele e pressão do teclado. O estudo observou que muitos destes canais variavam não unicamente pelas emoções, o que necessitou um estudo mais aprofundado. Segundo [Vachiratamporn et al. \(2015\)](#), que realizou estudos relacionando emoções e frequência cardíaca, os estados neutros e de diferentes níveis de medo estão diretamente relacionados a resposta negativa a um jogos de terror.

Outro estudo, avalia os diferentes níveis de dificuldade do jogo Tetris, relacionados a emoções distintas. São avaliada as diferenças fisiológicas do jogador usando sinais de eletroencefalograma (EEG) e questionários. Estes dados alimentam classificadores para avaliar as emoções ligadas a cada um dos níveis de dificuldade. O estudo observou que após a fusão das duas categorias de sinal a precisão

aumentou para 63% (CHANEL et al., 2011).

Há outras formas de medir a resposta à emoções, como por exemplo o que foi feito por Herrmann, Ehlis e Fallgatter (2003), que usando um fNIRS de 3 canais, observou ativação, na área pré-frontal do cérebro, evocadas por imagens específicas. Portanto, estes estudos comprovam a existência de indícios de ativações diferentes entre imagens positivas e negativas, que corrobora para o aprofundamento destes estudos de emoções positivas ou negativas.

O comportamento do corpo como resposta emocional apresenta diversos indícios de emoções por vezes superficiais, observado mais facilmente para os casos de desinteresse, tédio, desapresso, desatenção e sentimentos ligados ao foco e a vontade do jogador. Alguns jogos sérios usam a posição e a movimentação do usuário jogando (KLEINSMITH; BIANCHI-BERTHOUE; STEED, 2011). A movimentação dos braços e da cadeira demonstram desinteresse, e estes movimentos são classificados como sentimentos específicos (BIANCHI-BERTHOUE, 2013). O estudo evidencia o interesse na atividade e não as emoções, mas não se descarta o grande potencial relacionado a o uso destes tipos de métodos, aliado aos emocionais.

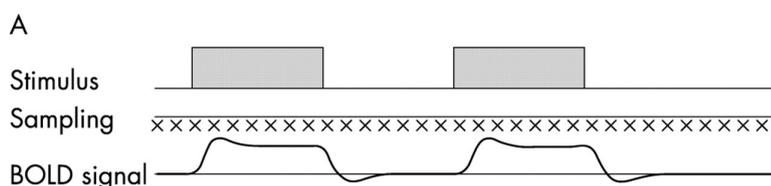
3 METODOLOGIA

3.1 Desenvolvimento do protocolo de evocação

Quando se busca respostas corporais, deve-se estimular e observar as áreas de interesse que demonstrem o que se busca como respostas. No caso dos experimentos em *fNIRS*, sempre tem-se como objetivo medir uma região do cérebro. Por exemplo, a região motora do cérebro incrementará sua atividade caso o voluntário mexa a mão ou a perna. Já para o caso das expressões faciais, a face reage aos estímulos, moldando suas micro expressões de uma forma específica. Como demonstrado no referencial teórico, um conjunto de micro expressões formam a resposta emocional correspondente. As principais formas de estímulos são conhecidas como evento em bloco e evento relacionado. A diferença básica entre os dois, reside no fato de que o evento em bloco apresenta diversos estímulos, causando um patamar de sinal *bold* ou uma contínua expressão facial. Já o evento relacionado, gera um único estímulo e logo após é verificada a resposta da atividade (HUETTEL et al., 2004).

O experimento em bloco caracteriza-se por uma sequência de estímulos seguidos de breves pausas. Portanto, o evento em bloco demonstra um número de estímulos de um mesmo critério, como apresentado na Figura 22. O experimento em bloco normalmente causa um incremento maior na atividade cerebral devido ao fato de que o estímulo é mantido por um tempo maior.

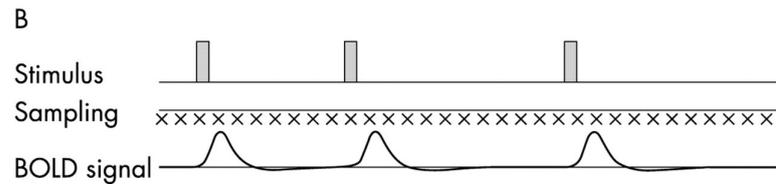
Figura 22 – Estímulos em blocos.



Fonte: (MATTHEWS; JEZZARD, 2004)

Outra forma de estimular é com eventos relacionados, como apresentado na Figura 23, onde o estímulo é único para cada critério e permanece por um tempo específico para provocá-lo. Após o estímulo, da mesma forma que ocorre em blocos, usa-se uma breve pausa para estabilizar os sinais e realizar outro teste.

Figura 23 – Estímulos Independentes.



Fonte: (MATTHEWS; JEZZARD, 2004)

Para este estudo foi utilizado estímulos em evento, mas este evento será mantido por um período de 10s para cada estímulo.

3.1.1 Escolha do método de estímulos

O teste consiste na apresentação de diversas imagens com caráter positivas, negativas e neutras aos voluntários. Estes voluntários estarão suscetíveis a demonstração em caráter experimental de suas emoções em relação a cada imagem. As imagens usadas no estudo foram retiradas do banco mundial de imagens *International Affective Picture System* (IAPS) (<http://csea.php.ufl.edu/downloadiaps.html>). Estas imagens já são amplamente utilizadas para pesquisas (LANG; BRADLEY, 2007). Apresentamos na Figura 24 um exemplo de imagem similar ao encontrado no banco de imagens.

Figura 24 – Imagens similares as de evocação.



Fonte: Imagens meramente ilustrativas retiradas da internet

As imagens escolhidas têm por princípio forçar o voluntário a expressar sentimentos. Portanto, a escolha correta das imagens foi primordial para que os estímulos sejam adequados ao estudo. Para essas escolhas se utilizou os índices fornecidos pelo banco de onde as imagens foram retiradas, classificando-as por estes índices. Para validar a escolha, ao final de cada estímulo o voluntário irá classificar a imagem com sua definição entre os três analisados no teste, evidenciando se a imagem é positiva, negativa ou neutra. Esta resposta é primordial para o estudo das relações das imagens posteriormente.

3.1.2 Método de escolha das imagens de estímulo

O banco da *IAPS* apresenta para o contexto das emoções um total de 1193 imagens. Analisando as imagens presentes, de imediato foi necessário retirar algumas imagens, devido a conteúdo explícito. Portanto todas as imagens de caráter sexual ou extremamente violento foram retirados.

Dentro deste banco, as imagens sofrem classificação em três áreas distintas: *dominância*, *valência* e *excitação*. As imagens são apresentadas em forma de tabela como demonstrado na Figura 25, onde podemos observar, na primeira coluna, uma breve descrição da imagem e, na segunda coluna, o índice numérico que nomeia as figuras dentro do banco. Já as colunas subsequentes apresentam as médias e os desvios padrão de cada uma das áreas de classificação.

Figura 25 – Imagens da apresentação dos dados da IAPS.

Description	Slide No.	Valence Mean(SD)	Arousal Mean(SD)	Dominance1 Mean (SD)	Dominance2 Mean (SD)	Pict Set
Puppies	1710	8.34(1.12)	5.41(2.34)	6.55(1.98)		3
Lion	1720	6.79(1.56)	5.32(1.82)	4.63(2.28)		3
Lion	1721	7.30(1.39)	4.53(2.21)	5.57(2.07)		10
Jaguars	1722	7.04(2.02)	5.22(2.49)	6.12(2.29)		12
Tiger	1726	4.79(2.10)	6.23(2.19)	.	4(2.33)	14
Lion	1731	7.07(1.58)	4.56(2.50)	.	5.85(2.19)	14
Owl	1740	6.91(1.38)	4.27(2.03)	5.85(1.79)		3
Bunnies	1750	8.28(1.07)	4.10(2.31)	6.15(2.01)		3
Hippo	1810	6.52(1.49)	4.45(2.11)	4.55(1.81)		5
Monkeys	1811	7.62(1.59)	5.12(2.25)	6.07(1.96)	5.99(1.95)	7
Elephants	1812	6.83(1.33)	3.60(2.11)	5.91(1.84)		9
Crocodile	1820	5.35(2.05)	5.67(2.09)	4.66(2.30)		19
Camels	1850	6.15(1.52)	4.06(2.14)	5.94(1.91)		12

Fonte: (LANG; BRADLEY, 2008)

Com as imagens adequadas para esta pesquisa, passou-se então a etapa de

escolha das melhores imagens dentre as que se enquadraram na pesquisa. Para esta seleção, criou-se um algoritmo no software *Matlab*, responsável pela classificação das imagens dentro de seus índices de dominância, valência e excitação.

Para a classificação dos índices foi confeccionado um programa que seleciona as imagens dentro do banco. Como todos os dados são apresentados em tabela, optou-se por transformar os dados em matriz. A matriz se nomeou de **Alls**. Dentro desta matriz, a coluna dois apresenta a valência, já a coluna quatro, apresenta a excitação, as demais colunas não serão utilizadas na classificação.

Usou-se 15 imagens consideradas positivas. Os valores dos parâmetros foram estipulados até encontrar-se a quantidade de imagens necessárias. Portanto usou-se valência maior que 7.5, e dominância maior que 4, como apresentado no algoritmo 1.

```

1 Igual = 0;
2 As = zeros(1994,1);
3 K = 1;
4 for j = 1:1994
5     if Alls(i,1) == As(j,1)
6         Igual = Igual+1;
7     end
8     if Alls(i,2)>7.5 & Alls(i,4)>4;
9         As(k,1) = Alls(i,1);
10        AVN_String(k,1)=Strcat(C:\',num2str(Alls(i,1),'.jpg'));
11        movefile(AVN_String(k),'C:\ IAPS\ Banco\ Positivo');
12        AVP(k,1) = Alls(i,1);
13        AVP(k,2) = Alls(i,2);
14        AVP(k,3) = Alls(i,4);
15        k= k+1;
16    end
17    Igual = 0
18 end

```

Algoritmo 1: Seleção das imagens positivas

O resultado do programa apresentou as 15 imagens selecionadas, as quais, devido ao termo de sigilo, serão tratadas pelos seus índices e apresentadas em forma de tabela, como mostrado na Tabela 2. Temos a valência média representada por Val_{mn} , a excitação por $arom_{mn}$.

Tabela 2 – Imagens selecionadas como positivas.

Positiva	IAPS	valmn	valsd	aromn	arosd	dom1mn	dom1sd
Seal	1440	8,19	1,53	4,61	2,54	6,05	2,38
Kitten	1460	8,21	1,21	4,31	2,63	6,00	2,00
Puppies	1710	8,34	1,12	5,41	2,34	6,55	1,98
Money	8501	7,91	1,66	6,44	2,29	6,05	2,52
Baby	2040	8,17	1,60	4,64	2,54	7,33	2,54
Baby	2050	8,20	1,31	4,57	2,53	7,71	2,53
Baby	2070	8,17	1,46	4,51	2,74	7,14	2,10
Babies	2080	8,09	1,47	4,70	2,59	7,08	1,96
Family	2154	8,03	1,13	4,48	2,31	6,38	1,90
NeutBaby	2260	8,06	1,42	4,26	2,44	7,47	2,44
Family	2340	8,03	1,26	4,90	2,20	6,18	1,86
Beach	5833	8,22	1,08	5,71	2,66	6,97	1,93
Seaside	5210	8,03	1,09	4,60	2,48	6,19	2,08
Sea	5825	8,03	1,18	5,46	2,72	6,61	2,11
Skier	8190	8,10	1,39	6,28	2,57	6,14	2,74

Fonte: O Autor (2018)

Pode-se observar, pela descrição das imagens positivas selecionadas, a existência de pontos coincidentes entre algumas das imagens. Estas coincidências formam grupos de imagens, como por exemplo, imagens de bebês, animais ou paisagens paradisíacas. Já analisando os grupos de imagens, observa-se que esta similaridade não está mais presente, o que demonstra que as amostras são diferentes tendo como único ponto em comum, o fato de provocarem emoções positivas.

Na escolha das imagens neutras, o algoritmo utilizado é muito semelhante ao Algoritmo 1, trocando-se apenas parâmetros de valência, entre 4 e 6.5, e dominância menor que 3. As 15 imagens selecionadas para emoções neutras tem seus parâmetros apresentados na Tabela 3.

Tabela 3 – Imagens selecionadas como neutras.

Neutra	IAPS	valmn	valsd	aromn	arosd	dom1mn	dom1sd
Man	2190	4,83	1,28	2,41	1,80	5,92	2,01
Chess	2840	4,91	1,52	2,43	1,82	5,56	1,93
Rocks	5130	4,45	1,13	2,51	1,72	5,84	1,98
Plant	5740	5,21	1,38	2,59	1,99	6,27	2,21
Hammer	7110	4,55	0,93	2,27	1,70	6,07	1,86
RollingPin	7000	5,00	0,84	2,42	1,79	6,14	2,14
Spoon	7004	5,04	0,60	2,00	1,66	6,74	1,99
Bowl	7006	4,88	0,99	2,33	1,67	6,18	1,96
Basket	7010	4,94	1,07	1,76	1,48	6,70	1,48
Fork	7080	5,27	1,09	2,32	1,84	7,04	1,84
Lamp	7175	4,87	1,00	1,72	1,26	6,47	2,04
AbstractArt	7187	5,07	1,02	2,30	1,75	6,10	2,04
ClothesRack	7217	4,82	0,99	2,43	1,64	6,25	1,86
Window	7490	5,52	1,41	2,42	2,23	5,81	2,10
Tissue	7950	4,94	1,21	2,28	1,81	6,30	2,11

Fonte: O Autor (2018)

Diferente do que se observou no caso das imagens positivas, as imagens neutras apresentam quase que como um todo, um padrão de imagens de objetos, e com pouca frequência imagens de pessoas neutras. Neste caso, a resposta deverá ser mais homogênea e se intensificar quando do aparecimento de face de pessoas.

Na escolha das imagens negativas, teve-se novamente a utilização do algoritmo modificando-se a valência para menor que 3, e dominância para maior que 5. Este algoritmo resultou nas 15 imagens selecionadas para as emoções negativas, como mostrado abaixo na Tabela 4.

Tabela 4 – Imagens selecionadas como negativas.

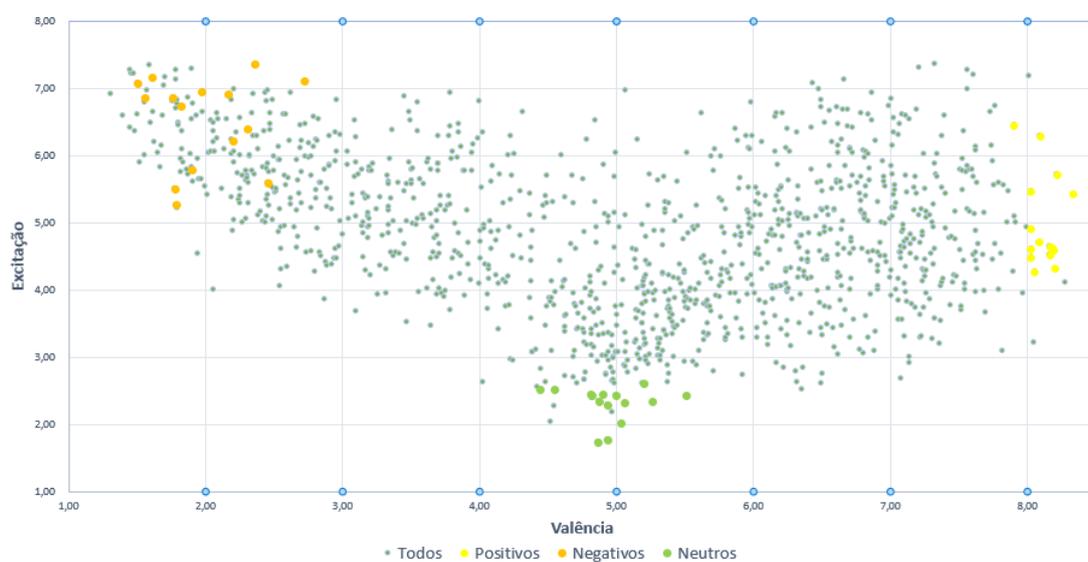
Negativas	IAPS	valmn	valsd	aromn	arosd	dom1mn	dom1sd
Toddler	2095	1,79	1,18	5,25	2,34	3,70	2,28
Gun	2811	2,17	1,38	6,90	2,22	2,55	2,21
Attack	6313	1,98	1,38	6,94	2,23	3,26	2,74
DeadMan	9412	1,83	1,37	6,72	2,07	3,00	2,32
CarAccident	9900	2,46	1,39	5,58	2,13	3,63	2,05
SadChildren	2703	1,91	1,26	5,78	2,25	3,15	1,95
DeadBody	3120	1,56	1,09	6,84	2,36	3,32	2,36
AimedGun	6230	2,37	1,57	7,35	2,01	2,15	2,09
BeatenFem	6315	2,31	1,69	6,38	2,39	3,79	2,69
Explosion	9940	1,62	1,20	7,15	2,24	2,45	2,22
SadChild	2800	1,78	1,14	5,49	2,11	3,40	2,32
DeadTiger	6415	2,21	1,51	6,20	2,31	3,79	2,47
Attack	6550	2,73	2,38	7,09	1,98	3,01	2,41
Attack	6563	1,77	1,23	6,85	2,18	2,93	2,12
Soldier	9410	1,51	1,15	7,07	2,06	2,81	1,99

Fonte: O Autor (2018)

As imagens negativas, como as positivas, apresentaram grupos de imagens com único ponto em comum o fato de provocarem emoções negativas.

Como pode-se ver utilizando os algoritmos construídos, pôde-se selecionar as 15 melhores imagens com as condições específicas para o experimento. A Figura 26 demonstra que as imagens estão bem selecionadas tendo em vista a posição que se encontram em relação ao banco total de imagens.

Figura 26 – Gráfico de todos as amostras.



Fonte: O Autor (2018)

Como esperado, os resultados condizem com o que buscava-se, sendo observado que em cada um dos casos, onde existem imagens melhores que as escolhidas, estas foram retiradas devido ao seu conteúdo inadequado.

3.1.3 Paradigma

Para a implementação das estruturas do paradigma, usou-se o software de NIRX, *NIRStim4.0* (TECHNOLOGIES., 2016b), montando uma estrutura como a representada na Figura 27.

Figura 27 – Fluxo de processo para um estímulo.



Fonte: O Autor (2018)

Como pode-se observar na Figura 27, cada estímulo ficará presente por 10 segundos, depois seguidos de duas perguntas, mostradas na Figura 28, com duração de 5 segundos cada uma. Para que o paciente volte ao estado normal, exibe-se uma cruz por mais 10 segundos, para somente então começar o próximo estímulo.

Figura 28 – Sequência de perguntas.



Fonte: O Autor (2018)

Cada exame usa cinco estímulos por emoção, gerando um tempo total de sete minutos e meio de exame. Cada voluntário passou por três exames, o que leva a 22 minutos e meio, que consiste em um experimento. Há também o tempo para colocação e calibração dos equipamentos.

Segundo o fluxograma dos exames, tem-se as 45 imagens selecionadas anteriormente, que foram novamente subdivididas em 3 exames com 3 emoções cada. Esta separação se deve ao fato de que serão realizados três exames em cada voluntário.

Com a segregação calculou-se as médias amostrais de cada um dos subgrupos para se avaliar a homogeneidade dos exames.

Tabela 5 – Médias para cada exames.

Exame	Positiva		Neutra		Negativa	
	valmn	aromn	valmn	aromn	valmn	aromn
1	8,118	4,568	4,946	2,232	2,046	6,278
2	8,126	4,952	4,824	2,492	1,954	6,7
3	8,112	5,456	5,09	2,162	2	6,54

Fonte: O Autor (2018)

Pode-se observar na Tabela 5, que os critérios emoções, relacionados entre eles, apresentaram índices muito próximos tanto para valência quanto excitação. O que demonstra a homogeneidade dos exames, viabilizando a comparação entre exames.

3.2 Análise hemodinâmica usando fNIRS

Dentro de reconhecimento de traços de emoções, encontra-se a possibilidade de analisar as áreas corticais do cérebro, analisando o organismo inconsciente através da variação da hemoglobina oxigenada. Emoções cerebrais em sua maioria estão ligadas as áreas do hipotálamo e da amígdala. A maioria dos estudos concentram-se nestas áreas para fins de levantamento emocional com o uso de ressonância magnética.

Há por outrora uma gama muito reduzida de estudos que direcionam a perspectiva de encontrar outras áreas de influência, áreas estas, mais superficiais no cérebro. Um grande empecilho a estes estudos relaciona-se às funções normais presentes a estas áreas corticais do cérebro. Ao mesmo tempo em que se monitora uma resposta emocional pode-se monitorar uma outra resposta, como no caso do córtex pre-frontal (PFC), que trata-se da área de processamento, bem como das experiências emocionais.

3.2.1 Área de ativação de emoções corticais

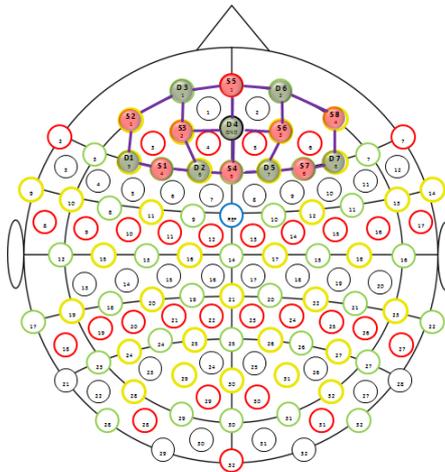
A escolha pela área a ser monitorada durante a estimulação teve como base os estudos que revelam que as principais áreas de ativação, interesse e investigação, encontram-se nas áreas do PFC, como mostrado por [Herrmann, Ehlis e Fallgatter](#)

(2003). O PFC esquerdo é ativado quando há emoções positivas e o PFC direito, emoções negativas. Este estudo comprova o aumento da atividade hemodinâmica na área do PFC, utilizando Tomografia por Emissão de Póstron (*PET*) ou *fNIRS*.

Com a escolha da área de interesse, voltou-se para a questão do equipamento a ser utilizado. Nesta pesquisa, utiliza-se o *NIRScout* de 8×8 canais, da *NIRx*. Com os critérios definidos, fez-se a montagem do mapa de posições dos sensores. As áreas possíveis de colocação dos optodos seguem em muito a utilizada nas toucas de Eletroencefalografia (*EEG*).

Junto ao equipamento temos o software de captura dos dados *NIRStar*. Neste software se introduz os critérios escolhidos para este estudo, que são: número de canais, tipo de análise e posição dos optodos. Para a montagem dos optodos, seguiu-se a distribuição demonstrada na Figura 29. A montagem escolhida é uma das já fornecidas pelo fabricante para análise da área do PFC.

Figura 29 – Mapa de disposições dos optodos.



Fonte: O Autor (2018)

A criação de cada canal consiste em um conjunto emissor-detector espaçados de 3 cm, de acordo com recomendação do fabricante. Emissores (*Source*) representados pela letra "S", e os Detectores (*Detector*) identificados pela letra "D".

Com os canais configurados, o software gera os canais em forma matricial, distribuídos como na Figura 30.

Figura 30 – Matriz de canais.

2 - 3	5 - 3		5 - 6	8 - 6
2 - 1	3 - 3	5 - 4	6 - 6	8 - 7
	3 - 4		6 - 4	
1 - 1	3 - 2	4 - 4	6 - 5	7 - 7
1 - 2	4 - 2		4 - 5	7 - 5

Fonte: O Autor (2018)

Podemos observar os 20 canais gerados na matriz de canais com a formação dos 8×8 emissores-receptores usados no estudo. O sinal coletado de cada canal é responsável por gerar o mapeamento cortical, contendo diversas intensidades e amostras. Utiliza-se o software *NIRStar*, que realiza a captura e o tratamento dos sinais, para gerar uma resposta aos estímulos diretamente nas áreas medidas.

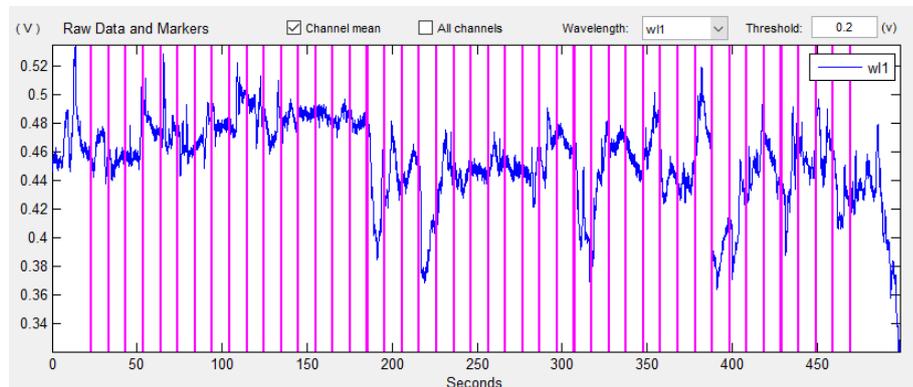
3.2.2 Pre-processamento dos dados de fNIRS

Para a obtenção da variação emocional evocada no cérebro, realizou-se a leitura dos coeficientes de variação da atenuação luminosa. Na estrutura do sinal, grava-se também os coeficientes e os ganhos que fazem parte do sinal global de atenuação. Com esta leitura obteve-se 20 canais, cada um com seus coeficientes e ganhos individuais. O valor dos ganhos se mantém igual ao longo de cada exame, pois são ajustado no momento da calibração da touca.

Para que os dados pudessem ser caracterizados no estudo, a saída de luz recebida pelos receptores foi convertida para uma resposta hemodinâmica, com variação nas áreas específicas de cada canal. Para esta conversão o software se baseia no diagrama de *Bold* seguindo um processo de quatro etapas.

A primeira etapa está relacionada a colocar as características do procedimento no software, usando o comando *Set markers*. Esta característica está ligada ao tempo de cada estímulo, como mostrado na Figura 31, onde pode-se observar o tempo ou *frame* em que foi realizado o trigger do estímulo.

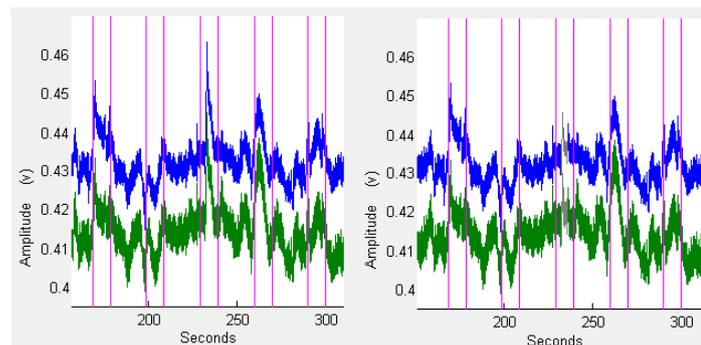
Figura 31 – Definição dos períodos.



Fonte: O Autor (2018)

Como visto no item do paradigma, cada etapa tem uma duração de 10 segundos. As etapas seguem a sequência de um "Baseline" + "Imagem" + "Perguntas" seguindo a ordem aleatória das emoções até completada as 15 repetições. Nesta etapa o dado foi analisado sobre sua conformidade.

Na segunda etapa foram amenizados os ruídos oriundos do movimento do sujeito, chamados na literatura de *Spike*. Estes ruídos são amenizados um a um manualmente pela plataforma no comando *Remove*, como mostrado na Figura 32 onde é apresentado o sinal antes e depois da remoção de um *Spike*.

Figura 32 – Gráfico original e após remover o *Spike*.

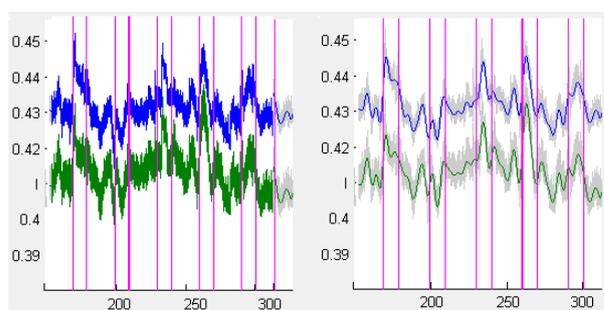
Fonte: O Autor (2018)

Pode-se observar claramente a amenização do sinal de ruído, entre os tem-

pos de (230 e 240) segundos. Esta retirada ajuda na conversão para um sinal hemodinâmico, já que não há tanto ruído externo.

O experimento captura todas as variações sanguíneas e respiratórias do sujeito. Estas variações tem frequências peculiares, como o ritmo cardíaco, e por não condizerem com a análise emocional do sujeito devem ser retiradas. Essas altas frequências são retiradas com o auxílio de um filtro passa-banda, onde todas as frequências acima de $0,2Hz$ e abaixo de $0,1Hz$ sofrem atenuação, restando, assim, em sua maioria, apenas as frequências relacionadas à ativação emocional do cérebro, como apresentado na Figura 33.

Figura 33 – Gráfico original e após passar o filtro.



Fonte: O Autor (2018)

A conversão do sinal de luz para um sinal de resposta hemodinâmica é a principal parte para a resposta com uso do equipamento fNIRS. Nesta etapa é realizada a conversão, onde se cruza os dados de ganho, sinal e coeficientes, formando o sinal de variação hemodinâmica.

O cálculo que redefine os dados gerados pelo equipamento, referentes ao estado hemodinâmico, é feito usando um método adaptado do modelo de *Beer Lambert*. Este modelo é baseado no método de medição da atenuação da luz, pelo efeito da influência das propriedades dos diferentes materiais, no meio em que a luz atravessa. No caso do estudo com fNIRS, o modelo é adaptado sobre algumas características como: comprimento de onda, comprimento do trajeto e valores padrão de oxigenação venosa. Estas características são adicionadas ao modelo no comando *Set parameters* e os valores dos parâmetros seguem os padrões apresentados por

Meyerding e Risius (2018).

Com todas as etapas concluídas e o sinal hemodinâmico gerado, exportou-se os dados no formato TXT. Este arquivo gerado e exportado pelo *NIRSlab* é constituído por 20 canais, com uma taxa de amostragem de 7,81 Hz, gerando 79 amostras em cada estímulo de 10 segundos. Os exames apresentam 15 sequências com 30 segundos cada. Um total para cada exame de 3555 amostras, sendo 395 para cada emoção, como apresentado na Tabela 6.

Tabela 6 – Conjunto de dados gerados por paciente.

Exame	Numero de amostras					
	Positivas	Negativas	Neutras	Baseline	Perguntas	Total
1	395	395	395	1185	1185	3555
2	395	395	395	1185	1185	3555
3	395	395	395	1185	1185	3555
Total	1185	1185	1185	3555	3555	10665

Fonte: O Autor (2018)

As perguntas estão sendo usadas para validar a evocação e o fato do sujeito processar a resposta. Os dados referentes ao período de leitura e respostas do questionário foram retirados, pois não apresentam resposta direta a evocação de uma emoção. Com a retirada das perguntas, cada exame fica com 2370 amostras e um total por paciente de 7110 amostras.

3.2.3 Teste estatístico dos dados fNIRS

O módulo de Análise de Dados do *NIRSlab* utiliza funções presentes no pacote *freeware SPM (Statistical Parametric Mapping)*, que fornecem a possibilidade de gerar estatísticas de séries temporais de estado hemodinâmico.

A análise ocorre de duas formas distintas, como mostrado no menu de seleção na Figura 34. Primeiro, analisa-se o sujeito de modo individual, seguido de uma análise de grupo, onde as respostas dos sujeitos de um mesmo experimento são cruzadas, analisando-se as respostas de forma conjunta.

Figura 34 – Índice de análise pelo NIRSlab



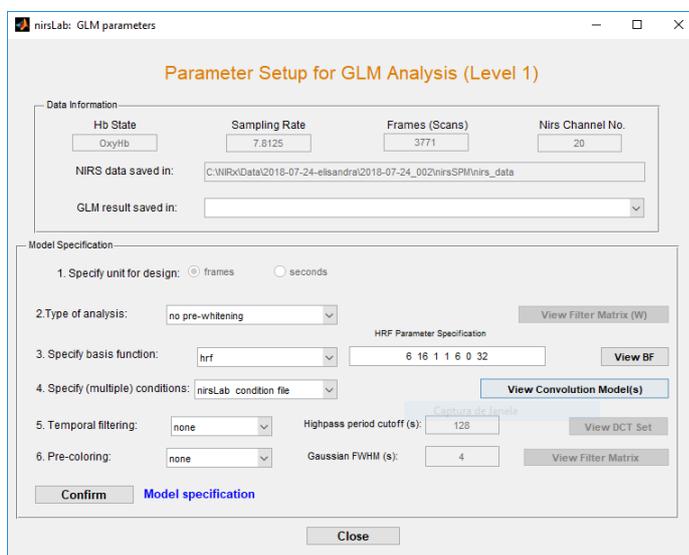
Fonte: O Autor (2018)

A análise dos sujeitos individualmente foi feita no teste de nível um, e os resultados de todos os sujeitos foram aplicados no método dois, como será descrito em seguida.

3.2.3.1 Teste SPM Nível 1

Para a análise estatística individual do sujeito, carrega-se os dados pre-processados. Realizando o teste, teve-se que configurar os critérios estatísticos como apresentado na Figura 35.

Figura 35 – Definindo os parâmetros



Fonte: O Autor (2018)

Teve-se que especificar uma função de base matemática, que modela a res-

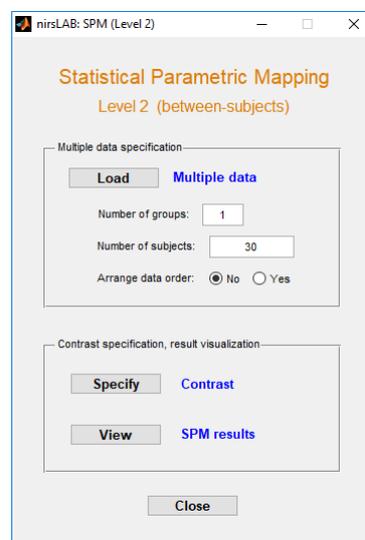
posta fisiológica da hemodinâmica. Como a atividade neural gera resultados de fluxo de sangue oxigenado, tipicamente após 1 a 2 segundos mais tarde, é recomendado que se use a função de resposta hemodinâmica HRF (do inglês *hemodynamic response functions*). Como o experimento utilizou *triggers*, foram selecionadas as condições segundo o *NIRSlab*, como configurado no pre-processamento.

Com os parâmetros configurados, gera-se a estatística no comando *Estimate*. Este comando gera a estatística e as respostas de cada sujeito. Este procedimento foi repetido para cada um dos 10 sujeitos analisados no estudo.

3.2.3.2 Teste SPM level 2

Para a análise de grupos ou múltiplos sujeitos deve-se definir quantos grupos e quantos serão os sujeitos por grupos. Para este estudo usou-se um grupo com 10 sujeitos. Com esta escolha, pode-se avaliar os sujeitos cruzando cada resposta entre sujeitos. Neste item, carregou-se os dados gerados no nível um de todos os sujeitos pelo comando *Load*, como mostrado na Figura 36.

Figura 36 – SPM - level 2



Fonte: O Autor (2018)

Após carregar os dados de todos os 10 sujeitos, criou-se o contraste que o

estudo deve usar para comparar os sujeitos. O tipo de contraste sugere a hipótese biológica que se está usando para a resposta emocional. Para este contraste, utilizou-se a comparação entre cada uma das respostas.

Pode-se observar, na Figura 37, a construção de cada contraste como descrito acima. Esta etapa é a principal, pois é onde definiu-se o que foi comparado. Nesta montagem, tem-se a comparação da resposta negativa com a resposta positiva.

Figura 37 – Contraste para comparações de fNIRS



Fonte: O Autor (2018)

Esta combinação de condições produz uma resposta hemodinâmica para cada emoção individualmente.

1. Condição de contraste para emoções negativas

Para emoções negativas temos o contraste $[10001]$, consideradas as ativações negativas (coluna um), versus a não ativação (coluna cinco).

2. Condição de contraste para emoções positivas

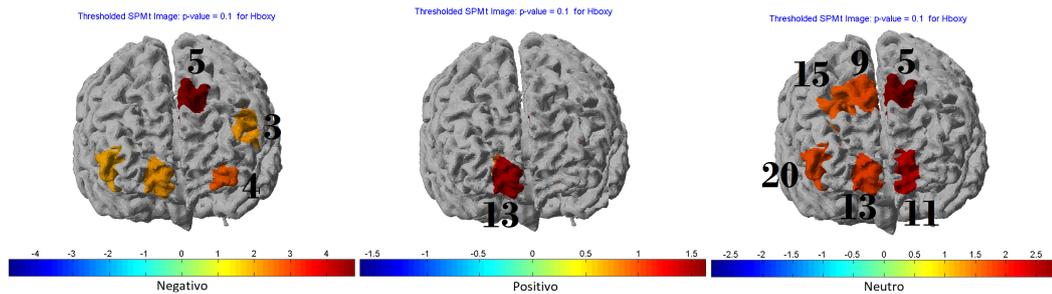
Para emoções positivas temos o contraste $[01001]$, consideradas as ativações negativas (coluna dois), versus a não ativação (coluna cinco).

3. Condição de contraste para emoções neutras

Para emoções positivas temos o contraste $[00101]$, considerada as ativações negativas (coluna três), versus a não ativação (coluna cinco).

Com a correta escolha das condições de contraste, gera-se os resultados do conjunto dos 10 sujeitos, como mostrado na Figura 38, onde apresenta-se uma resposta de ativação cerebral para cada estímulo emocional e quais canais para cada emoção.

Figura 38 – Respostas do dados fNIRS



Fonte: O Autor (2018)

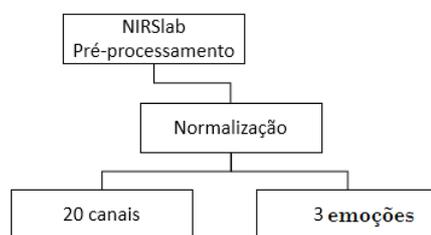
Pode-se observar, na Figura 38, as ativações de cada emoção. No entanto, é importante a definição específica das áreas que realmente sofrem ativação evocadas pelas emoções.

3.2.4 Normalização dos dados

Os dados gerados pelo software *NIRStar* tiveram que ser pre-processados no software *NIRSlab*, para estarem de acordo com os sinais de *triggers*. Como são 20 canais no experimento analisando toda a área pré-frontal, foram subdivididos os dados em dois grupos de interesse, os do lado esquerdo para emoções positivas e os do lado direito para emoções negativas. Para a análise das emoções neutras tem-se ativação em ambos os lados.

Os canais que tiveram ativação para emoção negativa foram os canais 3, 4 e 5 enquanto que para a emoção positiva somente o canal 13 sofreu ativação. Já as emoções neutras foram percebidas nos canais 9, 11, 15 e 20.

Figura 39 – Fluxograma de aquisição do fNIRS



Fonte: O Autor (2018)

A escolha dos canais pode ficar mais clara analisando a Figura 29. Em virtude do experimento avaliar três critérios distintos, os canais serão agrupados para cada uma das emoções, formando as três respostas esperadas. Neste momento, tem-se os dados gerados pelo equipamento em duas formas, como apresentado no fluxograma com as etapas na Figura 39.

3.3 Análise de emoções pelo reconhecimento facial

A análise emocional da face se baseia nas regiões como boca, olhos, sobrancelhas e micro expressões, como covas e marcas no rosto. O software escolhido classifica estas micro expressões e analisa quais as possíveis características encontradas para cada uma das expressões. Com os padrões da face, cria-se um banco de dados onde cada imagem apresenta uma resposta de emoção.

O banco de dados com grande número de amostras de treino, gera melhores índices de aprendizado de padrões e cria uma função de teste em que as entradas são expressões, e a saída é a emoção ativada. No caso de vídeos, a análise procede-se *frame a frame*, portanto necessita de um poder de processamento rápido o que é alcançado a partir de ferramentas de *Deep Learnig* e *Machine Learnig*.

Com o intuito de utilizar-se plataformas de reconhecimento de emoções faciais, realizou-se um estudo sobre a construção de ambas as plataformas abordadas e suas diferenças. Observou-se que a principal diferença é o método de construção dos detectores de emoção, que utilizam redes neurais convolucionais (CNN) ou máquinas de vetores de suporte (SVM).

3.3.1 Gerando índices de confiabilidade

Como as plataformas apresentam resultados diferentes, e diversas vezes contrários, analisando a mesma expressão facial, necessitou-se realizar uma análise das características dos resultados. A análise de processamentos das imagens observou que cada método tem suas vantagens e suas desvantagens, oriundos de sua construção.

O método CNN apresenta resultados piores para as emoções como um todo, obtendo, por exemplo, dificuldade na identificação da expressão de medo. Já o método SVM apresenta um bom resultado relacionado à emoção feliz. E em ambos os casos apresenta-se bem definidas as emoções felicidade, triste e surpresa. Em virtude de analisar estas variabilidades, analisou-se as respostas a partir de um banco de vídeos padronizados. Avalia-se a resposta com o uso de uma matriz de erros e acertos de cada emoção. Pode-se avaliar tanto a questão do qual classificador deve-se implementar para a resposta final do estudo, mas, também comparar as respostas entre os métodos, avaliando assim os rendimentos e as variabilidades dos métodos.

Para validar os resultados, utilizou-se o banco de dados padronizado *MMI Facial Expression*. Este banco está em desenvolvimento e visa fornecer volumes de dados, em formato de vídeo e imagens, para expressões faciais.

O banco contém gravações do padrão temporal completo de uma expressão facial: do neutro, ao início, ápice e deslocamento de volta ao neutro. O banco de dados é formado por mais de 2900 vídeos e imagens de alta resolução de 75 sujeitos. O banco de dados é fornecido gratuitamente para a comunidade científica ([VALSTAR](#); [MAAT](#); [GAIA, 2005](#)).

Neste estudo, foram usados vídeos de expressões emocionais, gerando mais de 2000 amostras emocionais, classificadas entre: feliz, triste, neutro, zangado, medo e surpreso. O banco não apresenta padrões de vídeos de expressão de medo, o que acarretou em não ser analisada a emoção medo.

3.3.1.1 Avaliando os melhores classificadores

Utilizou-se funções de classificação da ferramenta *Matlab*, para analisar as plataformas a serem utilizadas. Os classificadores usados foram: SVM e RF (*Random Florest*). A análise se procedeu em três etapas: avaliando a plataforma CNN, avaliando a plataforma SVM e a terceira etapa avaliando ambas as plataformas em conjunto. Para estas avaliações, utilizou-se os resultados do banco padronizado mencionado no item anterior.

1. Avaliando a plataforma CNN.

Para avaliar as 6 entradas emocionais do estudo usando CNN, utilizou-se os classificadores citados anteriormente avaliando as respostas individualmente. Com o uso da ferramenta *Learning classifier* na plataforma do *Matlab*, executou-se testes com classificadores, como mostrado no Tabela 7.

Tabela 7 – Classificadores da CNN pelo *Matlab*.

Modelo	Acurácia
Tree -Arvore aleatória	95.5%
SVM	88.7%

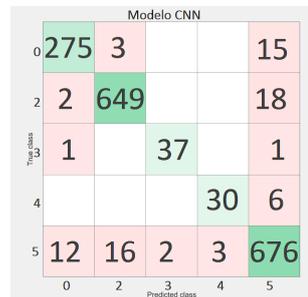
Fonte: O Autor (2018)

Pode-se observar que o melhor resultado encontrado foi 95,5% com o classificador RF. Para o classificador com o melhor rendimento, obteve-se a matriz de erros, para que se tivesse uma análise por emoções, como apresentada na Figura 40.

Pode-se observar pela matriz de confusão, que a diagonal principal concentra a maioria das respostas, o que demonstra que o classificador teve sucesso para todas as emoções testadas.

2. Avaliando modelo SVM. Para avaliar as 6 entradas emocionais do estudo usando SVM, utilizou-se novamente os classificadores citados anteriormente avaliando as respostas individualmente. Contudo, a plataforma SVM não realiza análise de emoção neutra, que não poderá ser analisada. As análises dos classificadores, RF e SVM, geraram os resultados mostrados na Figura 8.

Figura 40 – Matriz erros do classificador floresta aleatória.



Fonte: O Autor (2018)

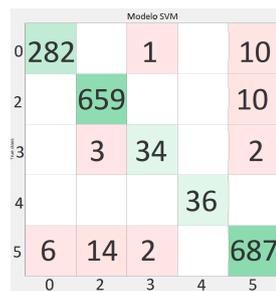
Tabela 8 – Classificadores da SVM pelo *Matlab*.

Modelo	Acurácia
Tree -Arvore aleatória	97.3%
SVM	97.1%

Fonte: O Autor (2018)

O melhor resultado encontrado pertence ao classificador RF, que apresentou acurácia de 97,3%, nos mesmos moldes da rede CNN. Usando o classificador com o melhor resultado, obteve-se a matriz de erros, para que se tivesse uma análise por emoções, como apresentada na Figura 41.

Figura 41 – Matriz erros do classificador floresta aleatória.



Fonte: O Autor (2018)

Como já se esperava, obteve-se resultados semelhantes aos gerados na CNN, usando o mesmo método para o classificador SVM.

3. Avaliando modelo CNN e SVM juntos.

Como o objetivo do estudo é a integração dos métodos de leitura da emoção pela face, foi realizado um classificador nos mesmos moldes do anterior, mas agora com a entrada de ambos os métodos. Tem-se, portanto, 12 canais de entrada para serem introduzidos no classificador. Seguindo o mesmo modelo anterior de análise, apresenta-se o resultados de cada classificador na Tabela 9.

Tabela 9 – Classificadores de ambos os métodos pelo *Matlab*.

Modelo	Acurácia
Tree -Arvore aleatória	98.1%
SVM	98.3%

Fonte: O Autor (2018)

Neste caso, o melhor resultado encontrado pertence ao classificador usando SVM. Este classificador apresentou acurácia de 98,3%. A partir deste classificador com o melhor rendimento, obteve-se a matriz de erros apresentada na Figura 42.

Figura 42 – Matriz erros do classificador SVM

Modelo SVMCNN					
	0	2	3	4	5
0	289				4
2		648			21
3			39		
4				36	
5	2	2			705

Fonte: O Autor (2018)

Como já se esperava encontrar, a união dos métodos melhorou a rede e elucidou qual o melhor modelo de classificador a ser utilizado na etapa final do estudo.

3.4 Voluntários e procedimentos

Por se tratar de um exame não-invasivo e que permite a utilização de voluntários saudáveis, foram selecionados 10 voluntários, 5 homens e 5 mulheres, estudantes da PUCRS. O início do recrutamento só ocorreu após o aprovação de comitê de ética da PUCRS, que ocorreu no dia 21 de Maio de 2018 onde o comitê validou parecer de número: 2.665.006, que consta em anexo.

3.4.1 Descrição dos voluntários

Os voluntários utilizados são caracterizados da seguinte maneira:

- Universitários;
- Saudável;
- Idade entre 20 e 30 anos;
- Ambos os sexos;
- Ativos.

3.4.2 Testes preliminares

Inicialmente utilizou-se um voluntário para testes, onde avaliou-se cinco critérios básicos do experimento: resposta ao estímulo, análise da montagem dos optodos, resposta da câmera, resposta dos métodos e a análise da relação dos dados gerados. A resposta gerada por este voluntário não foi integrada ao aprendizado das redes, por se tratar de teste preliminar.

3.4.3 Descrição do procedimento

Cada voluntário foi convidado a comparecer ao *INSCER* (Instituto do Cérebro), para a realização dos exames. Ao chegar foi lido o termo de consentimento livre e esclarecido e explicado o procedimento, bem como a proposta do estudo. Com o termo assinado o voluntário foi direcionado a sala de experimento, onde foi colocada a touca do *NIRx*. Antes do início, o voluntário passa por uma simulação para compreender bem o experimento.

Após a etapa de simulação, o equipamento é ligado e calibrado. Também é ligada a câmera utilizando o software de reconhecimento construído neste estudo, rodando ambos os processos simultaneamente:

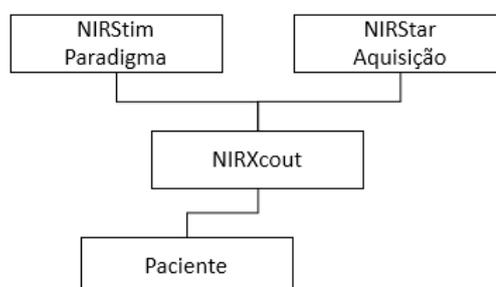
- *NIRStar*: Captura dos dados de *fNIRS*
- *FaceRecognizer*: Capturando dados da face.
- *NIRStim*: Gerando estímulos e *Trigger* de sincronia.

O voluntário passou por três sequências de experimentos, gerando certa quantidade de dados das três emoções distintas. Todos os dados foram armazenados, para uso posterior no treinamentos da rede.

3.4.4 Preparação do experimento usando fNIRS

Para o experimento, montou-se a touca segundo a sequência apresentada na Figura 29, onde utilizaram-se sete detectores e oito emissores, formando os 20 canais. A aquisição dos dados segue o fluxograma apresentado na Figura 43

Figura 43 – Fluxograma de aquisição do fNIRS.



Fonte: O Autor (2018)

Como descrito anteriormente no estudo, o software *NIRStim* aplica o paradigma criado, gerando os sinais de *Trigger* para o *NIRScout*. Em paralelo, o software *NIRStar* coleta os dados do equipamento *NIRScout*. E o equipamento, por sua vez, gerencia a coleta e o armazenamento dos dados pelo processo descrito no referencial teórico.

3.4.5 Preparação do experimento usando webcam

Para uma boa análise facial da câmera do computador, após a preparação do fNIRS, ajustou-se a distância e inclinação da tela e a calibração da aquisição, portanto é pedido ao voluntário que expresse emoções faciais forçadas enquanto realizam-se os ajustes.

3.5 Pre-processamento e adaptação de software de reconhecimento emocional

Para a alimentação dos classificadores, os dados brutos necessitam um pré-processamento. Uma parte deste processamento foi realizado na seção anterior onde os dados foram pré-processados para serem convertidos nas variáveis de interesse.

Na montagem da matriz de dados, defini-se quais são as classes ou dados de entrada do classificador, e quais os rótulos ou saídas do preditor. Neste estudo, as classe são formadas pelos 8 canais gerados pelo fNIRS e os 12 canais gerados pelos reconhecimentos de face, totalizando 20 entradas no sistema.

Já os rótulos são as saídas esperadas do sistema, representadas por positivo, negativo e neutro. Como a saída do classificador deve ser rotulada numericamente, ocorreu a conversão dos nomes para números, como apresentado na Equação 3.1.

$$Saída = \begin{cases} 1, & \text{Negativo} \\ 2, & \text{Positivo} \\ 3, & \text{Neutro} \\ 4, & \text{Baseline} \end{cases} \quad (3.1)$$

O conjunto de dados é separado em subconjuntos (treinamento e de teste). Com ele, é gerada a taxa de predições, acerto versus erro, aferida pelo classificador. Com isso os dados gerados no presente estudo foram separados de acordo com os modelos utilizados.

3.5.1 Pre-processamento - União das plataformas de dados

Para a correta união dos dados de *fNIRS* com os da câmera, teve-se que ajustar a frequência de amostragem e remover o ruído do tempo de resposta na composição dos dados.

1. Ajuste das frequências de amostragem

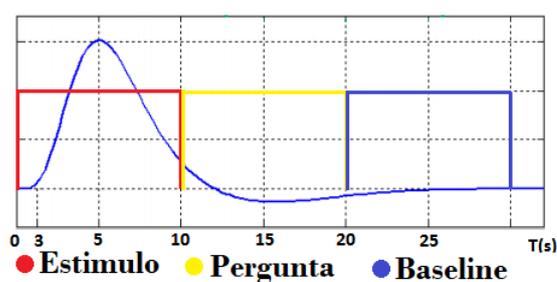
A frequência de amostragem do *fNIRS* é o dobro da encontrada na câmera. Para solucionar esta disparidade, a cada leitura do *fNIRS* duplicou-se a leitura da câmera. Com isso todas as amostras tem agora o mesmo tamanho, formando 79 amostras para cada estímulo de 10 segundos.

O ajuste tem por princípio movimentar matrizes e a cada captura de um dado da matriz de *fNIRS*, o sistema copia duas vezes a posição equivalente no tempo da matriz *câmera*. Este procedimento torna todas as matrizes homogêneas, pois os tempos estão equivalentes e o ajuste dos valores não afeta a qualidade do sinal.

2. Ajuste temporal por causa da resposta hemodinâmica

A resposta hemodinâmica do organismo ocorre após três segundos do início do estímulo, como demonstrado na Figura 44.

Figura 44 – Resposta hemodinâmica corporal (BOLD).



Fonte: O Autor (2018)

Como a saída inicia-se diretamente quando aplicado o estímulo, esta diferença de três segundos gera atraso de resposta, pois a saída apresenta a ativação, mas devido a ser uma resposta hemodinâmica ela ainda não gerou resposta

bold. Para este ajuste de dados, implementou-se um programa que realiza a retirada das primeiras 24 leituras que representam os primeiros três segundos, fazendo com que a matriz tenha 55 amostras para cada sequência de 10 segundos.

Retirou-se, também, os dados referentes à *baseline*. Estas tinham como objetivo que o paciente estabilizasse novamente suas respostas hemodinâmicas, como pode-se ver na Figura 44, antes de retornar com um novo estímulo. A *baseline* foi retirada pois não é utilizada no classificador.

3.5.2 Classificador pelo Matlab

Os métodos de classificação mais utilizados foram apresentados na fundamentação teórica, onde se observou que cada classificador apresenta suas características específicas, gerando respostas diferentes para o mesmo conjunto de dados. Portanto, para analisar qual destes métodos tem mais eficiência para nosso estudo, implementou-se os métodos, e observou-se que o classificador SVM, que apresentou os melhores resultados nos testes pela ferramenta *Matlab*.

3.5.2.1 Ajuste dos conjuntos de dados para uso no Matlab

A implementação dos classificadores foi realizada na ferramenta *Matlab*, utilizando os dados como ajustados anteriormente. Este ajuste é para que o classificador compreenda o que são as classe e o que são os rótulos. Podem-se também definir as classes como preditores e os rótulos como respostas ou saídas.

Os conjuntos de dados foram divididos em três conjuntos, para que pudessem ser analisados separadamente, comparando-os.

1. Conjunto de dados da câmera.

O subconjunto da câmera é formado por 12 canais de entrada e uma saída ligada a resposta emocional.

O número de amostras é variado dependendo da forma que for montado o classificador, se treinarmos em um sujeito, por exemplo, tem-se 2475 amostras, referente ao número de leituras feitas por sujeito.

Tabela 10 – Conjunto de dados das câmeras para treino.

Name	Type	Range	Import as
column_1	double	1 .. 3	Response
column_2	double	-2.33139 .. 3.90579	Predictor
column_3	double	-2.04392 .. 2.56796	Predictor
column_4	double	-0.497899 .. 8.85982	Predictor
column_5	double	-2.14583 .. 1.98951	Predictor
column_6	double	-1.47373 .. 4.34291	Predictor
column_7	double	-1.87648 .. 1.94799	Predictor
column_8	double	-1.30265 .. 2.10919	Predictor
column_9	double	-1.26046 .. 2.18039	Predictor
column_10	double	-2.02136 .. 1.3214	Predictor
column_11	double	-1.18089 .. 2.97331	Predictor
column_12	double	-1.49468 .. 3.38773	Predictor
column_13	double	-1.26027 .. 2.17462	Predictor

Fonte: O Autor (2018)

Teve-se para este subconjunto 13 colunas na matriz de dados, como mostrado na Tabela 10. A matriz foi construída de tal forma que a primeira coluna representa os rótulos e as demais colunas, os preditores ou classes do classificador.

2. Conjunto de dados da fNIRS.

Para o subconjunto de dados do fNIRS, estruturou-se os classificadores com oito canais de entrada, oriundos do fNIRS, e um canal de saída, com a resposta emocional, como apresentado na Tabela 11

Tabela 11 – Conjunto de dados de fNIRS para treino.

Name	Type	Range	Import as
column_1	double	1 .. 3	Response
column_2	double	-3.49994 .. 4.73439	Do not import
column_3	double	-3.18504 .. 3.37175	Do not import
column_4	double	-2.70223 .. 5.22363	Predictor
column_5	double	-2.96157 .. 4.23222	Predictor
column_6	double	-1.83713 .. 4.4977	Predictor
column_7	double	-3.29646 .. 3.4455	Do not import
column_8	double	-2.83332 .. 3.46388	Do not import
column_9	double	-4.76034 .. 4.72629	Do not import
column_10	double	-2.40359 .. 2.38598	Predictor
column_11	double	-2.99473 .. 4.12719	Do not import
column_12	double	-2.31436 .. 4.21372	Predictor
column_13	double	-2.65957 .. 3.02647	Do not import
column_14	double	-2.05287 .. 3.63659	Predictor
column_15	double	-2.44332 .. 3.14071	Do not import
column_16	double	-2.14806 .. 3.64871	Predictor
column_17	double	-2.70491 .. 3.05941	Do not import
column_18	double	-2.75598 .. 3.53188	Do not import
column_19	double	-2.23598 .. 3.64072	Do not import
column_20	double	-1.90102 .. 4.02616	Do not import
column_21	double	-1.94643 .. 4.04231	Predictor

Fonte: O Autor (2018)

Pode-se observar que só oito canais estão selecionados como preditores, por serem só estes os que demonstram emoções. Os demais canais que não geram interesse para este estudo foram desconsiderados.

3. Conjunto de ambos os dados.

Para este estudo, onde ambos os métodos de análise emocional foram utilizados, criou-se um terceiro subconjunto, no qual manteve-se os ajustes realizados nos dados de fNIRS. Portanto, este conjunto não apresenta mais 32 saídas, mas 20. Sendo estas, 12 da câmera e 8 do fNIRS, como apresentado na Tabela 12.

Tabela 12 – Treino dos classificadores

Name	Type	Range	Import as
column_1	double	1 .. 3	Response
column_2	double	-2.33139 .. 3.90579	Predictor
column_3	double	-2.04392 .. 2.56796	Predictor
column_4	double	-0.497899 .. 8.859...	Predictor
column_5	double	-2.14583 .. 1.98951	Predictor
column_6	double	-1.47373 .. 4.34291	Predictor
column_7	double	-1.87648 .. 1.94799	Predictor
column_8	double	-1.30265 .. 2.10919	Predictor
column_9	double	-1.26046 .. 2.18039	Predictor
column_10	double	-2.02136 .. 1.3214	Predictor
column_11	double	-1.18089 .. 2.97331	Predictor
column_12	double	-1.49468 .. 3.38773	Predictor
column_13	double	-1.26027 .. 2.17462	Predictor
column_14	double	-3.49994 .. 4.73439	Do not import
column_15	double	-3.18504 .. 3.37175	Do not import
column_16	double	-2.70223 .. 5.22363	Predictor
column_17	double	-2.96157 .. 4.23222	Predictor
column_18	double	-1.83713 .. 4.4977	Predictor
column_19	double	-3.29646 .. 3.4455	Do not import
column_20	double	-2.83332 .. 3.46388	Do not import
column_21	double	-4.76034 .. 4.72629	Do not import
column_22	double	-2.40359 .. 2.38598	Predictor
column_23	double	-2.99473 .. 4.12719	Do not import
column_24	double	-2.31436 .. 4.21372	Predictor
column_25	double	-2.65957 .. 3.02647	Do not import
column_26	double	-2.05287 .. 3.63659	Predictor
column_27	double	-2.44332 .. 3.14071	Do not import
column_28	double	-2.14806 .. 3.64871	Predictor
column_29	double	-2.70491 .. 3.05941	Do not import
column_30	double	-2.75598 .. 3.53188	Do not import
column_31	double	-2.23598 .. 3.64072	Do not import
column_32	double	-1.90102 .. 4.02616	Do not import
column_33	double	-1.94643 .. 4.04231	Predictor

Fonte: O Autor (2018)

No mesmo contexto do item anterior o conjunto montado apresentou algumas colunas desconsideradas, por estas não representarem uma resposta emocional. Lembrando que a retirada destes canais é devido a injeção de ruídos que não condizem com o estudo quando os canais se fazem presentes.

Estes são os três modelos de conjuntos de dados para os classificadores que serão usados no estudo. Ao longo da descrição dos classificadores exemplificou-se as variações ou mudanças em relação aos apresentados.

Para cada um dos classificadores gerou-se um código preditor. Usando o comando `yfit = trainedClassifier.predictFcn(Conjunto de teste)`, pode-se gerar as predições das respostas possíveis do outro sujeito. O código preditor por

sua vez gerou uma matriz de respostas chamada “yfit”, que foi comparado com o comando `confusionmat(Resposta real,yfit)`. Este comando tem por princípio comparar a resposta prevista pelo preditor com a resposta real. Esta comparação gera um índice de acertos, que demonstra a eficácia da rede em prever novos dados.

3.5.2.2 Conjunto de Classificadores 1 - Similaridade entre sujeitos.

Construiu-se classificadores para avaliar a similaridade entre os sujeitos, onde os subconjuntos de dados foram treinados em nove dos 10 sujeitos e testado no outro, criando 30 classificadores distintos. Com o uso deste método, pode-se mapear as respostas de cada sujeito comparado aos demais.

1. Subconjunto de dados da câmera.

A montagem do classificador do subconjunto câmera se procedeu seguindo-se a forma apresentada na Tabela 10, com 12 entradas e uma saída. Neste primeiro contexto temos 267300 pontos de treino para a predição, tendo 22275 saídas. Estes dados serviram para treinar cada um dos 10 classificadores, portanto este procedimento se repetiu por 10 vezes.

2. Subconjunto de dados fNIRS.

Para o subconjunto de dados de fNIRS, usou-se os canais selecionados anteriormente, mudando dos 20 possíveis canais, para 8 canais de interesse, como mostrado na Tabela 11. Neste subconjunto, cada classificador tem 200475 pontos em predição e responde a 22275 saídas de treino. Como no subconjunto de câmera, este subconjunto, também forneceu 10 preditores.

3. Subconjunto de ambos os dados.

Para o subconjunto completo dos dados se fez a união do subconjunto da câmera com o subconjunto do fNIRS, como mostrado na Tabela 12. Esta análise se encontra mais próxima da que será usada ao final, pois mescla ambos os métodos de aquisição de emoções para uma mesma resposta.

Para o terceiro conjunto de 10 classificadores, tem-se para cada, 445500 pontos em predição, já que foi realizada a soma dos canais de entrada. Como resposta para estes pontos usou-se as 22275 saídas geradas a partir do sujeito.

3.5.2.3 Conjunto de Classificadores 2 - Teste de validação cruzada

A validação cruzada é uma forma de representar todos as combinações possíveis dos dados, podendo assim apresentar uma medida mais confiável, para uma predição da respostas de outro sujeito. Portanto, neste teste, procurou-se observar e gerar os classificadores para o programa final, podendo evidenciar a acurácia do sistema. Utilizou-se um modelo que foi treinado em oito sujeitos e testado nos outros dois. Para tanto, temos 45 classificadores, formando todas as possíveis combinações.

Para este classificador, os dados foram organizados da seguinte forma: para treino teve-se 8 sujeitos, ou seja, um total de 19800 amostras. O conjunto de dados treinado apresenta as 20 entradas do conjunto completo, sendo 12 de câmeras e 8 de fNIRS além de uma saída.

Neste classificador, usou-se oito sujeitos para treino, o que totaliza 396000 pontos de análise. Com um número maior de dados a rede deveria apresentar uma acurácia maior, devido a ter simulado mais possibilidades. Os testes foram realizados nos outros dois sujeitos não utilizados para treino. Estes sujeitos apresentaram 4950 possibilidades de respostas que foram testadas pelos preditores.

3.5.2.4 Conjunto de Classificadores 3 - Similaridade entre exames

O conjunto três, foi separado em três exames, igualmente distribuídos como demonstrado na metodologia. Na busca de índices melhores de predição, treinou-se o sujeito nos dois primeiros exames e testou-se no terceiro exame. Este teste trabalhou com um único sujeito, gerando um índice de conformidade dos dados para o sujeito e avaliando a similaridade dos exames.

Neste conjunto temos 2145 saídas para a predição dentre as 2475 saídas totais. Estes dados servem para treinar os 10 classificadores de cada conjunto de dados, que depois foram testados em 330 respostas possíveis do sujeito que ainda não haviam sido usadas.

1. Conjunto de dados Câmera

Para o conjunto de dados responsável pelo reconhecimento facial, estruturou-

se da seguinte forma: os 12 canais das câmeras são as entradas para uma única saída emocional.

A construção do classificador se assemelha ao teste para o sujeito. Com a diferença de que são usados conjuntos menores de dados. Para este classificador teremos 25740 pontos de treino, para 2145 amostras.

2. Conjunto de dados fNIRS.

Para o segundo conjunto com os teste no fNIRS, os canais são selecionados seguindo as respostas de ativação obtidas na análise do estudo do fNIRS. O classificador apresenta 8 canais de entrada e uma saída emocional.

A construção do classificador em fNIRS se baseará em um conjunto com 17160 pontos de treino para as mesmas 2145 amostras.

3. Conjunto de ambos os dados.

Este terceiro conjunto de teste segue os moldes dos primeiros, com a diferença de usar todo o conjunto de dados.

O conjunto com ambas as entradas apresentou 42900 pontos de treino. Este classificador tende sempre a apresentar melhores resultados devido a conter mais pontos de treino, o que garante uma amostragem maior.

Os treinos dos classificadores seguiram todos o mesmo método e modelo de classificador. Usou-se o classificado SVM por este apresentar melhores resultados.

Os teste dos classificadores seguiram as mesmas características, como demonstrado antes do primeiro modelo de classificação ilustrado.

3.5.2.5 Conjunto de Classificadores 4 - integrando os exames e os sujeitos

Este conjunto de classificadores utiliza a resposta do próprio sujeito para prever as respostas, nos mesmos moldes do modelo anterior. A principal diferença está no uso de todos os sujeitos para o treinamento e para o teste. Para este conjunto, usou-se a validação cruzada, como já descrita. Dois exames foram utilizados como amostras para treino e um exame para teste. Os conjuntos de dados seguiram como os anteriores, separados em 3 conjuntos: câmera, fNIRS e ambos. Para o conjunto de treino, teve-se 16500 amostras e o conjunto de teste, 8260 amostras.

Com os dados organizados e montados, realizou-se a implementação dos classificadores, resultando em três classificadores para cada um dos subconjuntos, formando um total de nove classificadores.

Para cada uma das nove possibilidades, se criou um classificador nos mesmos moldes dos anteriores. Com o cruzamento destes dados, pode-se ter uma ideia de como será a resposta do preditor final para um novo exame de um novo paciente, validando o estudo.

Para avaliar a acurácia de predição de cada uma das emoções, utilizou-se as respostas independentes para cada emoção a partir da matriz de confusão. Os resultados de cada matriz foram somados gerando uma matriz de confusão para cada um dos três conjuntos, como mostrado nas Tabelas 13, 14, 15.

Tabela 13 – Matriz de confusão do conjunto câmera.

Câmera	Negativo	Positivo	Neutro	Precisão
Negativo	4445	3238	567	53,8 %
Positivo	1777	6067	406	73,5 %
Neutro	1548	3189	3513	42,5 %
Sensibilidade	57,2 %	48,5%	78,3%	56,6 %

Fonte: O Autor (2018)

Pode-se observar que para o conjunto da câmera, o preditor apresentou acurácia de $56,6 \pm 0,02\%$. O melhor índice encontrado foi de $73,5 \pm 0,21\%$, para o reconhecimento das emoções positivas, por outro lado a rede apresentou dificuldade em reconhecer emoções neutras com $42 \pm 0,06\%$ de acertos.

Tabela 14 – Matriz de confusão do conjunto fNIRS.

fNIRS	Negativo	Positivo	Neutro	Precisão
Negativo	3418	2586	2246	41,4 %
Positivo	1245	5027	1978	60,9 %
Neutro	1214	2931	4105	49,7 %
Sensibilidade	58,1 %	47,6%	49,2%	50,7 %

Fonte: O Autor (2018)

O classificador dos dados de fNIRS apresentou $50,7 \pm 0,021\%$ de acurácia. Diferente dos de câmera, os dados de fNIRS apresentam maior dificuldade em prever dados de emoções negativas, com $41,4 \pm 0,07\%$ acertos. Igualmente ao câmera, o melhor índice encontra-se em prever emoções positivas com $60,9 \pm 0,11\%$.

Tabela 15 – Matriz de confusão do conjunto Ambos.

Ambos	Negativo	Positivo	Neutro	Precisão
Negativo	6533	671	1046	79,1 %
Positivo	939	6246	1065	75,7 %
Neutro	1105	942	6203	75,1 %
Sensibilidade	76,1 %	76,1%	74,6%	76,6 %

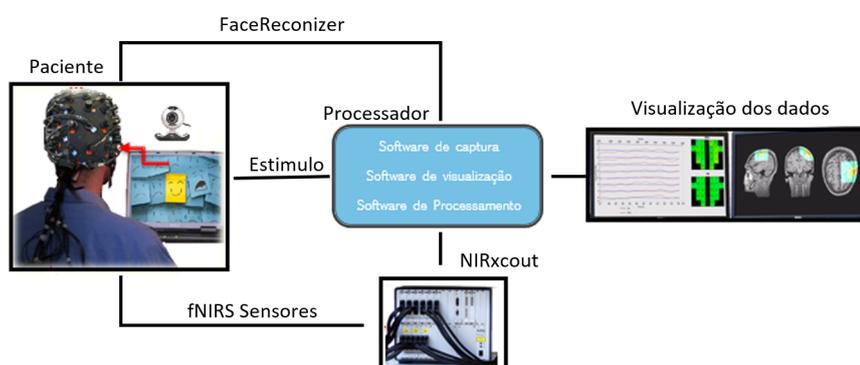
Fonte: O Autor (2018)

O classificador de ambos os dados apresentou índices superiores atingindo $76,6 \pm 0,02\%$ de acurácia. Para este classificador, diferente dos demais, a emoção que melhor foi prevista é a negativa, com $79,1 \pm 0,06\%$ de acertos. Este se assemelha ao classificador fNIRS, tendo a emoção neutra como a pior de se prever, com $75,1 \pm 0,01\%$ de acertos. Para este classificador, todos os resultados tiveram melhores índices de previsão, como esperado.

3.6 Sistema de reconhecimento de sinais

A união de todas as etapas do estudo cria a ferramenta de detecção de emoções faciais, integrando o uso de *fNIRS* e reconhecimento facial. A análise não ocorre em tempo real, adquirindo uma sequência de dados, que é analisada pelo *NIRx* e o sistema de câmera, como apresentado na Figura 47. Estes dados são processados posteriormente e utilizados pelo preditor no *Matlab*.

Figura 45 – Fluxograma do sistema proposto de Reconhecimento



Fonte: O Autor (2018)

O controle dos estímulos, ou seja, o paradigma, é associado ao fNIRS e ao Reconhecimento facial, para que os dados fiquem sincronizados com os estímulos que são enviados para o preditor.

Durante o processamento dos dados, são separados os canais de interesse que já estão dispostos no preditor. Para uso da ferramenta há a necessidade do treinamento prévio do usuário, o software captura a atividade e incorpora estes dados ao banco. Entrando, assim, o sujeito como um dado do classificador.

Com este procedimento a ferramenta está pronta a gerar resultados de emoções sobre uma atividade deste sujeito. Os dados devem passar pelo preditor após a realização da atividade, gerando as respostas emocionais que ocorreram durante a atividade.

4 RESULTADOS

Neste capítulo será demonstrado e avaliado os resultados encontrados para cada um dos objetivos desta dissertação. Primeiramente demonstrou-se os resultados do protocolo de evocação, em seguida avaliando os métodos de reconhecimento de fNIRS e facial, terminando pela análise do modelo de classificação gerado no estudo. Antes do resultado, apresenta-se detalhes dos 10 sujeitos que se voluntariaram a participar do estudo. Os voluntários são cinco homens com idades entre 23 a 30 anos e 5 mulheres com idades entre 19 a 28 anos. Todos os voluntários são saudáveis, estudantes e apresentam vidas ativas.

4.1 Resultados do protocolo de evocações

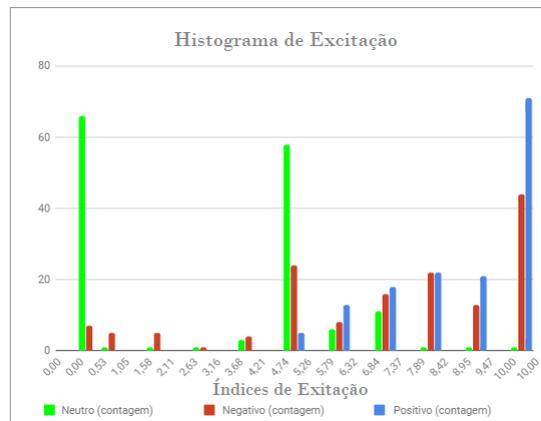
Visando avaliar o protocolo utilizado para evocar as emoções, foram realizadas duas perguntas após a exibição de cada imagem. Estas perguntas correspondiam a qual emoção provocada e qual intensidade desta emoção, que nada mais seria que a valência e a excitação, conceitos este definidos pelo IAPS.

O conjunto de dados apresenta 45 respostas para cada sujeito, totalizando 450 respostas para os 10 sujeitos. Para mostrar a frequência de cada resposta, utilizou-se um gráfico de histograma com todas as respostas obtidas para as perguntas.

Como pode-se observar o histograma da Figura 46 apresenta um acúmulo de respostas neutras com excitação próximo a zero, e algumas repetições com excitação em cinco, o que determina que os sujeitos acharam a imagem neutra. Para as respostas positivas e negativas, obteve-se todas as excitações acima de cinco, o que caracteriza imagens excitantes com maior acúmulo próximo a 10, o que caracteriza as imagens como muito excitante, como esperado.

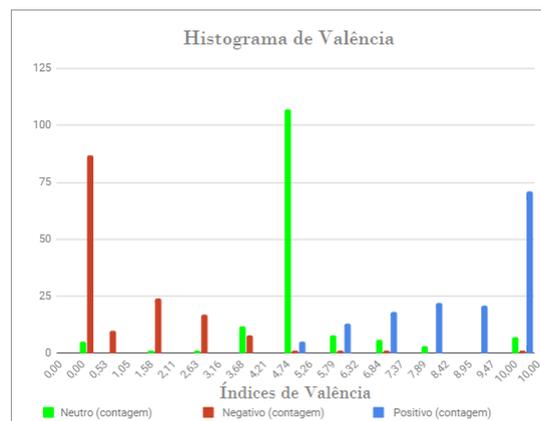
Para as respostas de valência obteve-se os dados apresentados no histograma da Figura 47.

Figura 46 – Histograma das respostas para a Excitação



Fonte: O Autor (2018)

Figura 47 – Histograma das respostas de Valência



Fonte: O Autor (2018)

Para o caso da valência, o gráfico de histograma revelou que as imagens com caráter neutro, encontram-se com seu resultado em cinco, o que caracteriza imagens neutras. Já as imagens negativas apresentam uma distribuição abaixo de cinco, e com maior acumulo próximo do zero. Por outro lado, como esperado, os resultados para as emoções positivas tiveram uma distribuição acima de cinco, da mesma forma, mais intensa no extremo próximo a 10.

Outra forma de analisar os resultados é comparado-os com os fornecidos

pela IAPS. Para esta comparação calculou-se a média de cada uma das respostas para cada uma das emoções como apresentado na Tabela 16.

Tabela 16 – Médias para cada atividade.

Evocação neutra		
Comparação	Valência	Excitação
IAPS	4,95	2,29
Sujeitos	5,15	2,99
Evocação positiva		
Comparação	Valência	Excitação
IAPS	8,12	4,92
Sujeitos	8,69	7,11
Evocação negativa		
Comparação	Valência	Excitação
IAPS	2,00	6,51
Sujeitos	1,13	6,98

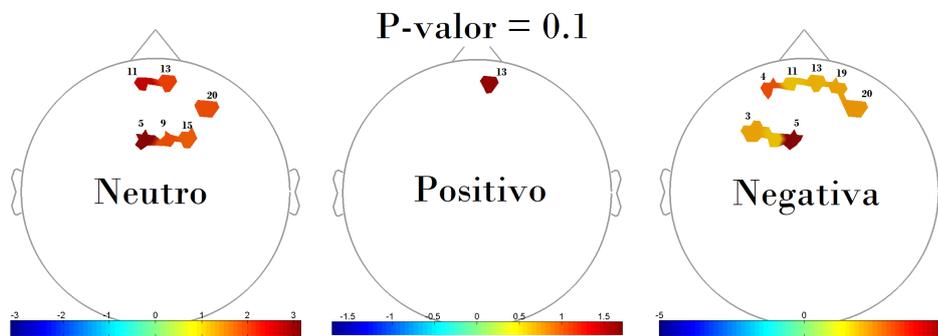
Fonte: O Autor (2018)

Pode-se observar pela análise dos resultados da Tabela 16, que os dados gerados pelas respostas dos sujeitos, em muito se aproxima dos dados gerados pelo IAPS. Esta aproximação das respostas demonstra que o público escolhido enquadra-se nas características do grupo utilizado pela IAPS. As respostas apresentadas sobre o protocolo de evocação, validando as escolhas das imagens para o estudo, bem como a disposição das mesmas no protocolo, evidencia que o mesmo e seus resultados encontram-se dentro do desejado pelo estudo.

4.2 Resultados do fNIRS

Para a análise das áreas de ativação, comparou-se todos as respostas dos 10 sujeitos, realizando-se uma análise estatísticas das ativações em cada canal, com o uso do *NIRSlab*. Estas análises estatísticas demonstraram quais são os canais de interesse para cada uma das emoções evocadas nos sujeitos. Com estas respostas realizou-se a minimização dos classificadores, tirando aqueles canais que não apresentaram ativação sobre as emoções, como ilustrado na Figura 48.

Figura 48 – Respostas de cada emoção do fNIRS



Fonte: O Autor (2018)

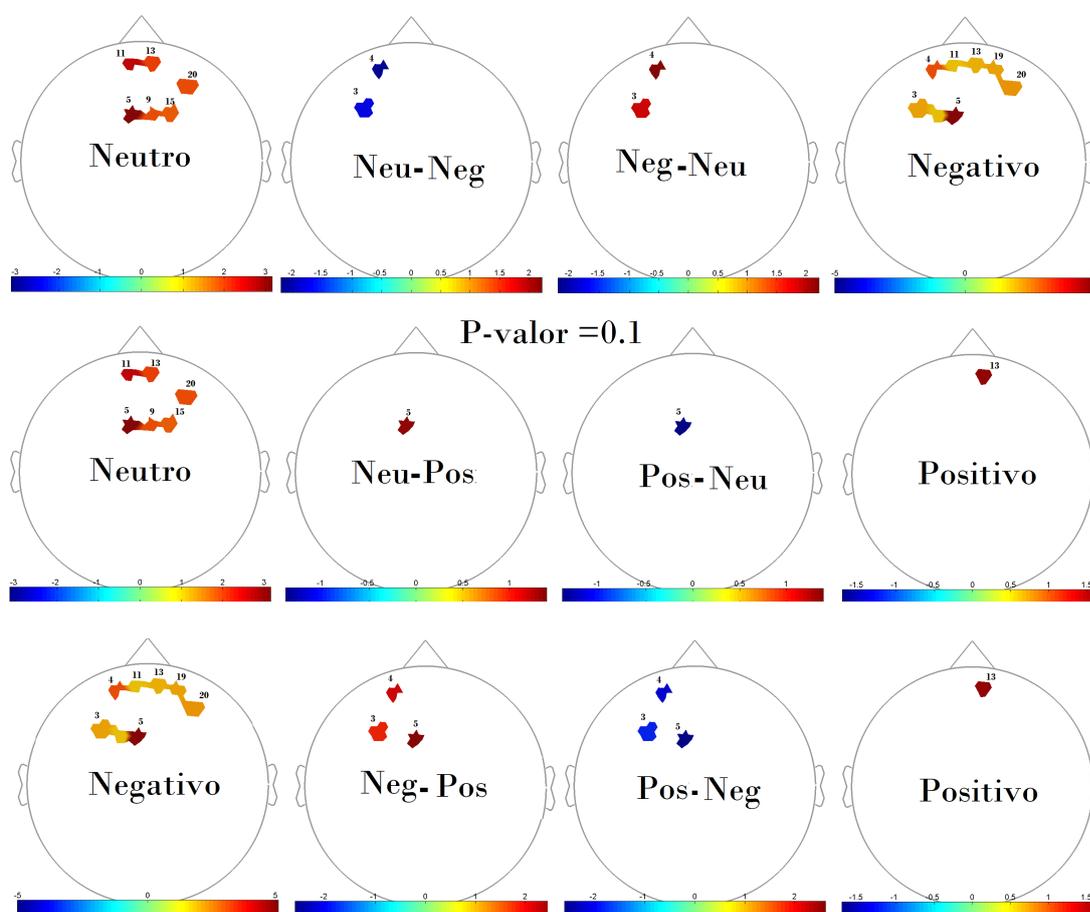
As respostas apresentadas na Figura 48 demonstram a ativação de cada uma das emoções, comparadas ao *baseline*. Esta configuração demonstra as ativações das emoções comparadas com os momentos em que o cérebro não recebia nenhum estímulo, estando em repouso, o que caracteriza baixa ativação. Esta comparação apresenta todas as respostas de ativação geradas no cérebro durante o período em que as imagens foram apresentadas. Com isto, pode-se observar que os canais 5, 9, 11, 13, 15 e 20 responderam à ativação neutra, já os canais 3, 4, 5, 11, 13, 19 e 20 à emoção Negativa. Contudo para emoção positiva apenas o canal 13 foi ativado.

Nesta configuração, obteve-se nove canais de ativação para as três emoções. Apresentou-se também canais que tiveram ativação para mais de uma emoção, como exemplo o canal 11, que consta nas duas emoções simultaneamente, o que demonstra que este canal não diferencia emoções.

Outra forma de analisar estes mesmos dados é realizando uma análise entre emoções. Nesta análise compara-se os canais ativados entre as emoções, demonstrando as diferenças de ativação entre os conjuntos de canais de cada emoção.

Como pode-se visualizar na Figura 49, a comparação entre resultados apresenta uma resposta mais direta quando busca-se classificar sinais entre emoções. Pode-se observar que o canal 5 demonstra ativação unicamente para emoções neutras, já os canais 3 e 4 demonstram unicamente ativação para emoções negativas. A emoção positiva não apresentou nenhum canal exclusivo.

Figura 49 – Respostas entre emoções no fNIRS



Fonte: O Autor (2018)

A resposta nos mostra os canais que realmente ativaram para cada emoção, se um canal ativa em mais de uma emoção ele foi retirado, pois não apresenta diretamente a emoção em específico. Ficando agora dois canais para emoções negativas, um para emoção neutra e nenhum para a emoção positiva. Escolheu-se o canal 13 para a emoção positiva, mesmo este canal estando presente nas demais emoções, pois ele apresentou uma maior intensidade para esta emoção.

4.3 Resultado do reconhecimento facial

O estudo analisou as plataformas de reconhecimento facial individualmente e comparou-as umas com as outras, usando um banco padrão de vídeos compondo expressões faciais. Esta técnica foi realizada para avaliar a qualidade das plataformas como também avaliar qual o melhor método de classificação que representaria os dados do estudo.

Como podemos observar nos resultados usando um classificador para a resposta da rede CNN, obteve-se uma melhor acurácia no preditor. Esta melhora vem tanto das melhorias implantadas na aquisição, como também pelo simples fato que a rede foi retreinada com novos valores, o que, por si só, já melhora os rendimentos. Observa-se, na Tabela 17, os resultados da plataforma CNN original (HO, 2016) e o do classificador utilizado para teste.

Tabela 17 – Comparação entre método original e re-treinado.

Emoção	CNN Original	Classificador	Diferença
Zangado	0,49	0,9385	0,4485
Medo	0,11		-0,11
Feliz	0,82	0,97	0,15
Triste	0,53	0,94	0,41
Surpreso	0,76	0,83	0,07
Neutro	0,619	0,95	0,331
Acurácia total	0,58	0,95	0,37

Fonte: O Autor (2018)

Pode-se observar melhora na acurácia de resposta do classificador comparado ao CNN original. Contudo no banco simulado não há emoções de medo. Portanto para esta emoção inicialmente não teremos conclusões. Já a avaliação global do classificador de teste apresentou melhora de cerca de 37% com relação ao original.

O resultado com o classificador de teste, comparado com a plataforma SVM original (MICHEL, 2003), apresentou índices melhores. Esta melhora, como a ocorrida na rede CNN, vem das melhorias implantadas na aquisição, como também pelo fato de que a rede foi treinada novamente com novos valores. Os resultados da comparação, podem ser observados na Tabela 18.

Tabela 18 – Comparação entre método original e re-treinado.

Emoção	SVM Original	Classificador	Diferença
Zangado	0,66	0,96	0,3
Medo	0,66		-0,66
Feliz	0,91	0,98	0,07
Triste	0,61	0,89	0,28
Surpreso	0,83	1	0,17
Neutro			
Acurácia total	0,71	0,97	0,26

Fonte: O Autor (2018)

Há uma melhora na acurácia encontrada como resposta pelo novo classificador. Da mesma forma, por se tratar do mesmo banco a emoção de medo não foi caracterizada. No contexto global do classificador, obteve-se uma melhora de cerca de 26% na predição.

Para o classificador usando ambas as plataformas ao mesmo tempo, obteve-se melhorias esperadas pelo incremento de classes de entrada, como mostrado na Tabela 19.

Tabela 19 – Comparação entre os métodos originais e o re-treinado.

Emoção	SVM Original	CNN Original	Classificador	Diferença
Zangado	0,66	0,49	0,98	0,49
Medo	0,66	0,11		-0,66
Feliz	0,91	0,82	0,97	0,15
Triste	0,61	0,53	1	0,47
Surpreso	0,83	0,76	1	0,24
Neutro		0,619	0,99	0,37
Acurácia total	0,71	0,58	0,97	0,4

Fonte: O Autor (2018)

O classificador com ambas as entradas apresentou cerca de 40% de melhoria na acurácia, índice superior ao realizadas anteriormente, como era esperado. Os resultados foram dispostos apresentando os índices iniciais das plataformas, os índices gerados pelo classificador montado e a diferença entre os três, como mostrado na Tabela 20.

Tabela 20 – Respostas dos testes dos reconhecimentos faciais.

Plataforma	Original	Classificador	Diferença
CNN	58 %	95 %	37 %
SVM	71 %	97 %	26 %
CNN e SVM	-	98 %	-

Fonte: O Autor (2018)

Pode-se observar que os índices tiveram grande melhoria, devido as modificações realizadas no estudo. Outro ponto importante relacionado ao aumento nos índices diz respeito ao retreino com novos dados, o que favorece o classificador a melhorar seus coeficientes. O uso do banco padronizado demonstrou também que o classificador que melhor responde ao uso com ambos os métodos é a SVM. Para o caso do estudo com a plataforma CNN a rede árvore de decisões apresentou um resultado levemente superior, mas levando em consideração as três formas, optou-se por escolher o classificador SVM.

4.4 Resultados do software de reconhecimento

Os resultados dos modelos de classificação para uso final do estudo foram descritos em separado nesta seção a fim de compreender as melhorias e os motivos de escolha entre eles.

4.4.1 Conjunto de Classificadores 1 - Similaridade dos sujeitos.

No estudo das semelhanças entre sujeitos analisou-se quanto um classificador treinado em nove sujeito conseguia prever o sujeito que faltava. Portanto, cada um dos classificadores foi utilizado para testar um sujeito, tendo sido treinado em nove. Para avaliar todas as respostas optou-se por apresentar os conjuntos câmera, fNIRS e um terceiro com ambos os conjuntos, como apresentado na Tabela 21.

Tabela 21 – Respostas dos classificadores de similaridade entre sujeitos.

Sujeito de teste	Câmera	fNIRS	Ambos
S10	36,9 %	36,2 %	38,2 %
S09	36,7 %	34,6 %	39,1 %
S08	36,1 %	36,3 %	40,6 %
S07	37,5 %	33,5 %	39,5 %
S06	30,9 %	31,5 %	33,7 %
S05	33,0 %	33,2 %	34,3 %
S04	39,8 %	38,7 %	40,4 %
S03	31,7 %	34,6 %	38,5 %
S02	38,3 %	37,6 %	40,6 %
S01	38,7 %	34,1 %	39,8 %
Média ± Desvio	35,9 ± 3,06 %	35,5 ± 2,17%	38,4 ± 2,50%

Fonte: O Autor (2018)

Com este estudo, pode-se perceber que as respostas entre sujeitos não são lineares. Os classificadores usados apresentaram uma aproximação média entre sujeitos de no máximo 38,4%, o que demonstra que há variação de respostas entre sujeitos, devido ao fato de serem respostas fisiológica de indivíduos diferentes. Este efeito fica mais evidente quando analisamos sujeitos individualmente, observando a qualidade do classificador para o próprio sujeito, como apresentado na Tabela 22.

Tabela 22 – Respostas do classificador para os dados de treino.

Sujeito de teste	Câmera	fNIRS	Ambos
S10	87 %	83 %	89 %
S09	85 %	84 %	90 %
S08	71 %	90 %	94 %
S07	89 %	90 %	98 %
S06	91 %	82 %	93 %
S05	86 %	79 %	94 %
S04	89 %	95 %	97 %
S03	92 %	92 %	94 %
S02	83 %	86 %	92 %
S01	81 %	75 %	87 %
Média ± Desvio	85,4 ± 6,14%	85,6 ± 6,21%	92,8 ± 3,42%

Fonte: O Autor (2018)

Estes resultados demonstram que quando usa-se os dados dos próprios sujeitos os índices de acurácia encontrados são superiores aos obtidos com um sujeito fora do treino. Este fato está relacionado com as diferenças das respostas entre sujeitos.

4.4.2 Conjunto de Classificadores 2 - Teste de validação cruzada

Na busca de um classificador para uso robusto utilizou-se validação cruzada para gerar índices de previsão do conjunto com ambos dados de entrada. O uso da validação cruzada coloca os resultados mais próximos dos reais. Foram realizados os classificadores para cada uma das combinações possíveis. Para melhor visualizar os dados apresenta-se as respostas dos testes para cada um dos 45 classificadores na Tabela 23. Podendo assim ter uma visão mais direta das diferenças e dos resultados individuais de cada combinação.

Tabela 23 – Respostas do classificador treino em oito e testa em dois.

Testados em	Ambos	Testados em	Ambos
S01-S02	40,23 %	S03-S10	38,50%
S01-S03	34,58 %	S04-S05	28,61%
S01-S04	36,26 %	S04-S06	37,74%
S01-S05	35,06 %	S04-S07	33,29%
S01-S06	32,43 %	S04-S08	34,48%
S01-S07	40,88 %	S04-S09	35,41%
S01-S08	38,68 %	S04-S10	33,07%
S01-S09	40,50 %	S05-S06	34,77%
S01-S10	31,97 %	S05-S07	32,77%
S02-S03	33,92 %	S05-S08	30,89%
S02-S04	37,66 %	S05-S09	33,15%
S02-S05	34,44 %	S05-S10	35,45%
S02-S06	34,18 %	S06-S07	28,97%
S02-S07	39,56 %	S06-S08	25,92%
S02-S08	36,67 %	S06-S09	27,31%
S02-S09	34,36 %	S06-S10	37,47%
S02-S10	30,73 %	S07-S08	35,03%
S03-S04	31,70 %	S07-S09	37,11%
S03-S05	34,40 %	S07-S10	31,90%
S03-S06	38,40 %	S08-S09	32,84%
S03-S07	31,54 %	S08-S10	39,43%
S03-S08	30,59 %	S09-S10	40,01%
S03-S09	32,68 %	Média±Desvio	33,9±3,588 %

Fonte: O Autor (2018)

Estes classificadores, da mesma forma, não apresentaram resultados aceitáveis para este estudo, com $33,9 \pm 3,588\%$ de acurácia.

4.4.3 Conjunto de Classificadores 3 - Similaridade dos exames

Como os resultados da validação cruzada não foram os esperados, partiu-se para uma avaliação sobre os exames dos sujeitos. Este conjunto busca avaliar a similaridade de cada um dos sujeitos, individualmente, em seus três exames. Os preditores foram treinados nos dois primeiros exames e testados no terceiro exame

não utilizado para treino. A partir destes testes gerou-se os resultados individuais da Tabela 24.

Tabela 24 – Respostas do classificador similaridade dos exames.

Sujeito	Câmera	fNIRS	Ambos
S10	51,8 %	51,5 %	65,8 %
S09	40,3 %	51,2 %	40,5 %
S08	49,8 %	58,3 %	41,6 %
S07	34,9 %	75,4 %	53,7 %
S06	35,8 %	36,9 %	55,7 %
S05	56,8 %	51,2 %	70,6 %
S04	43,9 %	75,3 %	68,8 %
S03	53,3 %	43,6 %	46,6 %
S02	36,1 %	40,3 %	48,2 %
S01	46,9 %	58,8 %	41,9 %
Média ± Desvio	44,9 ± 7,96 %	54,2 ± 13,17%	53,3 ± 11,6%

Fonte: O Autor (2018)

Pode-se observar uma melhora nos resultados dos classificadores, com índices de até 75% de acurácia. Estes resultados apresentam uma perspectiva de análise mais adequada para se utilizar neste estudo. Avaliando o sujeito individualmente observou-se resultados melhores, o que novamente fortalece a ideia de que os sujeitos apresentam respostas fisiológicas não homogêneas.

4.4.4 Conjunto de Classificadores 4 - Treino em exames com todos os sujeitos

Para uso do classificador foi realizado um quarto experimento onde se uniu os dados de todos os sujeitos levando em consideração o modelo da estrutura do classificador de número três, que apresentou resultados melhores. Junto a este foi realizada a validação cruzada dos exames 1, 2 e 3, como mostrada na Tabela 25.

Tabela 25 – Respostas dos testes validação cruzada.

Treino	Câmera	fNIRS	Ambos
Exame 1 e 2	55,4%	51,5%	78,8 %
Exame 1 e 3	57,8%	50,9%	76,3 %
Exame 2 e 3	54,9%	53,9%	76,5 %
Média \pm Desvio	56,0 \pm 1,55%	52,1 \pm 1,59%	77,2 \pm 1,39%

Fonte: O Autor (2018)

Pode-se observar uma melhora visível nos índices de classificação, principalmente no classificador de ambos os conjuntos, o que comprova a hipótese de que os métodos juntos apresentam resultados melhores. Outro ponto muito importante que fica claro neste classificador é que os sujeitos tendem a ser diferentes, mas quando usamos dados dos mesmos para treinar e testar os índices de acerto aumentam, pois a rede já treinou com dados similares.

Para avaliar as emoções em separado, deste último conjunto de classificação, apresentou-se as acurácias de cada emoção com demonstrado na Tabela 26.

Tabela 26 – Matriz de confusão dos conjuntos.

Emoção	Câmera	fNIRS	Ambos
Negativo	53,8 %	41,4 %	79,1 %
Positivo	73,5 %	60,9 %	75,7 %
Neutro	42,5 %	49,7 %	75,1 %
Acurácia média	56,6 %	50,7 %	76,6 %

Fonte: O Autor (2018)

Pode-se observar que os índices apresentados no conjunto ambos são muito superiores aos individuais para todas as emoções. Nos conjuntos câmera e fNIRS a emoção positiva era de melhor compreensão, já para o ambos a com melhor índice foi a negativa.

5 DISCUSSÃO

O objetivo deste estudo foi utilizar métodos de reconhecimento emocional para a implementação em jogos sérios. Os métodos utilizados foram o de reconhecimento facial e o de fNIRS. Estes métodos foram inseridos nas etapas de análise, antes mesmo do uso direto para o estudo. Depois foram amplamente utilizados para a criação e montagem de classificadores de emoções, atendendo as expectativas e gerando bons resultados. Para validar a eficácia deste algoritmo, ele foi comparado com as etapas dos estudos de pré-criação, utilizando diferentes técnicas de análises estatísticas. Todas as análises deste estudo levaram em conta apenas as regiões que são de interesse no estudo da emoção.

5.1 Desenvolvimento do protocolo de evocação

A escolha e a montagem do protocolo geraram ótimos resultados, com ressalva de alguns momentos onde os sujeitos acabavam por não expressar reações faciais, devido a congelar suas expressões, focando toda a atenção em analisar a imagem. Esta percepção se deu em análise visual dos vídeos dos voluntários durante o exame. Deve-se analisar formas para que o sujeito não se prenda por completo a imagem. O uso de imagens mais diretas, onde o sujeito não precise analisar a imagem para compreender o contexto, ajuda a minimizar estes ocorridos. Outra observação que ficou bem evidente é que quando do uso de imagens com grande necessidade de processamento de compreensão, gera-se uma forte ativação de toda área pré-frontal. Estudos futuros podem ser usados para a retirada do processamento de compreensão do estudo emocional.

O projeto se baseou em estudos envolvendo emoções realizados pela IAPS, por este motivo as imagens utilizadas foram fornecidas pela IAPS. A escolha por este, se deve as imagens já serem amplamente estudadas e apresentarem índices consistentes. As perguntas utilizadas para averiguar os estímulos e a dominância das imagens representaram por completo o que se buscava, apresentando índices

semelhantes aos fornecidos pela IAPS. Este resultado caracterizou como válidas as imagens escolhidas, bem como o público alvo utilizado.

O estudo sempre buscou imagens com baixos índices de excitação e índices médios de valência, para representar as imagens neutras. Para as imagens negativas os índices de valência eram baixos, próximos a zero, e para as positivas os índices eram de valência alta, próximos a 10. Sendo que para emoções negativas e positivas as imagens apresentam índices de excitação altos, próximos a 10. O uso de imagens negativas com uma excitação alta acarretou ruídos no sistema, pois as imagens causavam incômodo aos sujeitos. No estudo já havia sido retirado imagens com índices de violência maiores, mas é recomendado que este filtro seja ampliado para mais algumas imagens, evitando assim estes ruídos indesejados.

5.2 Análise hemodinâmica usando fNIRS

A análise de fNIRS nos demonstra as ativações emocionais geradas na área cortical, pela variação de hemoglobina oxigenada. Essas respostas representam o contexto do subconsciente emocional de cada indivíduo sobre sua perspectiva emocional. Trata-se de medir uma resposta empírica do sujeito, baseada na experiência e na observação do mesmo. Portanto, as respostas apresentadas são de caráter pessoal. O uso de sensores na área pré-frontal claramente analisa processamentos, que apresentam ligações com as emoções. A análise destes resultados gera evidências suficientes para diferenciar as emoções. A partir destas evidências pode-se observar quais canais realmente sofrem influência emocionais.

A análise utilizada comparou as respostas para dois conjuntos distintos, a emoção e o contraste, verificando estatisticamente quais os canais apresentavam ativação em comparação ao contraste. No estudo, usou-se duas formas de contraste para perceber o conjunto de canais que ativavam e para diferenciar os canais ativados para cada emoção evocada.

Optou-se por utilizar o conjunto de canais que era ativado emocionalmente, e não o conjunto que diferenciava entre emoções. Esta escolha se fez para manter um número considerado de canais em que o classificador pudesse analisar as diferenças dos dados. A percepção de laterização das emoções ficou evidente, sendo

possível observar a distinção de emoções positivas na direita e emoções negativas na esquerda. Não podendo, no entanto, definir tão claramente entre emoções neutras e processamento simples. Com o acréscimo futuro de imagens sem caráter emocional, sendo puramente processamento, será possível reavaliar as respostas excluindo delas o processamento não emocional.

5.3 Análise de emoções pelo reconhecimento facial

Para a plataforma CNN, a emoção de surpresa foi a que menos apresentou melhoria com as modificações, devido à utilização de um filtro de suavização do ruído com a média de 10 aquisições. Como a resposta da emoção de surpresa é de curta duração para a resposta facial, o filtro acabou suavizando também a sua análise, mas mesmo assim os índices encontrados foram melhores que os originais.

Já para a plataforma SVM, os resultados demonstraram melhoras consideráveis para todos os índices avaliados. Como a plataforma já apresentava bons resultados para a emoção feliz, teve pouco acréscimo e manteve-se ainda como uma das emoções que melhor classifica junto a emoção de surpresa. As emoções zangado e triste, que apresentavam índices baixos, tiveram uma excelente melhora, ficando com um índice final de acertos muito superior.

O classificador em que os dois métodos entram juntos como classes apresentaram todos os índices excelentes, comprovando a validade do estudo. Outro ponto importante é a análise da emoção neutra, que não existia na plataforma SVM, e apresenta resultados ótimos.

As melhorias implementadas, bem como as modificações e união dos métodos de reconhecimento facial, geraram resultados consideráveis para o reconhecimento facial atingindo índices excelentes para o banco padronizado. Índices estes muito superiores aos gerados pelos métodos tradicionais que foram usados como base. Estes resultados fortalecem o conceito de que a união de métodos e formas diferentes para uma análise emocional torna o reconhecimento mais eficiente e robusto.

Os índices atingidos foram excelentes, contudo a análise foi realizada sobre

expressões padronizadas e bem definidas, o que levou a atingir índices elevados. O uso comum provavelmente não apresentará índices tão elevados, mas certamente terá ainda índices excelentes e de ótima validade para a classificação emocional.

5.4 Pré-processamento e adaptação de software de reconhecimento emocional

Os índices encontrados não atingiram o potencial esperado de um classificador. Quando trabalha-se com análises biológicas, como respostas fisiológicas e expressões faciais, interage-se com muitas variáveis que influenciam as respostas dos sistemas. Estas variáveis podem vir do fato de que por mais que as pessoas sejam parecidas, suas respostas fisiológicas para uma análise usando fNIRS não são idênticas, o que acarreta em variações nas respostas diretamente. Em virtude destes fatos, há a necessidade de constantemente se ampliar o número de sujeitos no banco de dados. Portanto, quanto mais é ampliado o número de sujeitos, maiores serão os índices de acerto e melhores serão os resultados do estudo. Outro ponto importante está no contexto de que se o sujeito for incluído no banco antes de fazer uso da ferramenta também este terá uma grande influência nas respostas, pois, como vimos no estudo o classificador de um sujeito testado nele mesmo apresenta resultados melhores.

Quando compara-se os classificadores dos conjuntos câmera e fNIRS com a resposta do conjunto de ambos, pode-se observar que em cinco sujeitos a acurácia encontrada é maior para o conjunto fNIRS que para o conjunto câmera. E apenas em um sujeito teve-se uma acurácia maior no conjunto câmera. Esta evidência demonstra que para o protocolo utilizado os dados de fNIRS apresentam maior confiança para um possível preditor de emoções. A resposta facial não ficou tão evidente devido aos sujeitos não expressarem variações faciais. Esta evidência pode ser causada pelo protocolo utilizado se tratar de um exame, onde os sujeitos reagem à um conjunto de imagens. O protocolo é excelente para diferenciar emoções, que pode não ser o mesmo para o interesse do sujeito pela atividade. O que não ocorreria em um jogo que funciona de forma contínua, despertando curiosidade e variações de cena que causem uma sequência lógica de interesse.

Quando se utiliza maior conjuntos de classes, deve-se por lógica obter maior acurácia no preditor, devido ao classificador se basear em um número maior de entradas. Estas características ficam evidentes nas respostas dos sujeitos dez, seis, cinco e dois, que apresentaram evolução no classificador do conjunto ambos, como esperado, apresentando uma melhora considerada nos resultados.

Por outro lado, podemos observar pela análise dos sujeitos nove, oito, sete, quatro e um, um resultado diferente, onde o classificador do conjunto ambos, apresenta resultados inferiores aos outros dois. Este evento ocorre devido a canais da câmera e do fNIRS entrarem em conflito quanto a sua resposta. Teve-se, neste caso, classes independentes que apresentam mais de uma resposta, quando comparado a outras classes que apresentam uma única resposta. Este erro se deve também a algumas vezes os sujeitos não demonstrarem a expressão facial devido a congelar a expressão, isto ocorre quando é ativado uma janela traumática do sujeito, também chamada janela de *Killer*.

Em análise posterior aos testes, pode-se observar que dois sujeitos não apresentaram respostas faciais em relação a um grupo de imagens específicas, demonstrando as janelas traumáticas. O sujeito sete apresentou inatividade para as duas imagens de acidentes de trânsito e uma de acidente trágico. Quando questionado posteriormente o sujeito relatou ter sofrido recentemente um acidente violento de trânsito, o que causou maior impacto, fazendo lembrar da situação. O sujeito nove apresentou inatividade em quatro imagens fortes de ferimentos e quando questionado relatou que as imagens o incomodaram muito e impactaram a ponto deste desviar o olhar e não querer observar a imagem no período de 10 segundos. Para ambos os sujeitos fica claro, no vídeo gravado, o congelamento das expressões e a não ativação nos dados, pois o classificador segue demonstrando atitude neutra. Estes ruídos também ocorreram em alguns momentos eventuais para os demais sujeitos.

6 CONCLUSÃO

Neste estudo, criou-se protocolos de evocação capazes de estimularem emoções positivas, negativas e neutras. Estímulos estes, responsáveis por provocar uma resposta emocional direta no voluntário e assim é possível realizar leituras emocionais específicas destes. Pode-se analisar quais áreas específicas respondem a estes estímulos e quais suas funcionalidades emocionais. Analisou-se o uso do equipamento *NIRx* como um reconhecedor emocional, encontrando as áreas corticais que apresentam respostas emocionais distintas e classificáveis. Evidenciou-se o uso de plataformas de reconhecimento facial para reconhecimento de padrões emocionais, unificando diferentes métodos de construção, para a obtenção de melhores índices de predição do reconhecimento facial. Unificou-se o uso de reconhecimentos distintos, avaliando áreas fisiológicas diferentes para a obtenção de um sistema de predição emocional mais robusto e com índices melhores de acurácia.

Os resultados obtidos para o conjunto 4, demonstraram a validade dos sistemas de reconhecimento emocional. Quando são analisadas respostas cerebrais relacionados ao subconsciente emocional de cada indivíduo, analisa-se a resposta empírica do sujeito, baseada nas experiências vividas pelo mesmo. Esta análise trata-se de algo extremamente pessoal e individual, o que dificulta a classificação de padrões. O estudo pôde superar este desafio, atingindo índices de $77,2 \pm 1,39\%$, para prever as emoções vivenciadas durante os experimentos.

O processamento final para este estudo se deu em caráter exploratório, podendo-se construir ferramentas que juntas trabalhem para a obtenção de respostas emocionais. Em primeira instância, estas respostas ainda não são geradas em tempo real, o que de nada prejudica os achados encontrados, já que a conversão para uma análise em tempo real necessita apenas de uma união destas ferramentas. Esta união é possível com a utilização de um software chamado *OpenViBE*, que tem por sua funcionalidade a execução dos pre-processamentos gerados no NIRS-lab. Contudo, os demais processamentos e melhorias estão ligados à linguagem *Python*, já devidamente programados no software de classificação.

6.1 TRABALHOS FUTUROS

O uso de uma leitura cerebral com o intuito de medir emoções mostrou-se muito interessante e apresentou ótimos resultados, podendo se reconhecer que os achados da literatura sobre este tipo de análise, são consistentes e devem ser amplamente trabalhados. O uso do equipamento fNIRS, devido a ser de baixo custo quando comparado à ressonância magnética, permite a realização de exames em diferentes locais, facilitando a realização de novas pesquisas. O estudo pode ser aperfeiçoado futuramente acrescentando análises de nível emocional do participante, bem como análise dos movimentos sacádicos para retirar possíveis movimentos de olhos que influenciam na análise pré frontal.

A proposta de unir o presente estudo, a uma forma de obtenção de padrões emocionais a partir da voz, irá criar sistemas de reconhecimento de emoções cada vez mais robustos e precisos. Além disso o uso de preditores similares ao desenvolvido nesta dissertação tem o potencial de serem utilizados também para o reconhecimento de padrões psicológicos ligados ao estudo de psicopatias.

A construção de classificadores para diferentes formas de obtenção de uma mesma característica sentimental, são ao mesmo tempo desafiadores e motivantes, pela abertura que isso pode promover para um contexto de ampliação dos conhecimentos e das técnicas de aprendizados supervisionados.

Referências

- BIANCHI-BERTHOUBE, N. Understanding the role of body movement in player engagement. *Human-Computer Interaction*, Taylor & Francis, v. 28, n. 1, p. 40–75, 2013. 56
- BIANCHI-BERTHOUBE, N. et al. On posture as a modality for expressing and recognizing emotions. In: *Emotion and HCI workshop at BCS HCI London*. [S.l.: s.n.], 2006. 27
- BOYLE, E. A. et al. An update to the systematic literature review of empirical evidence of the impacts and outcomes of computer games and serious games. *Computers & Education*, Elsevier, v. 94, p. 178–192, 2016. 27
- BRUSILOVSKY, P.; KOBASA, A.; NEJDL, W. *The adaptive web*. [S.l.]: Springer-Verlag Berlin Heidelberg, 2007. 33
- BUSSO, C. et al. Analysis of emotion recognition using facial expressions, speech and multimodal information. In: *ACM. Proceedings of the 6th international conference on Multimodal interfaces*. [S.l.], 2004. p. 205–211. 55
- CACIOPPO, J. T. et al. The psychophysiology of emotion. *Handbook of emotions*, v. 2, p. 173–191, 2000. 33
- CHANEL, G. et al. Emotion assessment from physiological signals for adaptation of game difficulty. *IEEE Transactions on Systems, Man, and Cybernetics-Part A: Systems and Humans*, IEEE, v. 41, n. 6, p. 1052–1063, 2011. 56
- DERRYBERRY, A. *Debunking myths about serious games*. [S.l.], 2008. 32
- DÖRNER, R. et al. *Serious games: foundations, concepts and practice*. [S.l.]: Springer, 2016. 33
- FRIDLUND, A. J.; EKMAN, P.; OSTER, H. Facial expressions of emotion. Lawrence Erlbaum Associates, Inc, 1987. 45
- FUSAR-POLI, P. et al. Functional atlas of emotional faces processing: a voxel-based meta-analysis of 105 functional magnetic resonance imaging studies. *Journal of psychiatry & neuroscience: JPN*, Canadian Medical Association, v. 34, n. 6, p. 418, 2009. 31
- HARLEY, J. M.; BOUCHET, F.; AZEVEDO, R. Aligning and comparing data on emotions experienced during learning with metatutor. In: *SPRINGER. International Conference on Artificial Intelligence in Education*. [S.l.], 2013. p. 61–70. 54

- HAZLETT, R. L. Measuring emotional valence during interactive experiences: boys at video game play. In: ACM. *Proceedings of the SIGCHI conference on Human Factors in computing systems*. [S.l.], 2006. p. 1023–1026. 27
- HEATMAP, I. *Real-time Analytics*. 2017. Disponível em: <<https://heatmap.me/>>. 31
- HERRMANN, M. J.; EHLIS, A.-C.; FALLGATTER, A. J. Prefrontal activation through task requirements of emotional induction measured with nirs. *Biological psychology*, Elsevier, v. 64, n. 3, p. 255–263, 2003. 56, 66
- HO, J. *Facial Emotion Recognition*. [S.l.]: GitHub, 2016. <https://github.com/JostineHo/mememoji>. 53, 54, 100
- HUETTEL, S. A. et al. *Functional magnetic resonance imaging*. [S.l.]: Sinauer Associates Sunderland, 2004. v. 1. 57
- IHHP, I. *Test your Emotional Intelligence Free EQ Quiz*. 2017. Disponível em: <<http://www.ihhp.com/free-eq-quiz/>>. 31
- KIM, Y.; LEE, H.; PROVOST, E. M. Deep learning for robust feature generation in audiovisual emotion recognition. In: IEEE. *Acoustics, Speech and Signal Processing (ICASSP), 2013 IEEE International Conference on*. [S.l.], 2013. p. 3687–3691. 28
- KIRBY, M.; SIROVICH, L. Application of the karhunen-loeve procedure for the characterization of human faces. *IEEE Transactions on Pattern analysis and Machine intelligence*, IEEE, v. 12, n. 1, p. 103–108, 1990. 48
- KIVIKANGAS, J. M. et al. A review of the use of psychophysiological methods in game research. *journal of gaming & virtual worlds*, Intellect, v. 3, n. 3, p. 181–199, 2011. 27
- KLEINSMITH, A.; BIANCHI-BERTHOUBE, N.; STEED, A. Automatic recognition of non-acted affective postures. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, IEEE, v. 41, n. 4, p. 1027–1038, 2011. 56
- LANG, P.; BRADLEY, M. M. The international affective picture system (iaps) in the study of emotion and attention. *Handbook of emotion elicitation and assessment*, Oxford University Press USA New-York, NY, USA, v. 29, 2007. 58
- LANG, P.; BRADLEY, M. M. The international affective picture system (iaps) affective ratings of pictures and instruction manual. technical report a-8. *Handbook of emotion elicitation and assessment*, University of Florida, Gainesville, FL, v. 29, 2008. 59

- LARSEN, J. T. et al. The psychophysiology of emotion. *Handbook of emotions*, Guildford Press New York, NY, v. 3, p. 180–195, 2008. 33
- LECUN, Y.; BENGIO, Y. et al. Convolutional networks for images, speech, and time series. *The handbook of brain theory and neural networks*, v. 3361, n. 10, p. 1995, 1995. 51
- LIAW, A.; WIENER, M. et al. Classification and regression by randomforest. *R news*, v. 2, n. 3, p. 18–22, 2002. 51
- MA, M.; OIKONOMOU, A.; JAIN, L. C. *Serious games and edutainment applications*. [S.l.]: Springer, 2011. 32
- MAGERKO, B. et al. Intelligent adaptation of digital game-based learning. In: ACM. *Proceedings of the 2008 Conference on Future Play: Research, Play, Share*. [S.l.], 2008. p. 200–203. 33
- MATTHEWS, P. M.; JEZZARD, P. Functional magnetic resonance imaging. *Journal of Neurology, Neurosurgery & Psychiatry*, BMJ Publishing Group Ltd, v. 75, n. 1, p. 6–12, 2004. ISSN 0022-3050. Disponível em: <<http://jnnp.bmj.com/content/75/1/6>>. 57, 58
- MEYERDING, S. G.; RISIUS, A. Reading minds: Mobile functional near-infrared spectroscopy as a new neuroimaging method for economic and marketing research—a feasibility study. *Journal of Neuroscience, Psychology, and Economics*, Educational Publishing Foundation, 2018. 70
- MICHALSKI, R. S.; CARBONELL, J. G.; MITCHELL, T. M. *Machine learning: An artificial intelligence approach*. [S.l.]: Springer Science & Business Media, 2013. 48
- MICHEL, P. Support vector machines in automated emotion classification. *Churchill College, June*, 2003. 51, 52, 53, 100
- MOON, H.; PHILLIPS, P. J. Computational and performance aspects of pca-based face-recognition algorithms. *Perception*, SAGE Publications Sage UK: London, England, v. 30, n. 3, p. 303–321, 2001. 48
- PHILLIPS, P. J. et al. The feret evaluation methodology for face-recognition algorithms. *IEEE Transactions on pattern analysis and machine intelligence*, IEEE, v. 22, n. 10, p. 1090–1104, 2000. 49
- SALA, O. *Fundamentos da espectroscopia Raman e no infravermelho*. [S.l.]: Unesp, 1996. 34
- SCHMITZ, D. C. *NIRX & FNIRS:INTRODUCTION & OVERVIEW*. edicao. Local, 2016. 40 p. Disponível em: <<http://nirx.net/>>. Acesso em: 20 junho. de 2017. 35, 36, 37, 38, 39

- SMOLA, A. J. et al. *Advances in large margin classifiers*. [S.l.]: MIT press, 2000. 50
- ŠPAKOV, O.; MINIOTAS, D. Visualization of eye gaze data using heat maps. *Elektronika ir elektrotechnika*, Technologija, p. 55–58, 2007. 31
- SUSI, T.; JOHANNESSON, M.; BACKLUND, P. *Serious games: An overview*. [S.l.]: Institutionen för kommunikation och information, 2007. 27
- TECHNOLOGIES., . M. *THE NIRSCAP Manual*. [S.l.], 2016. v. 1, n. 1, 1–223 p. Disponível em: <<http://www.nirx.net>>. 37, 38, 42
- TECHNOLOGIES., . N. M. *NIRStim 4.0 User Manual*. [S.l.], 2016. v. 1, n. 1, 1–31 p. Disponível em: <<http://www.nirx.net>>. 63
- TECHNOLOGIES., . N. M. *THE NIRSCout Manual*. [S.l.], 2016. v. 1, n. 1, 1–223 p. Disponível em: <<http://www.nirx.net>>. 39, 40
- TIJS, T.; BROKKEN, D.; IJSSELSTEIJN, W. Creating an emotionally adaptive game. In: SPRINGER. *International Conference on Entertainment Computing*. [S.l.], 2008. p. 122–133. 55
- VACHIRATAMPORN, V. et al. An analysis of player affect transitions in survival horror games. *Journal on Multimodal User Interfaces*, Springer, v. 9, n. 1, p. 43–54, 2015. 27, 55
- VALSTAR, M.; MAAT, L.; GAIA, M. P. P. Mmi face database. 2005. 76
- WHALEN, P. J. et al. The emotional counting stroop paradigm: a functional magnetic resonance imaging probe of the anterior cingulate affective division. *Biological psychiatry*, Elsevier, v. 44, n. 12, p. 1219–1228, 1998. 55
- WIKLUND, M. et al. Evaluating educational games using facial expression recognition software: measurement of gaming emotion. In: ACADEMIC CONFERENCES INTERNATIONAL LIMITED. *European Conference on Games Based Learning*. [S.l.], 2015. p. 605. 54

Anexos

PARECER CONSUBSTANCIADO DO CEP

DADOS DO PROJETO DE PESQUISA

Título da Pesquisa: RECONHECIMENTO FACIAL E FNIRS PARA DETECÇÃO DE TRAÇOS DE EMOÇÕES PARA SEREM APLICADOS EM JOGOS SÉRIOS

Pesquisador: Dario Francisco Guimarães de Azevedo

Área Temática:

Versão: 2

CAAE: 86884418.8.0000.5336

Instituição Proponente: UNIAO BRASILEIRA DE EDUCACAO E ASSISTENCIA

Patrocinador Principal: Financiamento Próprio

DADOS DO PARECER

Número do Parecer: 2.665.006

Apresentação do Projeto:

O pesquisador principal do estudo: Dario Francisco Guimarães de Azevedo encaminhou ao CEP-PUCRS, em 04/05/2018, os seguintes documentos:

PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1092486.pdf;
DocumentoUnificadodoProjetodePesquisa_1522779419394.pdf;
cartaRespostaPendencias.doc;
Cartaresposta.pdf;
TCLE_V2_Milton.pdf;
FolhaDeRosto.pdf

Objetivo da Pesquisa:

O pesquisador principal do estudo: Dario Francisco Guimarães de Azevedo encaminhou ao CEP-PUCRS, em 04/05/2018, os seguintes documentos:

PB_INFORMAÇÕES_BÁSICAS_DO_PROJETO_1092486.pdf;
DocumentoUnificadodoProjetodePesquisa_1522779419394.pdf;
cartaRespostaPendencias.doc;
Cartaresposta.pdf;
TCLE_V2_Milton.pdf;

Endereço: Av.Ipiranga, 6681, prédio 50, sala 703

Bairro: Partenon

CEP: 90.619-900

UF: RS

Município: PORTO ALEGRE

Telefone: (51)3320-3345

Fax: (51)3320-3345

E-mail: cep@pucrs.br



Pontifícia Universidade Católica do Rio Grande do Sul
Pró-Reitoria Acadêmica
Av. Ipiranga, 6681 - Prédio 1 - 3º. andar
Porto Alegre - RS - Brasil
Fone: (51) 3320-3500 - Fax: (51) 3339-1564
E-mail: proacad@pucrs.br
Site: www.pucrs.br/proacad