

---

# **Abordagem Computacional para Aprimoramento das Habilidades com as Emoções em Indivíduos com Autismo**

---

**Adilmar Coelho Dantas**



UNIVERSIDADE FEDERAL DE UBERLÂNDIA  
FACULDADE DE COMPUTAÇÃO  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA COMPUTAÇÃO

Uberlândia  
19 de maio de 2022



**Adilmar Coelho Dantas**

**Abordagem Computacional para Aprimoramento das  
Habilidades com as Emoções em Indivíduos com  
Autismo**

Tese de doutorado apresentada ao Programa de Pós-graduação da Faculdade de Computação da Universidade Federal de Uberlândia como parte dos requisitos para a obtenção do título de Doutor em Ciência da Computação.

Área de concentração: Ciência da Computação

Orientador: Marcelo Zanchetta do Nascimento

Uberlândia  
19 de maio de 2022

Ficha Catalográfica Online do Sistema de Bibliotecas da UFU  
com dados informados pelo(a) próprio(a) autor(a).

D192 2022	<p>Dantas, Adilmar Coelho, 1992- Abordagem Computacional para Aprimoramento das Habilidades com as Emoções em Indivíduos com Autismo [recurso eletrônico] / Adilmar Coelho Dantas. - 2022.</p> <p>Orientador: Marcelo Zanchetta do Nascimento. Tese (Doutorado) - Universidade Federal de Uberlândia, Pós-graduação em Ciência da Computação. Modo de acesso: Internet. Disponível em: <a href="http://doi.org/10.14393/ufu.te.2022.211">http://doi.org/10.14393/ufu.te.2022.211</a> Inclui bibliografia.</p> <p>1. Computação. I. Nascimento, Marcelo Zanchetta do, 1976-, (Orient.). II. Universidade Federal de Uberlândia. Pós-graduação em Ciência da Computação. III. Título.</p> <p>CDU: 681.3</p>
--------------	---





**UNIVERSIDADE FEDERAL DE UBERLÂNDIA**  
 Coordenação do Programa de Pós-Graduação em Ciência da Computação  
 Av. João Naves de Ávila, nº 2121, Bloco 1A, Sala 243 - Bairro Santa Mônica, Uberlândia-MG, CEP 38400-902  
 Telefone: (34) 3239-4470 - www.ppgco.facom.ufu.br - cpqfacom@ufu.br



### ATA DE DEFESA - PÓS-GRADUAÇÃO

Programa de Pós-Graduação em:	Ciência da Computação				
Defesa de:	Tese, 6/2022, PPGCO				
Data:	6 de abril de 2022	Hora de início:	13:30	Hora de encerramento:	16:21
Matrícula do Discente:	11723CCP001				
Nome do Discente:	Adilmar Coelho Dantas				
Título do Trabalho:	Abordagem Computacional para aprimoramento das habilidades com as emoções em indivíduos com autismo				
Área de concentração:	Ciência da Computação				
Linha de pesquisa:	Ciência de Dados				
Projeto de Pesquisa de vinculação:	-				

Reuniu-se, por videoconferência, a Banca Examinadora, designada pelo Colegiado do Programa de Pós-graduação em Ciência da Computação, assim composta: Professores Doutores: Bruno Augusto Nassif Travençolo - FACOM/UFU, Fabiano Azevedo Dorça - FACOM/UFU, Márcia Aparecida Silva Bissaco - Núcleo de Pesquisas Tecnológicas/UMC, Francisco de Assis Zampirolli - CMCC/UFAB e Marcelo Zanchetta do Nascimento - FACOM/UFU orientador do candidato.

Os examinadores participaram desde as seguintes localidades: Márcia Aparecida Silva Bissaco - Mogi das Cruzes/SP; Francisco de Assis Zampirolli - Santo André/SP; Bruno Augusto Nassif Travençolo, Fabiano Azevedo Dorça e Marcelo Zanchetta do Nascimento - Uberlândia/MG. O discente participou da cidade de Uberlândia/MG.

Iniciando os trabalhos o presidente da mesa, Prof. Dr. Marcelo Zanchetta do Nascimento, apresentou a Comissão Examinadora e o candidato, agradeceu a presença do público, e concedeu ao Discente a palavra para a exposição do seu trabalho. A duração da apresentação do Discente e o tempo de arguição e resposta foram conforme as normas do Programa.

A seguir o senhor presidente concedeu a palavra, pela ordem sucessivamente, aos examinadores, que passaram a arguir o candidato. Ultimada a arguição, que se desenvolveu dentro dos termos regimentais, a Banca, em sessão secreta, atribuiu o resultado final, considerando o candidato:

**Aprovado.**

Esta defesa faz parte dos requisitos necessários à obtenção do título de Doutor.

O competente diploma será expedido após cumprimento dos demais requisitos, conforme as normas do Programa, a legislação pertinente e a regulamentação interna da UFU.

Nada mais havendo a tratar foram encerrados os trabalhos. Foi lavrada a presente ata que após lida e achada conforme foi assinada pela Banca Examinadora.



Documento assinado eletronicamente por **Francisco de Assis Zampirolli, Usuário Externo**, em 25/04/2022, às 10:30, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Bruno Augusto Nassif Travençolo, Professor(a) do Magistério Superior**, em 25/04/2022, às 13:37, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcelo Zanchetta do Nascimento, Professor(a) do Magistério Superior**, em 25/04/2022, às 13:52, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Marcia Aparecida Silva Bissaco, Usuário Externo**, em 25/04/2022, às 13:53, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



Documento assinado eletronicamente por **Fabiano Azevedo Dorça, Professor(a) do Magistério Superior**, em 02/05/2022, às 12:00, conforme horário oficial de Brasília, com fundamento no art. 6º, § 1º, do [Decreto nº 8.539, de 8 de outubro de 2015](#).



A autenticidade deste documento pode ser conferida no site [https://www.sei.ufu.br/sei/controlador\\_externo.php?acao=documento\\_conferir&id\\_orgao\\_acesso\\_externo=0](https://www.sei.ufu.br/sei/controlador_externo.php?acao=documento_conferir&id_orgao_acesso_externo=0), informando o código verificador **3500324** e o código CRC **43B3B9CE**.

---

# Agradecimentos

Início meus agradecimentos primeiramente a Deus por ter me proporcionado esse feito em minha vida. A minha amiga em especial **Taynara Prestes** que sempre me auxiliou em nas aplicações e nas realizações dos experimentos deste trabalho e colaborou diretamente e indiretamente para o desenvolvimento do mesmo.

Ao meu orientador **Marcelo Zanchetta do Nascimento**, por todo empenho e dedicação acadêmica junto ao projeto, sempre me guiando neste e em outros projetos desenvolvidos de forma direta e indireta.

A todos os meus amigos (**Michel Santos, Wanderson, Heidie, Guilherme, Guilherme Brilhante, etc**) que passaram pelo laboratório LIPAI - LITED e que me proporcionaram sempre bons momentos de estudo e descontração para todas as horas difíceis e acima de tudo pela amizade construída no decorrer desses anos.

Agradeço em especial a minha namorada **Norlene Cristina** que sempre esteve presente ao meu lado me apoiando e torcendo para a finalização desta etapa tão importante em minha vida.

A minha amiga **Sara**, que sempre me apoiou em todos os momentos e torceu pela minha vitória nessa jornada tão importante em minha vida.

Agradeço à Faculdade de Computação (FACOM) por ter sempre disponibilizado os melhores professores, funcionários para que este trabalho ocorresse com sucesso, aos funcionários **Erisvaldo e Sônia**, pela paciência e agilidade sempre que precisei.

Agradeço, também, à **CAPES** pelo apoio financeiro dedicado a este projeto, mesmo em momentos de crise em nosso país e com tantos cortes na Educação. Ao Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) (Projeto #311404/2021-9) e à Fundação de Amparo à Pesquisa do Estado de Minas Gerais (FAPEMIG) (Projeto #APQ-00578-18) pelos apoios financeiros. Finalmente, gostaria de agradecer à **Universidade Federal de Uberlândia** por ter me proporcionado amplo acesso aos laboratórios, espaço físico e todo apoio acadêmico necessário para o desenvolvimento deste trabalho.



*“Faça o dia valer as lembranças”  
(Bill Gates)*



---

## Resumo

A detecção das expressões faciais e o reconhecimento das emoções básicas são importantes nas relações interpessoais de indivíduos em uma sociedade. Os indivíduos neurotípicos têm essas habilidades aprimoradas de maneira natural e gradativa durante o ciclo da vida. No entanto, os indivíduos com Transtorno do Espectro Autista (TEA) apresentam dificuldades com a detecção das expressões faciais e representação das emoções. Essa restrição prejudica as interações sociais dos indivíduos com TEA em ambiente social. Este trabalho apresenta uma ferramenta computacional que emprega jogos sérios para auxiliar no ensino e aprimoramento dessas habilidades em indivíduos com TEA. Os algoritmos computacionais para a detecção e reconhecimento das emoções foram desenvolvidos numa abordagem que combina descritores de características manuais e aprendidos, por um modelo de redes neurais convolucionais. Essa ferramenta utiliza conceitos de jogos sérios, modelagem 3D para personagens, cenários e metodologias da aprendizagem baseada em jogos, com a finalidade de proporcionar o desenvolvimento dessas habilidades de maneira lúdica. Essa ferramenta funciona em diversas plataformas, com execução on-line, para interação com o usuário. Para a etapa de detecção e reconhecimento, quatro bancos de dados de imagens de domínio público (CK+, FER2013, RAF-DB e MMI) foram empregados para avaliação. O jogo foi investigado com voluntários, em cinco sessões, para as fases de investigação (referência, intervenção e manutenção). Os resultados mostraram que a ferramenta proposta contribuiu para o aprimoramento das habilidades de detecção e reconhecimento das emoções básicas em indivíduos com TEA.

**Palavras-chave:** Expressões Faciais, Emoções Básicas, Processamento de Imagens, Transtorno do Espectro Autista, Computação Afetiva, Jogos Sérios.





---

# Abstract

The detection of facial expressions and the recognition of basic emotions are important in the interpersonal relationships of individuals in a society. Neurotypical individuals have these skills enhanced naturally and gradually during the life cycle. However, individuals with Autism Spectrum Disorder (ASD) have difficulties with the detection of facial expressions and representation of emotions. This restriction affects the social interactions of individuals with ASD in a social environment. This work presents a computational tool that employs serious games to aid in the learning and improvement of these skills in individuals with ASD. The computational algorithms for emotion detection and recognition were developed in an approach that combines manual and learned feature descriptors, by a convolutional neural network model. This tool uses concepts from serious games, 3D modeling for characters, scenarios, and game-based learning methodologies, in order to provide the development of these skills in a playful way. This tool works on several platforms, running online, for interaction with the user. For the detection and recognition phase, four public domain image databases (CK+, FER2013, RAF-DB, and MMI) were used for evaluation. The game was investigated with volunteers in five sessions for the research phases (reference, intervention, and maintenance). The results showed that the proposed tool contributed to the improvement of basic emotion detection and recognition skills in individuals with ASD.

**Keywords:** Facial expressions. Basic emotions. Image processing. Autistic Spectrum Disorder. Affective Computing. Serious Games.



---

## Lista de ilustrações

Figura 1	–	Imagens obtidas com ressonância magnética das atividades cerebrais em pacientes típicos (a) e indivíduos com TEA (b).	30
Figura 2	–	Diagrama de fluxo de uma aplicação GBL.	34
Figura 3	–	Exemplos da variação de detalhes em objetos 3D de acordo com a distância, usando o algoritmo LOD.	38
Figura 4	–	UA identificadas no trabalho de (EKMAN; FRIESEN, 1978).	39
Figura 5	–	Sequência de imagens para representação da emoção surpresa, extraída da base (KANADE; COHN, 2005) com as combinações das UA 1, 2, 5 e 27.	40
Figura 6	–	Associação das UM com as UA.	41
Figura 7	–	Representação para calcular a imagem integral em (a) por meio da soma das áreas retangulares em (b).	42
Figura 8	–	Representação de movimento das regiões da boca por meio da técnica de deformação.	42
Figura 9	–	Exemplo de PCA aplicado em imagem da face para a extração de características.	43
Figura 10	–	Detecção dos pontos discriminantes na face para a detecção das emoções.	44
Figura 11	–	Representação das relações entre inteligência artificial, aprendizado de máquina e aprendizado profundo.	45
Figura 12	–	Rede Neural Artificial <i>Perceptron</i> com um único neurônio.	46
Figura 13	–	Exemplo de uma RNA MLP com características de aprendizagem.	47
Figura 14	–	Exemplo de operações realizadas durante o processo de convolução: (A) Representação de uma imagem 4x4x3 em RGB, (B) demonstração de uma convolução em uma imagem 5x5x1 com um <i>kernel</i> 3x3x1 para obter um recurso 3x3x1.	48
Figura 15	–	Representação do deslocamento do <i>kernel</i> em uma imagem.	48
Figura 16	–	Representação de convolução em uma matriz de imagem MxNx3 com um <i>kernel</i> 3x3x3.	49

Figura 17 – Animações desenvolvidas para ensinar a interpretação das emoções em indivíduos com TEA: (a) alegria, (b) desgosto, (c) raiva e (d) surpresa. . . . .	61
Figura 18 – Tabuleiro desenvolvido durante a produção do jogo. . . . .	62
Figura 19 – Processo de modelagem de personagem em 3D para o personagem Juninho. . . . .	63
Figura 20 – Desenvolvimento, estruturação e modelagem em 3D do personagem. . . . .	63
Figura 21 – Personagem desenvolvido em 3D expressando as emoções: alegria, raiva, tristeza e surpresa com base nos pontos chaves propostos por (EKMAN; ROSENBERG, 2005). . . . .	64
Figura 22 – Mapa mental da plataforma do jogo <i>Face Emotions</i> . . . . .	65
Figura 23 – Fluxograma das etapas da ferramenta proposta para aprimoramento das habilidades das emoções. . . . .	66
Figura 24 – A interface apresenta a personagem Ana escolhida para auxiliar no processo de aprimoramento das emoções. . . . .	67
Figura 25 – Interface para personalização do personagem com opções de roupas e outros acessórios (boné, chapéu e óculos). . . . .	68
Figura 26 – Interface do jogo para avaliar as competências do conhecimento da emoção. . . . .	68
Figura 27 – <i>Gameplay</i> do SG proposto: (a) identificação do usuário (ID); escolha de um dos personagens; (c) avaliação das habilidades sobre as emoções básicas; (d) cenário com 6 diferentes interações sobre as emoções; (e) o usuário pode melhorar as habilidades emocionais; e (f) a cada emoção expressa corretamente, um item surpresa é liberado para o personagem que pode avançar no mapa do jogo. . . . .	69
Figura 28 – Interface de treinamento de competências emocionais baseado nas expressões faciais e funcionalidades. . . . .	69
Figura 29 – Interface de treinamento para reconhecimento de emoções e capacidades de expressão. . . . .	70
Figura 30 – Informações extraídas com a técnica <i>Haar-Like</i> : (a), (b), (c) e (d) são características de borda; (e), (f), (g), (h), (i), (j), (k) e (l) são traços de linha; e (m) e (n) recursos <i>center surround</i> . . . . .	71
Figura 31 – Exemplo do detector de estruturas baseado no modelo em cascata. . . . .	71
Figura 32 – Etapas das regiões detectadas da face (face, olhos e boca). . . . .	72
Figura 33 – Marcadores inseridos na imagem do rosto do usuário. . . . .	72
Figura 34 – Demonstração das coordenadas dos pontos: a) pontos marcados em amarelo do contorno da face do usuário; b) pontos que representam o lado direito e esquerdo da face do usuário. . . . .	73
Figura 35 – Exemplo das coordenadas de pontos para detecção do contorno dos olhos direito (a) e esquerdo (b) do usuário. . . . .	73
Figura 36 – Pontos responsáveis pelo contorno das sobrancelhas: (a) lado direito do usuário e (b) lado esquerdo do usuário. . . . .	74

Figura 37 – Demonstração das coordenadas de pontos responsáveis pelo contorno de regiões: (a) boca e (b) nariz. . . . .	74
Figura 38 – Representação dos filtros convolucionais. . . . .	76
Figura 39 – Modelo proposto para obtenção das informações da face para classificação da emoção. . . . .	77
Figura 40 – Interface de ( <i>dashbord</i> ) com as informações de desempenho do usuário fornecido ao especialista. . . . .	78
Figura 41 – Tabelas e seus respectivos campos desenvolvidos para o módulo de dados. .	78
Figura 42 – Exemplo de imagens de pessoas expressando emoções para as bases: (a) CK+, (b) FER2013, (c) RAF e (d) MMI. . . . .	83
Figura 43 – Valor da acurácia em relação ao número de épocas empregada para avaliação: (a) treinamento e (b) teste. . . . .	84
Figura 44 – Valor da minimização de erro ( <i>Loss</i> ) do modelo proposto: (a) treinamento e (b) teste. . . . .	85
Figura 45 – Matriz de confusão do modelo proposto em amostra obtida do conjunto de dados CK+. . . . .	86
Figura 46 – Representação do Grad-CAM aplicado na emoção surpresa nas bases (a) CK+, (b) FER2013, (c) RAF-DB e (d) MMI. . . . .	87
Figura 47 – Taxa de avaliações corretas dos participantes com TEA para as três fases. .	91
Figura 48 – Taxa de avaliações corretas dos participantes P5, P6, P7 e P8 sem TEA para as três fases. . . . .	92
Figura 49 – Taxa de avaliações corretas dos participantes P9, P10, P11 e P12 com TEA para as três fases. . . . .	94
Figura 50 – Taxa de avaliações corretas do participante P13, P14, P15 e P16 com TEA para as três fases. . . . .	95



---

## Lista de tabelas

Tabela 1	–	Princípios fundamentais para a aprendizagem por meio do GBL. . . . .	34
Tabela 2	–	Trabalhos realizados utilizando realidade aumentada, realidade virtual e jogos para auxiliar no tratamento do TEA ou deficiência intelectual . . . . .	52
Tabela 3	–	Resultados obtidos pelo método proposto nas bases: CK+, FER2013, RAF-DB e MMI. . . . .	83
Tabela 4	–	Resultados dos experimentos avaliando os parâmetros da CNN. . . . .	84
Tabela 5	–	Resultados obtidos nos experimentos para diferentes modelos de CNN: Modelo proposto, MobileNet, e Inception v3 na base CK+. . . . .	84
Tabela 6	–	Resultados obtidos com as métricas acurácia, precisão, especificidade e F1-Score para cada uma das emoções básicas com o modelo proposto na base de imagens CK+. . . . .	85
Tabela 7	–	Comparação de métodos usados pelos trabalhos relacionados com a base de imagens CK+. . . . .	87
Tabela 8	–	Dados dos participantes do cenário 1 investigado. . . . .	88
Tabela 9	–	Taxa de acerto (%) para os indivíduos investigados nesse cenário. . . . .	88
Tabela 10	–	Ocorrência do número de alerta que o aplicativo gerou para o usuário devido à falta de atenção. . . . .	89
Tabela 11	–	Total de erros em cada etapa para os participantes com TEA. . . . .	89
Tabela 12	–	Dados dos participantes do cenário 2 investigado. . . . .	90
Tabela 13	–	Resultados (%) para os indivíduos com TEA. . . . .	93
Tabela 14	–	Ocorrência do número de alerta que o aplicativo gerou para o usuário devido à falta de atenção. . . . .	96
Tabela 15	–	Total de erros em cada etapa para os participantes com TEA. . . . .	96





---

# Lista de siglas

**API** *Application Programming Interface*

**CG** Computação Gráfica

**EEG** Eletroencefalograma

**GBL** Games-Based Learning

**LOD** *level-of-detail*

**MLP** *Multi Layer Perceptron*

**PCA** *Principal Component Analysis*

**QI** Quociente de Inteligência

**RPG** *Role-Playing Game*

**RNA** Redes Neurais Artificiais

**SNA** Sistema Nervoso Autônomo

**TC** Tomografia Computadorizada

**UA** Unidades de Ação

**UM** Unidades de Movimento

---

# Sumário

<b>1</b>	<b>INTRODUÇÃO . . . . .</b>	<b>23</b>
<b>1.1</b>	<b>Motivação . . . . .</b>	<b>25</b>
<b>1.2</b>	<b>Objetivos e Desafios da Pesquisa . . . . .</b>	<b>27</b>
<b>1.3</b>	<b>Hipótese . . . . .</b>	<b>27</b>
<b>1.4</b>	<b>Contribuições . . . . .</b>	<b>28</b>
<b>1.5</b>	<b>Organização desta Tese . . . . .</b>	<b>28</b>
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA . . . . .</b>	<b>29</b>
<b>2.1</b>	<b>Transtorno do Espectro Autístico . . . . .</b>	<b>29</b>
<b>2.2</b>	<b>Teoria da Mente . . . . .</b>	<b>31</b>
<b>2.3</b>	<b>Emoções . . . . .</b>	<b>32</b>
<b>2.4</b>	<b>Aprendizagem Baseada em Jogos . . . . .</b>	<b>33</b>
<b>2.5</b>	<b>Modelagem Computacional . . . . .</b>	<b>35</b>
<b>2.6</b>	<b>Reconhecimento das Expressões Faciais . . . . .</b>	<b>37</b>
<b>2.7</b>	<b>Análise de Imagens Digitais . . . . .</b>	<b>40</b>
2.7.1	Detecção de Regiões Candidatas . . . . .	41
2.7.2	Extração de Características da Face . . . . .	42
<b>2.8</b>	<b>Técnicas de Inteligência Artificial . . . . .</b>	<b>44</b>
2.8.1	Aprendizagem Profunda . . . . .	45
<b>2.9</b>	<b>Considerações Finais . . . . .</b>	<b>50</b>
<b>3</b>	<b>ESTADO DA ARTE . . . . .</b>	<b>51</b>
<b>3.1</b>	<b>Ensino de Competências Emocionais em Indivíduos com TEA . . . . .</b>	<b>51</b>
<b>3.2</b>	<b>Considerações Finais . . . . .</b>	<b>57</b>
<b>4</b>	<b>METODOLOGIA . . . . .</b>	<b>59</b>
<b>4.1</b>	<b>Projeto do Jogo . . . . .</b>	<b>59</b>
4.1.1	Conceito . . . . .	59

4.1.2	Pré-produção . . . . .	60
4.1.3	Protótipo . . . . .	60
4.1.4	Produção . . . . .	61
4.1.5	Modelagem Computacional . . . . .	62
<b>4.2</b>	<b>Jogo para detecção de emoções baseadas em expressões faciais . . . . .</b>	<b>64</b>
4.2.1	Módulo Detecção e Classificação . . . . .	70
4.2.2	Módulo Dados . . . . .	77
<b>4.3</b>	<b>Métricas de Avaliação . . . . .</b>	<b>79</b>
<b>5</b>	<b>RESULTADOS E DISCUSSÃO . . . . .</b>	<b>81</b>
<b>5.1</b>	<b>Avaliação dos Algoritmos de Detecção e Reconhecimento de Emoções em Bancos de Dados de Domínio Público . . . . .</b>	<b>81</b>
<b>5.2</b>	<b>Avaliando o SG <i>Face Emotions</i> com os Indivíduos . . . . .</b>	<b>87</b>
5.2.1	Cenário 1: Associação dos Deficientes da Liberdade . . . . .	87
5.2.2	Cenário 2: Clínica de Terapia Ocupacional . . . . .	90
<b>6</b>	<b>CONCLUSÃO . . . . .</b>	<b>99</b>
<b>6.1</b>	<b>Conclusões . . . . .</b>	<b>99</b>
<b>6.2</b>	<b>Trabalhos Futuros . . . . .</b>	<b>100</b>
<b>6.3</b>	<b>Produções Bibliográficas Obtidas no Doutorado . . . . .</b>	<b>101</b>
	<b>REFERÊNCIAS . . . . .</b>	<b>105</b>

---

## Introdução

A capacidade de reconhecer estados emocionais com base nas expressões faciais tem contribuído para interações sociais e comunicação interpessoais. As emoções - alegria, tristeza, raiva, medo, desgosto, surpresa e neutra - são expressões faciais naturais dos seres humanos, as quais permitem demonstrar os sentimentos. O reconhecimento e a manifestação dessas expressões são essenciais para as interações sociais entre as pessoas. Um dos primeiros estudos sobre esses estados emocionais foram explorados por Darwin e Garcia (DARWIN; GARCIA, 2000), em que investigaram como reconhecer as expressões faciais dos seres humanos e sua influência nas interações sociais.

Em 2002, surgiu uma teoria denominada Sistema de Codificação de Ação Facial (do inglês, *Facial Action Coding System* - FACS), que permite tratar as emoções como uma representação universal entre os seres humanos de diferentes grupos sociais e etnias (EKMAN; FRIESEN; HAGER, 2002). Essa teoria permite descrever todos os movimentos faciais por meio de quarenta e quatro representações denominadas Unidades de Ação (UA). Cada UA tem um código numérico que mapeia o grupo de músculos da face para formar diferentes mímicas faciais relacionadas às emoções.

O transtorno do espectro autista (TEA) é um transtorno do neurodesenvolvimento comum e grave que envolve dificuldades na comunicação não verbal e na interação social dos indivíduos. Esses sintomas estão presentes no início do desenvolvimento e afetam o funcionamento diário. O termo espectro é comumente usado devido à heterogeneidade na apresentação e gravidade dos sintomas, bem como nas habilidades e nível de funcionamento desses indivíduos (BLACK; GRANT, 2014). Estudos realizados por Lorna Wing e Judith Gould sobre a epidemiologia do autismo e condições relacionadas levaram ao conceito de TEA em 1979 (BOSA, 2002). Fatores genéticos e variáveis ambientais também podem contribuir para o surgimento ou agravamento de características em indivíduos com TEA (SANDIN et al., 2014).

A prevalência do Transtorno do Espectro Autista (TEA) é um tema polêmico na sociedade moderna. De acordo com dados estatísticos publicados pela Organização Mundial da Saúde (OMS) (OMS, 2021), uma em cada 160 crianças apresenta características que estão relacionadas ao TEA. Essa estimativa representa um valor médio e a prevalência relatada varia substancialmente.

almente entre os estudos. Alguns estudos bem controlados, no entanto, relataram números que são substancialmente mais altos. Um indivíduo com TEA pode apresentar limitações relacionadas a relação interpessoal, comunicação social e imaginação (HEWITT, 2006). A principal causa está relacionada a capacidade de reconhecer e expressar as emoções, o que pode comprometer de forma negativa as interações sociais (BARON-COHEN et al., 2005).

Proporcionar o aprimoramento das habilidades para o reconhecimento e expressão de emoções, com soluções que explorem cenários cotidianos do mundo real, por meio de estratégias computacionais pode ser relevante, conforme relato em (MOTA, 2012). Se considerar a capacidade cognitiva dos indivíduos, essas soluções podem favorecer aprendizado, explorando as competências e aumentando a motivação do indivíduo por meio de recursos multimídia. Além disso, essas soluções possibilitam obter dados estatísticos importantes para compreender o progresso e a evolução do indivíduo durante a fase de um tratamento, tais como: depressões, ansiedade, distúrbios sociais e do neuro desenvolvimento.

Diversas pesquisas foram realizadas com o objetivo de investigar com mais profundidade os aspectos cognitivos das habilidades de emoções em indivíduos com TEA. Conforme relato dos autores em (DAPOGNY et al., 2018), o ensino das emoções para pessoas com TEA é um grande desafio, pois as respostas sobre essas habilidades são limitadas e provocam dificuldades para a interação social. Uma das principais causas desse problema, é dada pela diferença das intensidades cerebrais e das regiões ativadas durante o processo de reconhecimento e expressão das emoções. Além disso, pode ocorrer a existência de anormalidades nas amígdalas, o que contribui de forma negativa para as interações sociais, conforme mostram os trabalhos de (DAPRETTO et al., 2006; BARON-COHEN et al., 2005; QIN et al., 2020; SEGUIN et al., 2021). De acordo com uma pesquisa publicada em (BAPTISTA; BOSA, 2002), o desenvolvimento intelectual de pessoas com TEA deve ocorrer de forma especial em relação a grande parte das pessoas. Normalmente, esses indivíduos têm dificuldade no processo de iniciar e manter um relacionamento social com outras pessoas.

Diante da importância dessas competências, várias ferramentas são propostas para auxiliar os especialistas no ensino dessas emoções aos indivíduos com TEA. Em uma revisão apresentada por (GROSSARD et al., 2017), soluções baseadas em jogos sérios foram identificadas com o objetivo de ensinar habilidades para pessoas com TEA. Essas soluções são empregadas em intervenções por meio de fotografias planas ou desenhos esquemáticos, explorando as habilidades em representações das situações do cotidiano, desejos e crenças (THIBES-UEM; ALENCAR-UEM; AZEVEDO-UEM, 2009; ALMEIDA et al., 2019).

Os avanços tecnológicos têm provocado o surgimento de diversas propostas com objetivo de melhorar as interações sociais no reconhecimento de emoções (METRI; GHORPADE; BUTALIA, 2012). Nesse caso, ferramentas computacionais são desenvolvidas explorando recursos multimídia, inteligência artificial, figuras planas ou jogos (TANG et al., 2021). Na literatura, há uma quantidade de jogos sérios desenvolvidos para auxiliar no aprimoramento das habilidades emocionais e sociais dos indivíduos com TEA conforme relatos nos trabalhos de (SILVER; OA-

KES, 2001; FRIDENSON-HAYO et al., 2017; DAPOGNY et al., 2018; BOLTE et al., 2002; ALMEIDA et al., 2019). Apesar de relevantes contribuições, a maioria dessas soluções oferece suporte ao aprimoramento de apenas uma das habilidades com protocolos não personalizados a cada indivíduo. Outras limitações e desafios são destacados nesses trabalhos: tamanho das amostras, a dificuldade de convocação desse público, dependência de *hardware* ou auxílio de terceiros para uso da ferramenta, relatórios insuficientes ou inexistentes, o que não permite o acompanhamento da evolução dos indivíduos de maneira personalizável às habilidades. Também é importante considerar as restrições relacionadas às características de um indivíduo com TEA quanto a resistência à mudança e imprevisibilidade, a qual pode ocorrer durante as intervenções (MALINVERNI et al., 2017). Com isso, esses desafios têm motivado a criação de novos métodos de tratamento e ferramentas computacionais que possam ajudar os profissionais no processo de tratamento de pessoas com TEA.

## 1.1 Motivação

Em 2020, um documento publicado pelos Centros de Controle e Prevenção de Doenças dos EUA relatou um aumento de 178% no número de crianças diagnosticadas com TEA, em comparação com dados publicados em 2000 (MAENNER et al., 2020). No Brasil, com uma população de mais de 200 milhões, estima-se que existam 2 milhões de indivíduos com TEA (CONTROL; PREVENTION, 2018).

A capacidade de reconhecer os estados mentais das expressões faciais desempenha um papel vital na interação social e na comunicação interpessoal. Entre as principais características estão padrões comportamentais repetitivos, deficiências em termos de interação social e comunicação verbal e não verbal e falta de reciprocidade socio-emocional, presente desde a infância (SCASSELLATI; ADMONI; MATARIĆ, 2012; LIMA et al., 2019). Indivíduos com TEA muitas vezes enfrentam dificuldades no desenvolvimento de habilidades relacionadas à emoção em ambientes colaborativos. Eles têm problemas de longo prazo ao participar de atividades como trabalho em equipe ou falar em público (ELSHAHAWY; ABOELNAGA; SHARAF, 2020; RIBU, 2010). A representação das expressões são mais fáceis quando indivíduos com TEA são motivados a produzi-las, em vez de esperar que essas representações ocorram espontaneamente (TREVISAN; HOSKYN; BIRMINGHAM, 2018).

De acordo com os autores de (GROSSMAN; EDELSON; TAGER-FLUSBERG, 2013), as maneiras pelas quais os indivíduos com TEA produzem as emoções por meio das expressões faciais podem influenciar suas interações sociais no mundo real. Os métodos tradicionais de ensino e interações em grupo podem não ser adequados para esta atividade, conforme descrito em (PENNINGTON; CARPENTER, 2019). Nos últimos anos, vários estudos têm sido realizados com o objetivo de apresentar novas soluções baseadas no uso de tecnologias de informação e comunicação para o ensino e aprimoramento das emoções de forma individualizada (SAGAYARAJ; GOPAL; KARTHIKEYAN, 2020; MANTA et al., 2020; STRICKROTH et al.,

2020; PENNINGTON; CARPENTER, 2019). Diferentes algoritmos de visão computacional têm sido empregados para prever informações de imagem da face humana (BASTANFARD; TAKAHASHI; NAKAJIMA, 2004). Várias estratégias para melhorar o reconhecimento de expressões faciais têm sido aplicadas para melhorar essas habilidades em indivíduos com TEA (GROSSARD et al., 2019; DAPOGNY et al., 2018; LIMA et al., 2019).

Os esforços para desenvolver essas ferramentas vêm aumentando, conforme dados apresentados em pesquisas sobre o panorama global dessas soluções (LIMA et al., 2019). Essa estratégia também está sendo empregada em outros domínios. Crianças com deficiência auditiva que têm uma compreensão limitada de conceitos concretos e especialmente abstratos na vida cotidiana têm problemas para interagir com o mundo exterior (BASTANFARD et al., 2010). A utilização de metodologias educativas com conteúdos multimídia é considerada uma estratégia que facilita o processo de melhoria das competências. Uma das principais vantagens do uso dessa estratégia é a construção de ambientes de jogos digitais. As crianças demonstraram ter um alto nível de interesse em aplicativos digitais, sendo, portanto, uma opção atraente em termos de habilidades de ensino (BOUCENNA et al., 2014). De acordo com (CONSTANTIN et al., 2017; BLACK; GRANT, 2014), indivíduos com TEA possuem afinidade com a tecnologia e seus interesses especiais podem ser receptivos a recompensas digitais inseridas no contexto do jogo.

O tratamento de indivíduos com TEA, quando realizado nos primeiros anos de vida, dão ao indivíduo a possibilidade de melhora no grau clínico e na qualidade de vida (GROSSARD et al., 2019; DAPOGNY et al., 2018; LIMA et al., 2019). O desenvolvimento de técnicas computacionais pode contribuir para o aprimoramento da representação dos sentimentos e das relações sociais (GROSSARD et al., 2019; DAPOGNY et al., 2018). Essas estratégias fornecem características particulares e relevantes que são difíceis de ser observadas em tratamentos convencionais. Outra característica é a possibilidade de resolver uma série de limitações existentes no processo manual de análise das informações obtidas no período de tratamento, tais como: registro automático dos dados, recuperação e visualização das informações (BENKHERRAT, 2018). Além disso permitir investigar essas informações de forma quantitativa para aprimorar o tratamento.

Atualmente, parte dos métodos tradicionais empregados pelos especialistas exige uma grande carga de trabalho para análise das informações coletadas manualmente (BENKHERRAT, 2018). O desenvolvimento de ferramentas computacionais possibilita o aumento do número de intervenções durante o tratamento clínico. O desenvolvimento de abordagens com essas características permite que o indivíduo possa treinar e aperfeiçoar as habilidades em diferentes ambientes, não se restringindo apenas aos procedimentos nos consultórios. Na literatura há trabalhos que exploram jogos sérios para o ensino de emoções nos indivíduos com TEA (FRIDENSON-HAYO et al., 2017; ALMEIDA et al., 2019; MALINVERNI et al., 2017; CHEN; LEE; LIN, 2016; MAZUREK; ENGELHARDT; CLARK, 2015; TSANGOURI et al., 2016; STURM; PEPPE; PLOOG, 2016).



Com base nas informações relatadas, vale também destacar que essas ferramentas desenvolvidas são de domínio proprietário, sendo inexistentes em instituições públicas (TREVISAN et al., 2021; SILVA; ARTUSO; TORTATO, 2020). Além disso, parte dessas ferramentas não oferecem *feedback* no processo de ensino dessas emoções e não exploram outros recursos computacionais que podem contribuir no ensino de emoções. Então, o desenvolvimento de novas ferramentas computacionais, que permitam fornecer *feedback* das expressões faciais produzidas pelo indivíduo e quantificar o aprimoramento durante o tratamento, torna-se relevante para o cenário atual.

## 1.2 Objetivos e Desafios da Pesquisa

O presente trabalho traz o desenvolvimento de uma ferramenta computacional para auxiliar no treinamento e aprimoramento das habilidades emocionais da face de indivíduos com TEA. Pontualmente, pretende-se neste trabalho:

- ❑ Desenvolver uma ferramenta computacional composta por um jogo sério, com características voltadas às habilidades educacionais, para aprimorar a capacidade de reconhecer e expressar emoções em indivíduos com TEA;
- ❑ Desenvolver um método para detecção e reconhecimento de emoções baseados em um modelo de aprendizagem profunda e descritores manuais;
- ❑ Desenvolver um sistema para coleta de dados dos indivíduos com TEA para análise quantitativa do aprimoramento no ensino das emoções;
- ❑ Investigação do aprimoramento das habilidades emocionais em indivíduos com TEA em ambientes educacionais com uso da ferramenta computacional.

## 1.3 Hipótese

O ensino dos estados emocionais com base nas expressões faciais é complexo para os indivíduos com TEA. Existem ainda questões em aberto nessa área que permitam melhorar as relações intersociais. O emprego de técnicas computacionais baseadas em jogos pode auxiliar nas terapias de maneira positiva? As ferramentas computacionais, que forneçam dados quantitativos, podem colaborar com os especialistas no aprimoramento das habilidades e avanço nas terapias?

Com base nessas questões, este trabalho será desenvolvido, tendo como primícias o emprego de técnicas computacionais no desenvolvimento de jogos sérios que auxiliem no ensino e tratamento dos indivíduos com TEA. Outras questões secundárias importantes nesse processo, são:

- ❑ Quais métodos podem contribuir na detecção e análise da face e emoções? Esses métodos podem retornar um *feedback* para o aprimoramento dessas emoções?

## 1.4 Contribuições

As contribuições do trabalho proposto são:

- ❑ Uma nova abordagem para o reconhecimento e classificação das emoções *online* para indivíduos com TEA;
- ❑ O desenvolvimento de jogos sérios que possibilitem o ensino das emoções para indivíduos com TEA;
- ❑ Um sistema computacional para quantificação das sessões no processo de aprimoramento do ensino das emoções para especialistas.

## 1.5 Organização desta Tese

Neste capítulo, foram apresentadas as considerações iniciais, motivação e contribuições do desenvolvimento deste projeto. O restante deste trabalho possui a seguinte organização:

- ❑ Capítulo 2: Introduz os principais conceitos para a compreensão e validação do trabalho proposto;
- ❑ Capítulo 3: Apresenta os trabalhos correlatos relacionados com a proposta;
- ❑ Capítulo 4: Descreve o trabalho proposto, bem como as tecnologias, técnicas e metodologias desenvolvidas;
- ❑ Capítulo 5: Apresenta os resultados e as respectivas análises dos métodos computacionais para aprendizagem das emoções em indivíduos com TEA;
- ❑ Capítulo 6: São apresentados as conclusões e contribuições obtidas, considerações finais e trabalhos futuros.

---

## Fundamentação Teórica

Nesta seção são apresentados os conceitos teóricos necessários para a descrição dessa pesquisa. São tratados os principais assuntos relacionados ao ensino de emoções e às técnicas empregadas para a elaboração de jogos para aprendizagem das emoções.

### 2.1 Transtorno do Espectro Autístico

A síndrome, inicialmente, conhecida pelo termo “autismo”, em 1911, foi empregada para descrever a perda de contato com a realidade. Essa síndrome é muito abrangente, incluindo vários sintomas e com diversas variações clínicas. Além disso, apresenta um amplo grau de desenvolvimento e funcionamentos distintos para cada pessoa (KAMP-BECKER et al., 2010). Esses sintomas representam a incapacidade ou extrema dificuldade de comunicação num convívio social (AJURIAGUERRA; MARCELLI, 1986). Com base nessas observações, em 1979, Lorna Wing e Judith Gould (BOSA, 2002), definiram o termo Transtorno do Espectro Autista (TEA), para descrever o problema de neurodesenvolvimento, caracterizado por impedimentos relacionados à comunicação social, às dificuldades com interações sociais recíprocas e aos padrões incomuns de comportamentos repetitivos (MUÑOZ, 2018).

Para a maioria dos especialistas, esse transtorno envolve atrasos e comprometimentos das áreas de interação social e linguagem, que está relacionado à disfunção na cognição emocional e social. Também pode incluir limitações relacionadas a emoção, movimentação motora e sensorial de um indivíduo (GREENSPAN; WIEDER, 2006). As primeiras manifestações dos sintomas podem ser observadas antes dos 36 meses de idade em uma criança, mas podem persistir ao longo da vida, de maneira mais silenciosa (PLIMLEY, 2007).

Segundo os pesquisadores em (WEITLAUF et al., 2014), embora existam diferenças qualitativas como Quociente de Inteligência (QI), interações sociais, e comportamentais, os métodos quantitativos ou práticas para diferenciar os níveis do TEA ainda permanecem indeterminados. Os especialistas utilizam como padrão as recomendações do documento DSM-5, um manual de diagnóstico e estatístico de transtornos mentais (ASSOCIATION et al., 2014). Nesse documento, o TEA é categorizado com base em observações realizadas durante a investigação da

comunicação social, comportamentos restritos e repetitivos de cada indivíduo com TEA. Os níveis podem ser leves, ou nível 1, onde tem-se os indivíduos que requerem suporte; os moderados, ou nível 2, indivíduos que precisam de suporte substancial; e os severos, ou nível 3, são pessoas que requerem suporte muito substancial. Os sintomas e as características comportamentais de um indivíduo com TEA podem ser modificadas durante o ciclo de sua vida (KLIN, 2006).

A obtenção de um diagnóstico e a realização de tratamentos possibilitam uma melhor qualidade de vida para os indivíduos com TEA. As metodologias de análise na definição do diagnóstico e tratamento têm passado por uma série de transformações ao longo dos últimos anos. Para diagnosticar um indivíduo com TEA, atualmente, os especialistas usam critérios que analisam as capacidades sociais e comportamentais. Esses critérios exploram as alterações na capacidade de linguagem e interações sociais, e os mecanismos neurobiológicos (WEITLAUF et al., 2014). Essas informações possibilitam identificar deficiências nas expressões da face durante a representação das emoções, as posturas corporais e os gestos utilizados durante as interações sociais e as dificuldades de respostas com outras pessoas (BEGEER et al., 2008). Essas modificações podem ser observadas no trabalho de (DAPRETTO et al., 2006), o qual mostra um exemplo do comportamento da região do cérebro de um indivíduo com TEA. Na Figura 1 do estudo são apresentadas imagens obtidas com ressonância magnética das regiões do cérebro de diferentes indivíduos. Esse estudo mostra a diferença das atividades de um paciente típico e um paciente com TEA, ao reconhecer as expressões faciais durante uma interação social.

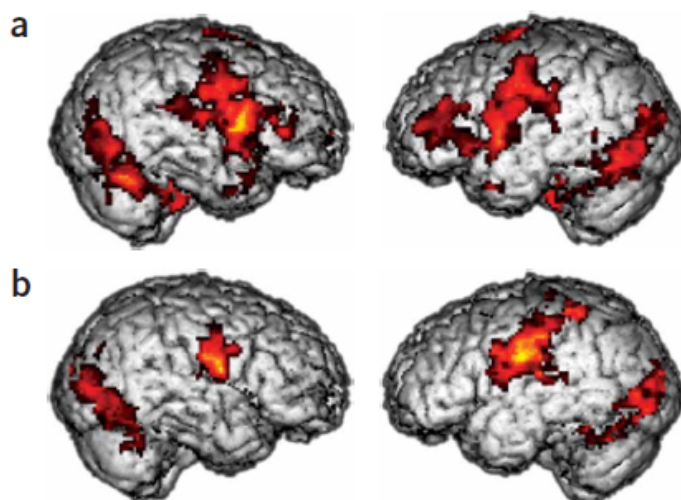


Figura 1 – Imagens obtidas com ressonância magnética das atividades cerebrais em pacientes típicos (a) e indivíduos com TEA (b). **Fonte:** (DAPRETTO et al., 2006).

As regiões demarcadas pelas cores, vermelha e amarela, mostram as diferenças de comportamento entre esses indivíduos. Na Figura 1 (a) são mostradas as atividades cerebrais em pessoas típicas, em que é possível notar uma maior ativação das regiões próximas da área frontal, o que não acontece nas imagens da Figura 1(b). Esse mesmo estudo mostra que é possível

melhorar as regiões afetadas na análise de expressões das emoções, por meio de técnicas de aprendizagem de emoções e auxílio de profissionais da área da psicologia.

Um tratamento para melhorar a qualidade de vida do indivíduo com TEA é aprimorar as atividades cerebrais. Isso deve-se iniciar o mais cedo possível e ser adaptado às necessidades específicas de cada um, de acordo com a criança e a família, de forma a contribuir a aprimorar as funcionalidades das crianças com TEA (ZABLOTSKY; BRADSHAW; STUART, 2013).

## 2.2 Teoria da Mente

Segundo (ASTINGTON; GOPNIK, 1988), a Teoria da Mente pode ser definida como a área que investiga a habilidade dos indivíduos em compreenderem seus próprios estados mentais e os de outra pessoa. Dessa maneira, podem prever suas ações ou comportamentos em indivíduos típicos, mas pode ter limitações em pessoas com TEA, esquizofrenia, pessoas com déficit de atenção, etc.

Em teoria da mente, tem-se como suposição que cada indivíduo tem uma mente e cada humano só pode intuir a existência de sua própria mente por meio da introspecção e ninguém tem acesso direto à mente de outra pessoa. Quando há referência a outro indivíduo há uma analogia com a própria mente, que se baseia na natureza recíproca da interação social, no uso funcional da linguagem e na compreensão das emoções e ações dos outros (PREMACK; WOODRUFF, 1978).

Essa teoria pode ser considerada uma habilidade potencialmente natural em humanos mas, para que ocorra, são necessárias experiências sociais durante muitos anos para ativá-la e aprimorá-la. Diferentes pessoas podem desenvolver teorias mais ou menos efetivas. Além disso, a empatia é um conceito que está diretamente relacionado, significando a capacidade de reconhecimento e compreensão dos estados mentais, incluindo crenças, desejos e particularmente as emoções dos outros, frequentemente denominada como a habilidade de "compreender o ponto de vista do outro" (JOU; SPERB, 1999).

De acordo com os autores (BARON-COHEN et al., 1985; BARON-COHEN, 2001), os indivíduos com TEA ou transtornos mentais, em geral, apresentam severas dificuldades em conceber estados mentais, ou seja, saber o que outras pessoas pensam, querem, sentem ou acreditam. Essas habilidades não aparecem antes do segundo ano de vida. Estudos realizados nos trabalhos (BARON-COHEN et al., 1985; BARON-COHEN, 2001) demonstraram que esses indivíduos tem severas dificuldades em empregar a teoria da mente e, conseqüentemente, uma incapacidade de representar estados mentais e, por isso, os indivíduos com TEA são incapazes de imputar crenças a outras pessoas e estão em grande desvantagem ao prever o comportamento de outros sujeitos.

Desenvolver essas habilidades cognitivas é extremamente importante para o desenvolvimento social dos indivíduos, pois são fundamentais para a criação e manutenção das interações sociais no cotidiano. Para contornar essa situação, a psicologia, em conjunto com especialis-

tas da educação e medicina, vem ao longo dos anos desenvolvendo terapias e metodologias para trabalhar essas habilidades com a intenção de proporcionar melhores condições de vida às pessoas, principalmente, os indivíduos com TEA (BARON-COHEN, 1991).

## 2.3 Emoções

As emoções podem ser interpretadas como uma reação tipicamente breve e intensa relacionada a um determinado fato ou evento e podem ser classificadas por alegria, tristeza, raiva, desgosto, surpresa e medo. De forma natural, as emoções podem surgir de alguma situação específica com a finalidade de preparar o organismo para reagir a um estímulo.

As emoções fazem parte do estado “humor”, a qual pode ser afetiva e difusa e com duração indeterminada (PERGHER et al., 2006). Diferente da emoção, o estado humor engloba uma série de estratégias cognitivas de longo prazo que podem ser observadas por inúmeros eventos. Geralmente, esse estado determina quais emoções uma pessoa está mais propícia a inferir durante um determinado evento em relacionamentos interpessoais. Um exemplo da diferença entre emoção e humor pode ser observado em uma pessoa, quando ela é classificada como um indivíduo triste, se os músculos de sua face representam a emoção tristeza, mas se esse indivíduo está deprimido, tem-se uma representação de seu estado de humor.

O uso de tecnologias tem contribuído na análise das emoções em determinados casos. Na literatura, existem técnicas para reconhecer as emoções por meio de algoritmos computacionais, sensores e sinais biológicos. Conforme descrito em (NASS; BRAVE, 2007), as principais estratégias para analisar as emoções são:

- ❑ **Respostas neurológicas:** análise de sinais de Eletroencefalograma (EEG) da região do cérebro pode ser empregado em estudos que investigam as emoções. Essa é a primeira parte do ser humano a ser estimulada para ativação dos músculos da face na representação das emoções;
- ❑ **Atividades autonômicas:** alguns componentes do Sistema Nervoso Autônomo (SNA) podem ser empregados, tais como, frequência cardíaca, pressão cardíaca, respiração, dilatação das pupilas, tensão muscular, etc;
- ❑ **Expressões faciais:** uma das formas menos invasivas para reconhecer as representações presentes das emoções por meio da análise dos músculos presentes na face;
- ❑ **Voz:** a entonação da voz e a velocidade da fala, podem ser utilizadas como características na representação das emoções.

Estudos recentes trazem diversas contribuições em relação ao uso de técnicas para análise das emoções das pessoas. Dentre as principais abordagens, destaca-se o trabalho de Konar e Chakraborty (KONAR; CHAKRABORTY, 2014) por investigarem e mostrarem estratégias para a detecção das emoções. No entanto, os autores apontaram que as expressões faciais

são capazes de proporcionar resultados mais expressivos no processo de reconhecimento das emoções. Além disso, o uso das expressões tem como outra vantagem ser um método menos invasivo, o que permite ao indivíduo um conforto maior no processo de sessões de terapia com o psicólogo.

A utilização de jogos para aprendizagem estão sendo explorados constantemente (LI et al., 2012; MONSALVE; WERNECK; LEITE, 2010), essa metodologia é capaz de proporcionar um ambiente que ensina e ao mesmo tempo é capaz de entreter os usuários por meio da aprendizagem baseada em jogos.

## 2.4 Aprendizagem Baseada em Jogos

A constante evolução das ferramentas de aprendizagem tem tornado acessível o emprego de jogos para o contexto educacional (MONSALVE; WERNECK; LEITE, 2010). Os jogos são, normalmente, utilizados como um material de apoio para as aulas, permitindo melhorar o processo de ensino, entendimento e visualização de elementos abstratos de difícil entendimento, usados pelos métodos tradicionais. Essa nova abordagem denominada aprendizagem baseada em jogos, que vem do termo em inglês *Games-Based Learning (GBL)*, permite empregar jogos para melhorar as habilidades de aprendizagem.

Ensinar, utilizando jogos, permite desenvolver atividades lúdicas que motivam os estudantes durante o processo de aprendizagem, trazendo benefícios para a tarefa educacional. Os jogos possuem um efeito de estimulação maior do que outras técnicas, pois se adaptam bem à diversas situações do mundo real, assim como inserem o estudante em diversos contextos (LI et al., 2012). As informações fornecidas por meio de um *feedback* dos usuários podem ser inseridas no aprimoramento dessas ferramentas. Um exemplo dessa abordagem foi apresentado por *Garrist et al.* (GARRIS; AHLERS; DRISKELL, 2002) para um contexto semelhante à imagem na Figura 2. Nesse caso, a entrada inclui os materiais educacionais e as características do jogo mostram os elementos utilizados no desenvolvimento, tais como: imagens, animações, vídeos, etc. Após o processamento dos ciclos internos, incluindo o *feedback*, julgamento e comportamento do usuário, os resultados mostram o aprendizado adquirido.

Segundo (COLEMAN; MONEY, 2020), o GBL permite uma aprendizagem mais eficaz por explorar três conceitos envolvidos: capacitação dos estudantes, resolução de problemas e a compreensão, conforme apresentado na Tabela 1. Na capacitação do estudante, o princípio *co – design*, é relevante para passar aos jogadores a sensibilidade de que suas escolhas são as principais características por trás da experiência que eles estão tendo ao interagir com o GBL. A personalização permite aos jogadores a capacidade de tomar decisões sobre o modo como joga uma partida, permitindo assim aos usuários criarem a identidade do jogador por meio de incorporação. Dessa maneira, as escolhas realizadas durante o jogo tornam-se pessoalmente relevantes para os jogadores, aumentando seu nível de interesse durante as partidas.

O princípio dos problemas ordenados, conforme a Tabela 1, pertencente à solução de pro-

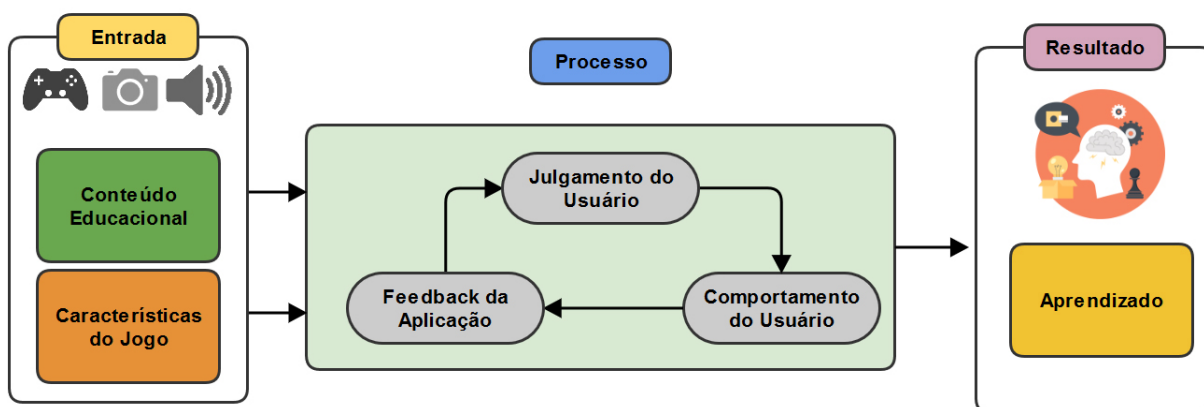


Figura 2 – Diagrama de fluxo de uma aplicação GBL. **Fonte:** Adaptado de (LI et al., 2012).

Tabela 1 – Princípios fundamentais para a aprendizagem por meio do GBL. **Fonte:** (COLEMAN; MONEY, 2020)

Conceito	Princípio
Capacitação dos estudantes	<i>Co-design</i> Customização Identidade do jogador Manipulação
Solução de problemas	Problemas ordenados Agradavelmente frustrante Ciclos de especialização Fornecimento de informações Aprendizagem de aquário Aprendizado em <i>sandbox</i> Habilidades como estratégias
Compreensão	Sistemas a pensamento Significado da experiência

blemas, diz respeito ao fato de que as pessoas geralmente perdem o interesse se confrontadas com algo muito além de sua capacidade de solucionar, devido a sua complexidade ou à falta de entendimento. Segundo (COLEMAN; MONEY, 2020), os problemas considerados ideias para os usuários são aqueles que estão no limite do conforto e elaborados de maneira agradavelmente frustrante. Por último, a compreensão ou entendimento é composto por dois princípios: o pensamento sistêmico, que se relaciona às habilidades e ideias que o jogo transmite e o significado da experiência, que propõe que o aprendizado dos usuários ocorra de maneira mais significativa, quando filtrado pelas experiências que um aluno teve.

Os GBL possuem vantagens e desvantagens em relação ao uso no processo educacional. Uma das características de limitação é a acessibilidade aos recursos tecnológicos e a possibilidade de perda de foco dos usuários, se não forem utilizados adequadamente, conforme apresentado por (HWANG et al., 2012). No entanto, Prensky (PRENSKY, 2003) mostra as vantagens dos jogos digitais no ensino, que são capazes de proporcionar melhora na atenção seletiva visual do usuário, podendo ser aliada do profissional de educação no processo de aprendizagem. Essas



ferramentas, que exploram tecnologias digitais para a construção de jogos mais interativos têm contribuído para o processo de aprendizagem de indivíduos com TEA (TANG; HANNEGHAN; RHALIBI, 2009),

## 2.5 Modelagem Computacional

Na definição de personagens em jogos sempre é necessária a elaboração de modelos computacionais. Uma das áreas mais exploradas para a elaboração desses modelos é a Computação Gráfica (CG). Essa área define o uso de computadores ou dispositivos eletrônicos para criação e manipulação de imagens vetoriais por meio de algoritmos e técnicas. A CG está em constante evolução e possui terminologias e conceitos padrões que facilitam seu estudo (MARSCHNER; SHIRLEY, 2015). Na verdade, definir essa área é extremamente complexo, mas a maioria dos pesquisadores e profissionais categorizam como:

- ❑ **Modelagem:** Trabalha com a representação matemática e representação do objeto e sua aparência, de forma que essas representações possam ser armazenadas nos computadores. Um exemplo são os personagens em um jogo que pode ser representada como um conjunto de pontos 3D ordenados por uma regra de interpolação (MARSCHNER; SHIRLEY, 2015);
- ❑ **Renderização:** Assim como na área de artes, a renderização está relacionada com a criação de sombreamento da imagem para objetos 3D, convertendo símbolos gráficos em um resultado visual, seja por meio de imagens ou vídeos;
- ❑ **Animação:** Técnica utilizada para criar a ilusão de movimento nos objetos por meio de uma sequência de imagens. A animação usa elementos de modelagem, renderização e movimento em função do tempo.

A CG é composta por várias áreas da ciência da computação, sendo difícil reuni-las em um único trabalho devido aos seus conceitos e particularidades. Com isso, são apresentadas as principais áreas da CG exploradas nesta pesquisa:

- ❑ **Interação com o usuário:** trabalha com a interface entre os dispositivos de entrada, dos quais tem-se: o microfone, as câmeras, o teclado e *feedback* visual de imagens;
- ❑ **Realidade Virtual:** essa técnica tenta imergir o usuário em um mundo virtual 3D. Esse tipo de recurso requer um dispositivo gráfico e um equipamento que forneça resposta do movimento da cabeça e um *hardware* que tenha relevante desempenho de capacidade de processamento gráfico;
- ❑ **Visualização:** fornece aos usuários uma visão complexa de informações por meio da exibição visual.

Atualmente, a CG vem sendo empregada em diversas áreas do conhecimento, como:

- ❑ **Videogames:** A cada dia os jogos digitais possuem mais gráficos 3D e algoritmos de renderização mais aperfeiçoados;
- ❑ **Desenhos animados:** atualmente, os desenhos animados são renderizados a partir de modelos 3D, permitindo assim um ponto de vista móvel contínuo, economizando tempo do artista;
- ❑ **CAD/CAM:** em *Computer-aided design* e *computer-aided manufacturing*, a CG é utilizada para orientar a produção de peças, componentes e acessórios durante a elaboração de equipamentos. Um exemplo, é na produção de peças de carros projetadas no computador;
- ❑ **Simulação:** em aplicações que propõem situações inviáveis ou altamente perigosas fisicamente, como por exemplo, uma situação de incêndio. A CG permite modelar detalhes semelhantes ao mundo real em situação de perigo;
- ❑ **Imagens Médicas:** criam imagens importantes para a área médica. Um exemplo é uma imagem de Tomografia Computadorizada (TC), a qual é composta por uma grande matriz retangular 3D contendo valores de densidades ópticas daquelas regiões;
- ❑ **Visualização de Informações:** a computação visual auxilia áreas como a matemática e a estatística a criarem uma visualização de dados que necessariamente não possuem uma representação visual, como por exemplo, a tendência temporal de uma bolsa de valores.

As técnicas empregadas em CG podem influenciar no desempenho em um jogo, como por exemplo, motor gráfico, *hardware*, mecânica do jogo e jogabilidade (MARSCHNER; SHIRLEY, 2015). Em relação às plataformas, há uma série de dispositivos que podem comprometer o desempenho de um jogo: *hardware*, *software*, sistema operacional e *Application Programming Interface* (API).

Antigamente, os jogos eram projetados apenas para uma plataforma específica, mas com o passar do tempo e, devido à grande evolução dos dispositivos de reprodução, o desenvolvimento de jogos multiplataformas se tornaram realidade. Uma das vantagens de desenvolver jogo multiplataforma está relacionada ao seu potencial de base de clientes, uma vez que muito mais usuários podem ter acesso ao jogo. Algumas dessas plataformas não são bem definidas, como por exemplo, em uma aplicação para sistemas operacionais Windows, o desenvolvedor deve considerar uma grande variedade de possíveis configurações de *hardware*. Outra característica é que se espera que a aplicação execute bem em novas configurações que serão desenvolvidas no futuro.

Uma solução para esse problema é o uso de interface de programa de aplicação (do inglês, *Application programming interface* - API) devidamente implementada que faz o escalonamento dos recursos disponíveis com o objetivo de garantir um desempenho razoável em sistemas de

baixo custo. Assim, pode fornecer um alto nível de qualidade visual e processamento em sistemas de alto desempenho. Nessa estratégia, os ajustes são realizados automaticamente pelo próprio sistema, ao fazer uma análise dos equipamentos disponíveis, mas o usuário também pode defini-los por meio de uma interface para essa finalidade.

Um outro ponto relevante em soluções de CG são os gráficos dos jogos que têm como principal problema gerenciar as limitações de recursos. Cada plataforma impõe suas próprias restrições de acesso ao *hardware*, tempo de processamento, armazenamento e largura de banda. Para contornar essas restrições, os programadores utilizam diversas técnicas de otimização, como por exemplo, o *Pixel Shader*, que permite trabalhar na estrutura dos *pixels* após um modelo 3D ser rasterizado, possibilitando criar efeitos na imagem. Exemplos de uso do *Pixel Shader* na simulação de rugosidade e desfoque apenas podem ocorrer devido à sua execução diretamente sobre um processador gráfico, o que possibilita o uso de placas gráficas (do inglês, *graphics processing unit* - GPU) para os cálculos matemáticos e lógicos relacionados ao jogo. (MARSCHNER; SHIRLEY, 2015).

Na etapa de renderização de objeto também há desafios a serem tratados durante o desenvolvimento de jogos, em que a maneira mais rápida de explorar essa técnica é não renderizar. Uma técnica para determinar quando um objeto deve ser ou não renderizado pode ser uma boa proposta que permita economizar o processamento de *pixels*, processamento de vértices e tempo de CPU. A técnica *Frustum Culling* é muito explorada no desenvolvimento de jogos para fazer a seleção dos elementos que devem ser renderizados. No entanto, somente essa técnica não permite uma boa representação dos objetos. Para tanto, há algoritmos de remoção de oclusão de alto nível que empregam estruturas de dados como *Potentially Visible Sets* (PVS) e *Binary spatial partitioning* (BSP). Algoritmos como o *level-of-detail* (LOD) renderizam diferentes representações do objeto baseado em sua distância (cobertura da tela) economizando assim um processamento significativo (MARSCHNER; SHIRLEY, 2015). Na Figura 3 são apresentadas ilustrações do uso desse algoritmo.

O processamento gráfico também pode ser utilizado para armazenar dados de cada quadro e carregá-los antes de sua inicialização em um processo de renderização. Um exemplo são as informações de iluminação através de algoritmos, que calculam a iluminação em toda a cena e as armazenam em estruturas de dados como o *lightmaps*. Essas informações são relevantes no processo de desenvolvimento de uma plataforma de jogos para o processo de aprendizagem das emoções em indivíduos com TEA.

## 2.6 Reconhecimento das Expressões Faciais

Charles Robert Darwin, 1872, foi um dos primeiros naturalistas a estudar e trabalhar com o reconhecimento facial das emoções em seres humanos como fator social, dentro da perspectiva evolucionista (DARWIN, 1965). As emoções podem ser tratadas como eventos descritos ou inferidos por outros, com observações ou evidências do comportamento.

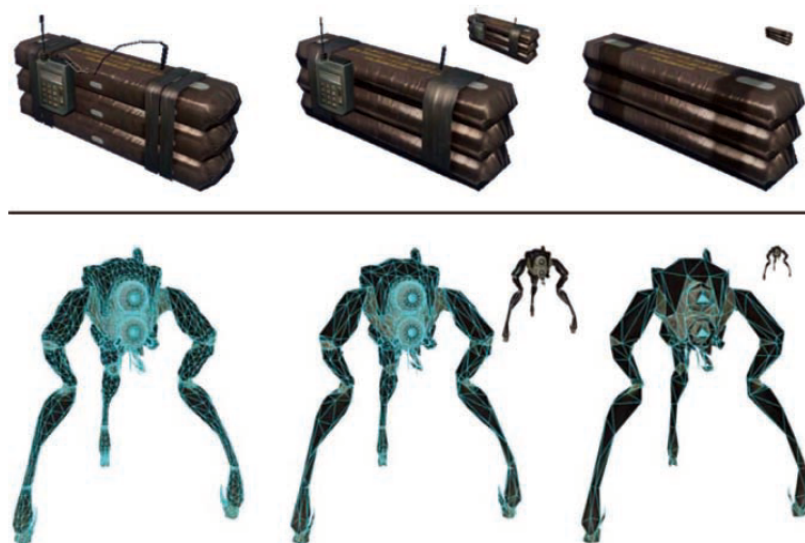


Figura 3 – Exemplos da variação de detalhes em objetos 3D de acordo com a distância, usando o algoritmo LOD. **Fonte:** (MARSCHNER; SHIRLEY, 2015).

Darwin, em suas observações e experimentos, defendeu que as emoções básicas são frutos de um processo evolutivo e que esse processo permitiu a sobrevivência da nossa espécie, expressando a reação adequada para cada tipo de situação. Por exemplo, o homem, ao confrontar um animal que colocasse sua vida em risco, a reação de medo não apenas dava ao homem primitivo a necessidade de fugir e esconder, mas também gerava as descargas hormonais de adrenalina e movimento do fluxo sanguíneo para que os membros inferiores auxiliem nesta tarefa, para garantir sua sobrevivência. Essa ideia inicial permitiu o desenvolvimento de um padrão de reconhecimento para auxiliar os seres humanos (GARCIA, 2009). As emoções quando combinadas com os sentimentos, constituem os estados emocionais. Os estudos apresentados por Darwin contribuíram para o reconhecimento e análise dos estados mentais, mas não ocorreu uma observação intercultural e sistemática nessas análises. Em seus experimentos, Darwin procurou demonstrar o funcionamento das expressões faciais como a principal manifestação das emoções. Então, Darwin viajou para diversos países com diferentes populações para realizar uma série de observações. Na maioria das observações utilizou fotos e gravuras para exemplificar as contrações musculares da face dessas populações. Nesses experimentos foram também observadas as posturas corporais que denotam determinadas emoções (GARCIA, 2009).

Com o objetivo de explorar novas pesquisas com uma abordagem mais sistemática transcultural surgiu o trabalho de Ekman e Friesen (EKMAN; FRIESEN, 1978). Ekman e Friesen realizaram um estudo na Nova Guiné, na Oceania, em 1968, com uma civilização isolada e sem a influência da cultura globalizada, com a finalidade de evitar influências sociais na formação e expressão facial das emoções básicas. Cada indivíduo foi fotografado e filmado sem que suas emoções sofressem alguma interferência, pois desconheciam as funções de uma câmera. Nesse experimento, sua pesquisa consistia em contar uma história a uma das pessoas dessa sociedade e, em seguida, mostrar fotos de expressões faciais, perguntando qual dos rostos a pessoa

representaria se estivesse naquela situação.

Ekman e Friesen constataram, em sua pesquisa, que a expressão facial de seis emoções básicas (alegria, tristeza, medo, surpresa, aversão e raiva) eram as mesmas para toda espécie humana, independentemente de sua cultura. Isso permitiu a criação da teoria denominada *Facial Action Coding System* (FACS) (EKMAN; FRIESEN, 1978), a qual demonstra que as emoções são consideradas universais em diferentes grupos sociais e etnias. A teoria FACS (EKMAN; FRIESEN, 1978) é um sistema com base nos músculos faciais, que permite descrever todos os movimentos por meio de quarenta e quatro Unidades de Ação (UA). Cada UA possui um código numérico para identificar os respectivos músculos contraídos para cada movimento da face. Um exemplo dessa representação é a denominada UA-1, que permite representar o levantamento da sobrancelha interna durante a representação de uma emoção. As representações desenvolvidas por Ekman com as UA são exibidas na Figura 4.































Unidades de ação da face superior					
AU 1	AU 2	AU 4	AU 5	AU 6	AU 7
 Levantador de sobrancelha <b>Levantador</b>	 Sobrancelha <b>Levantador</b>	 Testa <b>Inferior</b>	 Tampa superior <b>Levantador</b>	 Bochecha <b>Levantador</b>	 Apertador de Pálpebra
<b>*AU 41</b>	<b>*AU 42</b>	<b>*AU 43</b>	<b>AU 44</b>	<b>AU 45</b>	<b>AU 46</b>
 Inclinação da tampa	 Fenda	 Olhos fechados	 Olhos Semicerrados	 Piscar	 Piscadela
Unidades de ação da face inferior					
AU 9	AU 10	AU 11	AU 12	AU 13	AU 14
 Enrugador do nariz	 Levantador de lábio superior	 Aprofundador Nasolabial	 Extrator de canto labial	 Inchador de Bochecha	 Fazedor de Covinhas
AU 15	AU 16	AU 17	AU 18	AU 20	AU 22
 Depressor de canto labial	 Depressor de lábio inferior	 Levantador de queixo	 Franzedor de Lábio	 Esticador de Lábio	 Afunilador de Lábio
AU 23	AU 24	*AU 25	*AU 26	*AU 27	AU 28
 Endurecedor de Lábio	 Pressor labial	 Separador de Lábios	 Queda de Mandíbula	 Alongamento da boca	 Sucção de Lábios

Figura 4 – UA identificadas no trabalho de (EKMAN; FRIESEN, 1978).

Na Figura 4 é possível observar que cada um dos movimentos faciais foi bem definido e dividido em duas categorias para facilitar o entendimento: as UA da face superior e as UA da face inferior. Além disso, há uma combinação desses movimentos faciais que determinam as emoções por meio das expressões.

A Figura 5 demonstra um outro exemplo do uso das expressões faciais frontais extraídas de imagens para representação da emoção surpresa, obtidas do banco de dados Cohn e Kanade (KANADE; COHN, 2005). Nesse banco há uma sequência para representação de cada emoção em que no momento inicial tem-se a expressão neutra e evolui para a representação da emoção alvo.

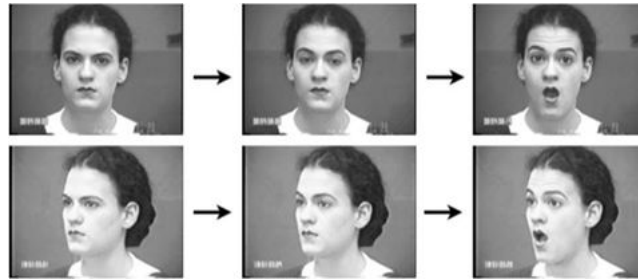


Figura 5 – Sequência de imagens para representação da emoção surpresa, extraída da base (KANADE; COHN, 2005) com as combinações das UA 1, 2, 5 e 27.

Os eventos faciais codificados pelo sistema FACS podem ser classificados em categorias emocionais e não-emocionais. Com base neste modelo, surgiu o *EMotion FACS* (EMFACS). Esse sistema possibilita a classificação dos movimentos da face em um sistema FACS (EKMAN; FRIESEN, 1978). Nesse modelo há um número menor de combinações de UA e uma versão modificada de regras e procedimentos de medição. Isso ocorre para diminuir o tempo de resposta quando o interesse é somente classificar as emoções inferidas pelas expressões faciais. Na Figura 5 é demonstrado uma combinação da codificação desses movimentos faciais que permite inferir a emoção expressada.

Esses métodos são empregados na detecção de pontos da face para construção de soluções que possam obter interação e comunicação homem-máquina. Na literatura há estudos que exploram o uso de associação entre as UA e as Unidades de Movimento (UM). Nesse caso, as UM podem ser empregadas para obter características de variações de volume extraídas para melhorar a interpretação e a localização das UA no reconhecimento de emoções (AZCARATE et al., 2005a; VALSTAR; PANTIC, 2012a; SAHA et al., 2015). Na Figura 6 é apresentado um exemplo dessa abordagem entre as UA e as UM, entre duas pessoas.

## 2.7 Análise de Imagens Digitais

Em processamento digital de imagens existem diversas técnicas que são empregadas na detecção e análise de informações. Essas técnicas são empregadas nas etapas de melhoramento, detecção, segmentação, extração de característica e classificação. Nessa seção, são apresentadas algumas abordagens computacionais para estudo e investigação de informações presentes na face dos indivíduos.



Figura 6 – Associação das UM com as UA.  
Adaptado de: (SAHYOUN, 2016).

### 2.7.1 Detecção de Regiões Candidatas

A face de um indivíduo é um objeto dinâmico e tem muita variabilidade em relação à sua aparência, sendo um desafio na construção de algoritmos capazes de fazer a detecção, baseados em métodos da área de visão computacional. Na literatura, técnicas são propostas baseadas na detecção de contorno, ou métodos mais complexos são apresentados combinando os algoritmos de reconhecimento de padrões e inteligência artificial. Essas soluções procuram explorar informações obtidas com a extração de características, uniformização de escala de cores, correção de iluminação e transformações de cores não lineares para localizar regiões que tratam a detecção da face. A maior parte desses algoritmos buscam, no primeiro momento, selecionar as regiões candidatas, os quais representem as áreas dos olhos, boca, nariz e contorno da face (HSU; ABDEL-MOTTALEB; JAIN, 2002), e construir mapas de informações para extração de características locais. Nessa etapa, a técnica denominada imagem integral (VIOLA; JONES, 2001; NGUYEN, 2007) é empregada para obtenção dessa informação. A imagem integral pode ser definida como tabelas de dados para pesquisas bidimensionais em que seu tamanho pode ser analisado até o tamanho da imagem. Cada elemento dessa imagem integral contém a soma de todos os *pixels* localizados entre a região superior esquerda da imagem original em relação ao objeto analisado. Essa técnica permite calcular a soma de áreas retangulares na imagem em qualquer posição ou escala da imagem original (ver Equação 1).

$$I(x, y) = i(x, y) + I(x - 1, y) + I(x, y - 1) - I(x - 1, y - 1), \quad (1)$$

sendo,  $I(x, y)$  o valor correspondente na imagem integral,  $i(x, y)$  representa a imagem original e  $(x, y)$  coordenadas de cada *pixel*.

Na Figura 7(a) tem-se a área retangular sombreada, e na Figura 7(b) se tem-se a representação de como a imagem integral pode ser calculada em uma única varredura da imagem.

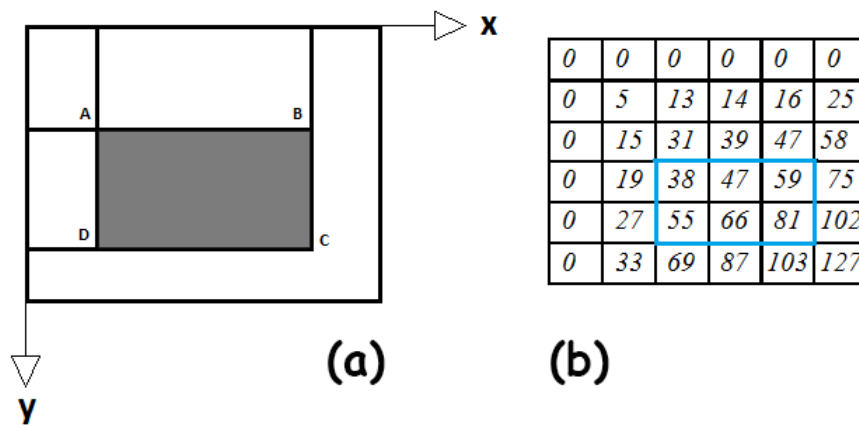


Figura 7 – Representação para calcular a imagem integral em (a) por meio da soma das áreas retangulares em (b).

### 2.7.2 Extração de Características da Face

A extração das características das expressões faciais é o primeiro passo para analisar e classificar as emoções presentes na face. Nesse processo, sempre ocorre o tratamento de ruídos presentes no processo de captura dos sinais da imagem. Então, são utilizadas técnicas como os filtros lineares e não lineares.

Tradicionalmente, existem duas classes de técnicas para a extração de informações das expressões faciais: a) a extração de características geométricas, que depende de parâmetros e informações faciais dos olhos, da boca e do nariz; e b) a análise da aparência, considerando toda a face como uma matriz de valores com a intensidade de cores, tons de cinza, etc. Esse processo busca comparar com outras matizes geradas a partir de uma imagem contendo a face do indivíduo (base).

Dessas soluções, um dos modelos mais utilizados nessa etapa é a abordagem geométrica, obtida por modelos deformáveis, os quais permitem segmentar e identificar os movimentos da região da boca, conforme demonstrado na Figura 8 (KASS; WITKIN; TERZOPOULOS, 1988).

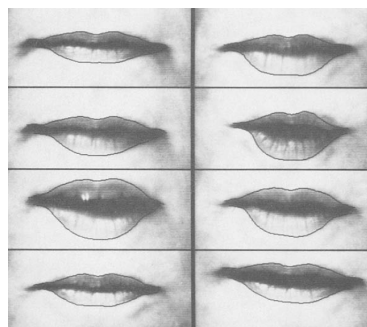


Figura 8 – Representação de movimento das regiões da boca por meio da técnica de deformação. **Fonte:** (KASS; WITKIN; TERZOPOULOS, 1988).

Apesar da técnica ter a capacidade de fazer a detecção de regiões, essa pode apresentar falhas no caso da presença de oclusões ou outros objetos na imagem. Por isso, os auto-



res em (BRUNELLI; POGGIO, 1993) propuseram uma solução para a localização de pontos na imagem, usando a relação com uma imagem-modelo com os pontos de interesse pré-selecionados, facilitando assim a localização e extração de características da face.

Também com o mesmo objetivo, os autores em (CHANG; HUANG, 2014) utilizaram a técnica *Principal Component Analysis* (PCA). O PCA é uma técnica que utiliza uma transformação ortogonal de vetores para converter um conjunto de observações de variáveis, possivelmente correlacionadas, num conjunto de valores de variáveis linearmente não correlacionadas. Esse conjunto é denominado componentes principais, onde o número de componentes principais é menor ou igual ao número de variáveis originais. O PCA também pode ser utilizado na redução de dimensionalidade em soluções de identificação da face, conforme proposta de (PADGETT; COTTRELL, 1997). Na Figura 9 é demonstrada a aplicação do PCA para análise das principais variações de cor, para cada indivíduo, no processo de extração de característica para identificação de uma nova imagem da face, com base em comparação no espaço denominado *eigenfaces* por meio da distância euclidiana (BRYANT; YARNOLD, 1995).



Figura 9 – Exemplo de PCA aplicado em imagem da face para a extração de características.

No processamento da informação, no PCA, uma imagem pode ser representada por um matriz  $Z$  de dimensão  $n \times m$ , em que  $n$  representa o número de linhas e  $m$  o número de colunas. Então, essa matriz pode ser representada por um vetor  $z^T$   $n$  dimensional em que  $n = mxn$  é representada na matriz  $Z$  concatenada, permitindo a extração de características na imagem.

Uma outra abordagem de extração de informações da face foi proposta por (COHN et al., 1998), em que uma função discriminante foi empregada para extração de informações nas regiões individuais da face: sobrancelha, olhos e boca. Nessa proposta, a função discriminante lambda de Wilk, um método estatístico usado para verificar se há diferenças entre as médias, em uma combinação de variáveis dependentes foi aplicada para três ações faciais da região dos olhos, cinco funções para a região das sobrancelhas e nove ações faciais da região do nariz

e boca. Essa função foi empregada para possibilitar o mapeamento das informações da face. Na Figura 10 é apresentado o mapeamento dessas regiões no reconhecimento das expressões faciais.

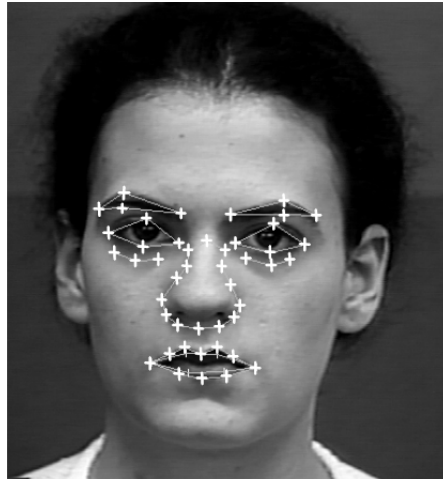


Figura 10 – Detecção dos pontos discriminantes na face para a detecção das emoções. **Fonte:** (COHN et al., 1998).

Dentre as diversas opções para a extração de características na literatura, nota-se que as abordagens geométricas possuem restrições quando aplicadas sobre um domínio com muita presença de ruído. Porém, essa abordagem tem resultados relevantes em relação ao desempenho no processamento, quando as plataformas adotadas apresentam limitações de recursos computacionais comparadas com as demais abordagens (ACEVEDO et al., 2017). Por isso, esse tipo de técnica torna-se relevante para aplicação a um contexto maior de equipamentos disponíveis para execução.

## 2.8 Técnicas de Inteligência Artificial

A Inteligência Artificial (IA) é uma área que surgiu nos anos 50, quando os pesquisadores buscavam uma maneira de fazer os computadores pensarem como os seres humanos (KETKAR et al., 2017). Essa área pode ser compreendida como um campo global que envolve o aprendizado de máquina e a aprendizagem profunda. Na Figura 11 é exemplificado essas áreas e as relações entre elas.

Uma das subáreas de IA é o Aprendizado de Máquina (AM) que começou a ser desenvolvida na década de 90 e rapidamente se tornou o subcampo mais popular, avançando em seu uso devido à maior disponibilidade de *hardware*, aumento do volume de dados e informações para serem processadas (KETKAR et al., 2017). Esse subcampo está diretamente relacionado com abordagens estatísticas e matemáticas, mas existem alguns pontos diferentes que devem ser levados em consideração. Um dos pontos mais importantes é que, ao contrário da área de estatística, a aprendizagem de máquinas tende a lidar com grandes e complexos conjuntos de

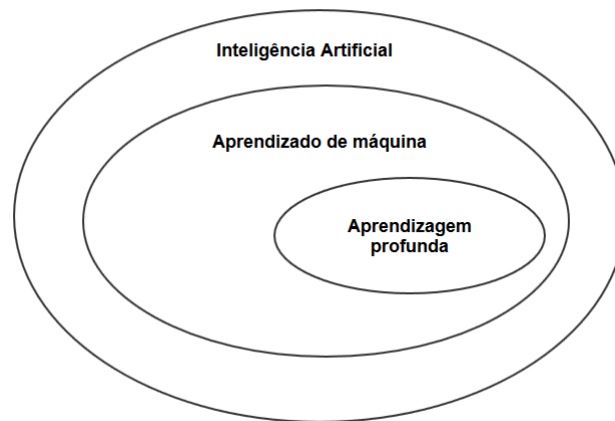


Figura 11 – Representação das relações entre inteligência artificial, aprendizado de máquina e aprendizado profundo. Adaptado de: (KETKAR et al., 2017).

dados, como por exemplo, grandes volumes de imagens com milhões de *pixels* em que na área da estatística clássica é impraticável.

Entre as técnicas de AM utilizadas para reconhecimento das expressões faciais, a Indução em Árvores de Decisão vem sendo uma das abordagens empregadas devido à capacidade de classificação se aproximar à tomada de decisão do ser humano, conforme descrito nos trabalhos (THOMSON et al., 2006; KOSUDA et al., 1998; OLANOW; WATTS; KOLLER, 2001). Recentemente, tem surgido também o uso de redes neurais convolucionais profundas, uma subárea de AM, na análise de face, conforme os trabalhos (DACHAPALLY, 2017; NG et al., 2015; KIM; LEE; PROVOST, 2013; SALMAM; MADANI; KISSI, 2016; ZAVAREZ; BERRIEL; OLIVEIRA-SANTOS, 2017).

Com base na Figura 11, outro subcampo de IA é a Aprendizagem Profunda (AP), a qual vem sendo estudada desde 2006, conforme descrito nos trabalhos (BENGIO et al., 2009; HINTON; OSINDERO; TEH, 2006). As abordagens e técnicas desenvolvidas nos últimos anos proporcionaram uma série de impactos positivos em diversas áreas do conhecimento, com uso dessas técnicas. Exemplos são as áreas de processamento de sinais, análise de imagens digitais, visão computacional, etc. (BENGIO et al., 2009). A AP pode ser descrita como a capacidade de aprender vários níveis de representação e abstração e permite melhorar a compreensão dos dados, como por exemplo, imagens, sons e texto (DENG; YU et al., 2014).

### 2.8.1 Aprendizagem Profunda

O conceito de AP teve origem a partir de estudos de Redes Neurais Artificiais (RNA), as quais podem ser definidas como uma estrutura complexa interligada por elementos de processamento ou neurônios. Esses neurônios possuem a capacidade de processar dados e informações para a representação de conhecimento. Os primeiros conceitos de RNA foram introduzidos em 1943, mas ganhou popularidade algumas décadas depois, com a introdução de algoritmos

de treinamento, como por exemplo, o *backpropagation*. Esse conceito permite a realização de um treinamento posterior para aperfeiçoar os resultados do modelo-base usado no treinamento inicial do modelo construído (RUSSELL; NORVIG, 2004).

No ano de 1958, a RNA *perceptron* foi então proposta e desenvolvida por Frank Rosenblatt com base nos trabalhos de Walter Pitts e Warren Sturgis McCulloch. Esse modelo possui um único neurônio que realiza as classificações de maneira linear. Na Figura 12 pode ser observado um exemplo do RNA *perceptron*. Nessa figura pode ser observado o termo  $y(n)$  que representa os parâmetros de entrada, o  $w(n)$  que define os respectivos pesos.

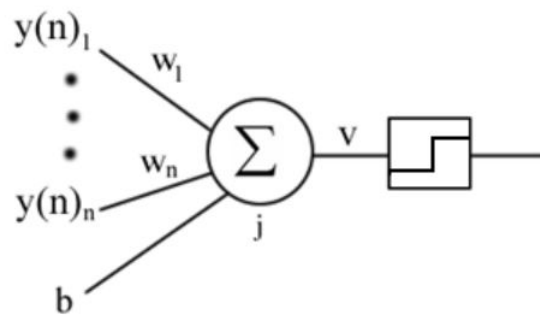


Figura 12 – Rede Neural Artificial *Perceptron* com um único neurônio.

Com o passar dos anos e a necessidade de lidar com dados não linearmente separáveis, os pesquisadores adicionaram novas camadas de neurônios, às quais denominou-se "camadas ocultas". As camadas ocultas são empregadas numa etapa para processamento da informação onde não é possível prever a saída desejada nas camadas intermediárias. Essa abordagem deu origem às redes neurais *Multi Layer Perceptron* (MLP) com várias camadas ocultas. Essas redes podem ser consideradas um exemplo clássico de aprendizagem profunda, conforme mostrado na Figura 13.

Segundo (RUSSELL; NORVIG, 2004), em geral, as redes MLP (ver Figura 13) possuem uma arquitetura composta pelos elementos fundamentais para seu funcionamento, dados por:

- ❑ **Camada de entrada:** Contém as informações que são fornecidas à rede no momento da sua construção, e esses dados são utilizados para treinar o modelo;
- ❑ **Camadas ocultas:** São responsáveis por realizar cálculos intermediários que auxiliam a rede a encontrar os valores finais para a camada de saída;
- ❑ **Peso:** Contém um valor de peso para cada neurônio, no valor do neurônio da próxima camada;
- ❑ **Função de ativação:** É um componente com a finalidade de permitir a solução de problemas complexos, por meio da transformação não linear dos dados, a partir da camada de entrada;
- ❑ **Camada de saída:** Contém o resultado da classificação final do processamento da rede.

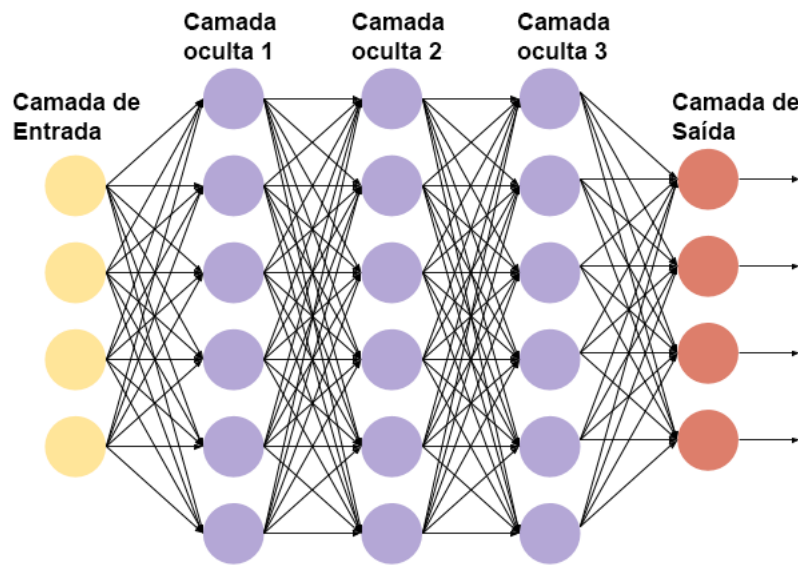


Figura 13 – Exemplo de uma RNA MLP com características de aprendizagem.

A AP tem em seu conjunto um aglomerado de técnicas e arquiteturas de algoritmos de inteligência artificial. O grande diferencial é que essas soluções têm como característica principal muitas camadas de processamento de informações não-lineares que são de natureza hierárquica. Em um modelo AP tem-se três tipos de estruturas:

- ❑ redes de aprendizagem não supervisionada: Nesse tipo de aprendizagem não se conhece a classe dos dados utilizados. Esse modelo identifica a correlação existente entre os dados observados e podem ser chamadas de redes generativas;
- ❑ redes de aprendizagem supervisionada: Esse tipo de rede recebe como parâmetro as classes dos dados para a realização da caracterização das distribuições posteriores por meio das classes dos dados visíveis. As classes dos dados estão sempre disponíveis de forma direta ou indireta. Essas redes são conhecidas como redes profundas discriminatórias;
- ❑ redes profundas híbridas: Nesse modelo ocorre a discriminação assistida dos dados, combinando os bons resultados das redes por meio da otimização e regularização das redes profundas supervisionadas.

A infraestrutura de um modelo AP tem alguns elementos envolvidos na construção. Primeiramente, são empregadas as camadas de convolução com a funcionalidade de filtrar as imagens com uma máscara (*kernel*) para realizar operações de convolução. Dessa maneira, a imagem é percorrida e os traços mais relevantes são representados em mapas de características (*features*). A convolução tem como funcionalidade melhorar as informações por meio de realce de detalhes sem perder recursos importantes. O uso desses filtros é relevante quando pretende-se projetar uma arquitetura que não seja apenas eficiente na obtenção de recursos, mas escalável para grandes volumes de dados (ROSEBROCK, 2017). Para exemplificar esse processo, a imagem RGB representada na Figura 14 (A) mostra parte desse processo para etapa de separar os três planos

de cores: vermelho, verde e azul (dos canais de cores RGB). Na Figura 14 (B), a região em verde se assemelha à imagem de entrada  $5 \times 5 \times 1$ . O elemento envolvido na execução da operação de convolução na primeira parte de uma camada convolucional *kernel* (K), representado na cor amarela, mostra K como uma matriz de tamanho  $3 \times 3 \times 1$ .

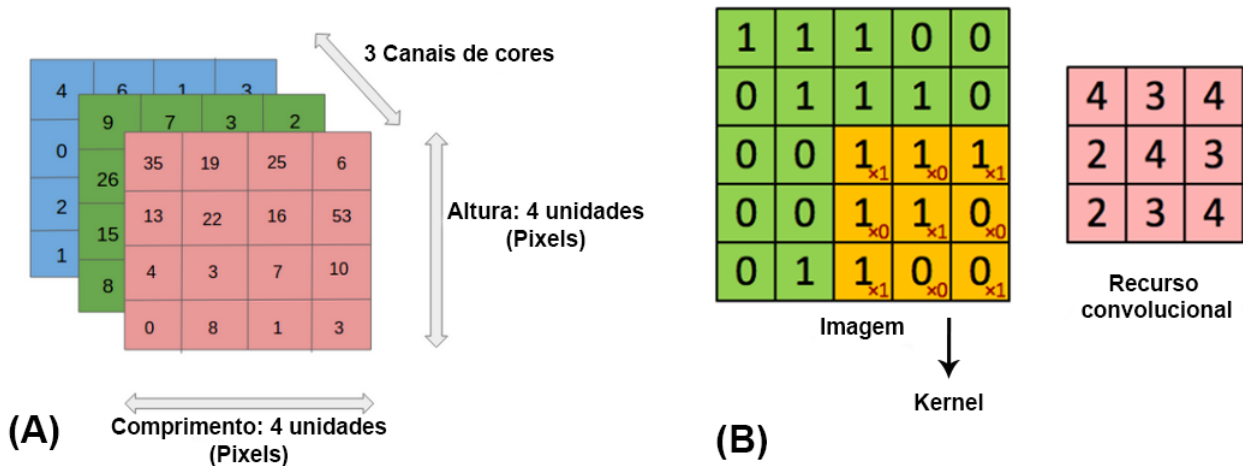


Figura 14 – Exemplo de operações realizadas durante o processo de convolução: (A) Representação de uma imagem  $4 \times 4 \times 3$  em RGB, (B) demonstração de uma convolução em uma imagem  $5 \times 5 \times 1$  com um *kernel*  $3 \times 3 \times 1$  para obter um recurso  $3 \times 3 \times 1$ .

Nesse processo de convolução, um *kernel* poderá se deslocar pela imagem, sempre executando uma operação de multiplicação de matrizes entre K e a parte P da imagem sobre a qual o *kernel* está percorrendo (ver Figura 15).

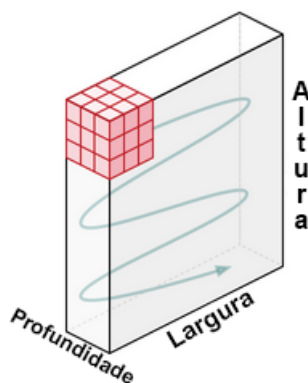


Figura 15 – Representação do deslocamento do *kernel* em uma imagem.

Dessa maneira, o filtro irá se deslocar no sentido da direção direita até percorrer a imagem conforme demonstrado na Figura 16. No caso de imagens com vários canais, como uma imagem RGB, o *kernel* tem uma representação para cada canal com as mesmas características de profundidade. A multiplicação das matrizes é então realizada e os resultados podem ser somados para fornecer uma saída com a devida profundidade (ROSEBROCK, 2017).

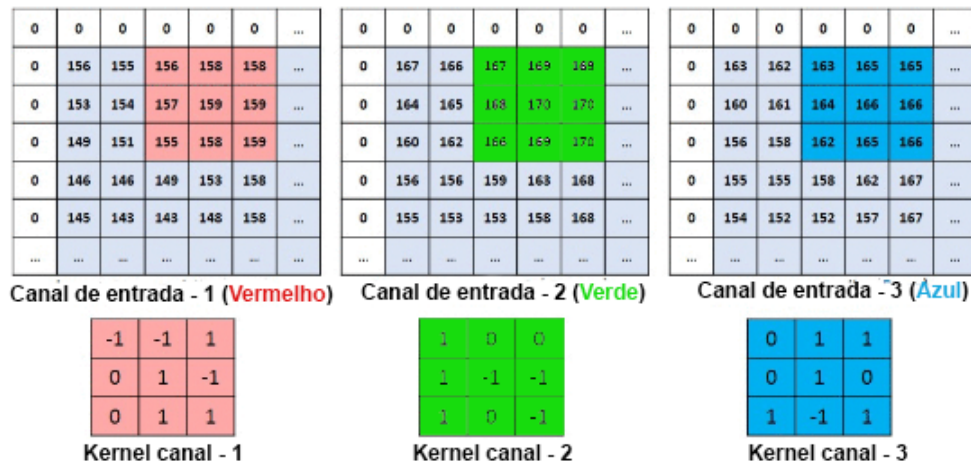


Figura 16 – Representação de convolução em uma matriz de imagem  $M \times N \times 3$  com um *kernel*  $3 \times 3 \times 3$ .

Semelhante à camada convolucional, uma camada denominada *Pooling* é empregada na sequência e tem como responsabilidade reduzir o tamanho espacial das informações. Isso diminuirá a necessidade de recursos computacionais para processar os dados. Essa camada é útil para a extração de características importantes, mantendo assim o processo de treinamento eficaz do modelo (ROSEBROCK, 2017; NIELSEN, 2015). Dependendo da complexidade da imagem, o número de camadas pode ser inseridos na sequência, para capturar detalhes em níveis mais baixos. As duas abordagens mais empregadas na etapa é o *pool* máximo e *pool* médio. O *pool* máximo retorna o valor máximo de intensidade de brilho da parte da imagem. Por outro lado, o *pool* médio retorna a média de todos os valores da parte da imagem. Além disso, essa etapa funciona como um supressor de ruído, que descarta as ativações ruidosas e executa o cancelamento de ruído em conjunto com a redução de dimensionalidade. O *pool* médio também realiza a redução da dimensionalidade como um mecanismo de supressão de ruído (NIELSEN, 2015).

Outros elementos importantes que compõem um modelo AP são as funções de ativação. Essa etapa tem a funcionalidade de tratar a não-linearidades do sistema para que consiga aprender qualquer tipo de característica. Atualmente, diversas funções de ativação são empregadas, mas as principais são: *sigmóide*, *ReLU*, *tanh* e a *Softmax*. Por fim, a rede tem geralmente uma camada denominada *Fully connected*, onde os dados são vetorizados e a saída são  $N$  neurônios. Esse  $N$  pode ser a quantidade de classes do seu modelo utilizado para a classificação com uma rede neural artificial. Essa saída é alimentada a uma rede neural *feed-forward* e uma retropropagação é aplicada para as iterações no treinamento. Ao longo de uma série de épocas, o modelo é capaz de aprender a distinguir as características dominantes para classificá-las (ROSEBROCK, 2017).

Existem várias arquiteturas de modelos CNNs disponíveis, fundamentais no desenvolvimento de algoritmos. As principais arquiteturas são:

#### □ LeNet

- ❑ AlexNet
- ❑ VGGNet
- ❑ MobileNet
- ❑ GoogLeNet
- ❑ ResNet
- ❑ ZFNet

Dentre esses modelos, devido à característica da plataforma que será construída, a *MobileNet* será investigada para o modelo proposto. Essa rede é uma simplificação de redes neurais, que possibilita seu uso em aplicações *web* e em dispositivos móveis. A *MobileNet* é uma classe de convolução de redes neurais projetadas por pesquisadores da empresa Google. A intenção desse modelo é facilitar o uso de recursos de dispositivos móveis, onde parte da informação é processada na nuvem, de forma a economizar recursos do aparelho (NIELSEN, 2015; ROSEBROCK, 2017).

## 2.9 Considerações Finais

Segundo dados estatísticos já apresentados anteriormente por órgãos como o *Centers for Disease Control and Prevention* e a Organização Mundial da Saúde mostram que existe uma grande parcela da população que apresenta sintomas ligados ao TEA. Também há um número cada vez menor de políticas públicas voltadas para o desenvolvimento de ferramentas e tratamentos para esse público. Diversas pesquisas já realizadas demonstraram que os indivíduos com TEA possuem limitações nas habilidades para reconhecimento e expressão das emoções, o que pode comprometer as interações sociais.

O estudo das emoções apresentadas por Darwin, Ekman e Friesen demonstram que as emoções básicas são universais e podem ser reconhecidas e classificadas de acordo com as expressões faciais. Essas pesquisas são fundamentais para a compreensão dos gestos empregados para detectar as emoções. Como o comportamento das emoções em indivíduos com TEA é atípico, o estudo da teoria da mente auxilia na compreensão dos estados emocionais dessas pessoas para que permita investigar soluções que auxiliem no aprimoramento dessas habilidades.

A teoria da aprendizagem, baseada em jogos sérios, foi utilizada para fornecer uma experiência diferenciada para os indivíduos, aumentando o interesse e motivação. Isso possibilita o desenvolvimento de ferramenta de entretenimento e, ao mesmo tempo, contribui com avaliações clínicas voltadas para a saúde. A modelagem computacional e o uso de técnicas de processamento de imagem e redes neurais convolucionais permite o desenvolvimento de modelos para jogos com personagens que explorem as habilidades de detecção e reconhecimento das emoções.



---

## Estado da Arte

Neste capítulo são apresentados os principais estudos nos últimos anos que exploraram os métodos na aprendizagem das emoções básicas em pessoas com TEA.

### 3.1 Ensino de Competências Emocionais em Indivíduos com TEA

Esta seção apresenta trabalhos que exploram tecnologias de jogos voltados para auxiliar pessoas com TEA no aprendizado das competências emocionais. Na Tabela 2 são apresentadas, de forma sumarizada, as principais contribuições presentes na literatura.

Tabela 2 – Trabalhos realizados utilizando realidade aumentada, realidade virtual e jogos para auxiliar no tratamento do TEA ou deficiência intelectual

Referência	Tecnologia Utilizada	Método de Avaliação	Resultados
(CHEN; LEE; LIN, 2015)	Realidade Aumentada	Taxa de avaliação dos participantes durante três fases de testes.	A avaliação foi realizada em três etapas, com uma evolução média de 89,94% dos pacientes.
(MAZUREK; ENGELHARDT; CLARK, 2015)	Jogos Digitais	Um estudo qualitativo para examinar as preferências e motivações para jogos de videogame em adultos com TEA.	As análises permitiram identificar os benefícios e os efeitos negativos do uso de videogame.
(CHEN; LEE; LIN, 2016)	Realidade Aumentada	Taxa de avaliação dos participantes durante três fases de testes. Utilização de um livro com histórias e vídeos em RA.	A avaliação foi realizada em três etapas para avaliar a capacidade de identificar as emoções básicas.
(DIDEHBANI et al., 2016)	Realidade Virtual	Os participantes foram submetidos a três fases de avaliação no ambiente virtual.	Para validar a proposta foi avaliada a capacidade de identificar as emoções com seis participantes. Em todas as fases ocorreram melhoras significativas dos participantes.
(TSANGOURI et al., 2016)	Jogo Digital Móvel	Através da avaliação realizada antes e após as intervenções com nove participantes com TEA.	Os resultados produzidos pela utilização do jogo demonstraram um aumento significativo nas capacidades de reconhecer e expressar as emoções com poucas ações em relação as intervenções.
(STURM; PEPPE; PLOOG, 2016)	Jogo Digital Móvel	Um estudo qualitativo, levando em consideração os aspectos relacionados aos padrões atípicos da atenção em portadores do TEA.	Os resultados obtidos foram satisfatórios.
(MALINVERNI et al., 2017)	Jogo Digital	Foi realizado um estudo observacional com o acompanhamento de profissionais da área da saúde e designers de jogos, com 10 indivíduos com TEA com o objetivo de avaliar a aceitação e o prazer do jogo.	As análises em vídeo e comportamentais revelaram uma aceitação positiva de jogos pelos participantes.

Referencia	Tecnologia	Método de Avaliação	Resultados
(FRIDENSON-HAYO et al., 2017)	Jogo Digital	Uma avaliação transcultural, com crianças de 6 a 9 anos, e com diferentes níveis de TEA. Para a avaliação foram utilizados dados como o conhecimento vocabulário dos participantes.	Os resultados foram positivos para os aspectos qualitativos através dos relatos informados pelos pais dos indivíduos e também para os dados quantitativos.
(DAPOGNY et al., 2018)	Jogo Digital	Uma avaliação qualitativa para avaliar a capacidade de expressão de quatro emoções básicas: neutro, alegria, tristeza e raiva em situações sociais. Essa avaliação foi realizada por meio da participação de usuários com TEA.	Os resultados para a classificação das quatro emoções básicas se apresentaram satisfatórios.
(ALMEIDA et al., 2019)	Jogo Digital	Uma avaliação qualitativa realizada por meio de experimentos com uma equipe multidisciplinar com 20 pessoas, 10 crianças com TEA E 28 crianças com desenvolvimento neuro-típico.	Os resultados obtidos foram satisfatórios para a aceitação da aplicação desenvolvida.
(PAVEZ et al., 2021)	Jogo Digital	Uma avaliação qualitativa para avaliar a capacidade de expressão das cinco emoções básicas: raiva, medo, tristeza, alegria e neutra.	Os resultados obtidos foram satisfatórios, por meio de uma análise qualitativa, para o reconhecimento e o auxílio no desenvolvimento dos participantes.

No trabalho de (CHEN; LEE; LIN, 2015) foi desenvolvido um sistema para possibilitar que três adolescentes com TEA pudessem aperfeiçoar e praticar os conhecimentos a respeito das expressões faciais. No sistema proposto, foram disponibilizadas para os usuários algumas animações em formato 3D das expressões faciais básicas. Essas animações foram sobrepostas à face dos participantes. Para a criação desses modelos 3D, os autores utilizaram imagens frontais e laterais das faces de cada participante para gerar máscaras faciais nos modelos. Ainda na construção do modelo foi utilizada a teoria FACS para garantir a representação dos micros movimentos faciais. Essa sobreposição tinha como objetivo facilitar a definição das emoções e, consequentemente, o aperfeiçoamento das habilidades sociais desses indivíduos. O método proposto foi aplicado no ambiente escolar por meio do uso da tecnologia de Realidade Aumentada (RA). Os três participantes foram submetidos a sete etapas de intervenção, utilizando o método proposto. Os resultados obtidos foram analisados com a capacidade inicial de reconhecer essas emoções. O método obteve uma acurácia de 89,94% para as seis emoções básicas. Os autores destacaram que, apesar da dificuldade em manter a atenção e foco durante as intervenções, o uso de RA contribuiu no interesse em usar essa ferramenta. Em trabalhos futuros, os autores recomendam um número maior de indivíduos e o uso de métricas para quantificar as contribuições da utilização da tecnologia RA.

No trabalho apresentado em (MAZUREK; ENGELHARDT; CLARK, 2015), os autores realizaram uma análise qualitativa para examinar as preferências e motivações da utilização de jogos de videogame com adultos portadores de TEA. Nesse trabalho, foi realizada uma entrevista em que o questionário era composto por quatro perguntas sobre uso de videogame. Essas questões tinham como objetivo conhecer as habilidades dos adultos com jogos. Para esse estudo, um grupo com 58 adultos com TEA foram selecionados. O uso do videogame tinha como função aumentar as interações sociais entre as pessoas com diferentes tipos de jogos que pudessem explorar as habilidades emocionais. Essa avaliação permitiu identificar os benefícios do uso de videogame no contexto de interações sociais e a redução do estresse num processo avaliativo. Os autores mostraram que os gêneros de jogos mais frequentes foram o *Role-Playing* (31%) e o *Action-Adventure* (19%). Apesar dessas informações, o estudo ficou restrito a não quantificar medidas com o grupo. Além disso, os participantes já tinham habilidades cognitivas relativamente bem desenvolvidas e habilidades verbais.

Em (CHEN; LEE; LIN, 2016), os autores desenvolveram um livro de histórias com elementos de vídeo modelagem combinados com realidade aumentada. Essas tecnologias foram inseridas com o objetivo de fortalecer e atrair a atenção de crianças com sinais de TEA. Foi utilizado rastreamento natural sem marcadores para ensinar as crianças a reconhecer os padrões para expressar as emoções. A proposta tinha como característica a concentração em imagens visuais estáticas impressa em um livro de histórias. Posteriormente, ocorria uma fase de extensão para observação de animação de atores presentes na história. Uma investigação foi realizada com as fases de avaliação inicial, intervenção e acompanhamento. Durante as avaliações, os autores observaram que o sistema proposto se mostrou útil para ajudar os participantes com TEA

a reconhecer e compreender as emoções. Isso permitiu que esses participantes pudessem reconhecer as emoções de pessoas com diferentes idades. No entanto, apenas medidas qualitativas foram usadas e as medidas quantitativas não foram consideradas pelos autores.

O trabalho de (DIDEHBANI et al., 2016) tinha como objetivo desenvolver e avaliar a utilização de ambientes com realidade virtual no ensino de habilidades sociais e emocionais em crianças com TEA. Os autores avaliaram habilidades como o reconhecimento das emoções básicas, atribuição social, atenção e função executiva responsável por controlar as reações cognitivas nos seres humanos. Foi utilizado um ambiente virtual, desenvolvido especialmente para trabalhar as habilidades emocionais e sociais com RV e objetos 3D. Nessa aplicação, os indivíduos deveriam executar atividades em situações diversas, tais como: diálogos, interações sociais, atividades em grupo intermediadas pela interface da aplicação. Para avaliar o desempenho, os autores selecionaram um grupo com 30 crianças, com idades entre 7 e 16 anos, diagnosticadas com TEA. Esse grupo foi submetido a 10 sessões, com duração de 1 hora, em 5 semanas. Os resultados obtidos mostraram melhorias na capacidade de reconhecimento das emoções, na atribuição social e na função executiva do raciocínio. Segundo os autores, essas descobertas mostraram que o uso de plataforma com RV pode melhorar as deficiências sociais em indivíduos com TEA. Os autores destacaram que os ambientes virtuais podem contribuir com pessoas que não possuem acesso aos tratamentos tradicionais.

No trabalho proposto pelos autores em (TSANGOURI et al., 2016) um jogo para dispositivos móveis foi desenvolvido para auxiliar no desenvolvimento das capacidades emocionais de indivíduos com TEA. O jogo foi estruturado para proporcionar uma aprendizagem das emoções intermediada por técnicas computacionais. A estrutura da aplicação foi dividida em quatro partes: aplicação do dispositivo móvel, servidor de requisições, servidor para processamento das emoções e servidor *web* responsável por disponibilizar a classificação das emoções. A aplicação foi testada com nove indivíduos com TEA entre as idades de 18 e 25 anos. Essa avaliação ocorreu com as intervenções durante duas semanas. Os resultados demonstraram que os participantes tiveram as habilidades de reconhecer e expressar as emoções, aperfeiçoadas de forma significativa segundo os especialistas envolvidos no projeto. Os autores não relatam informações sobre as avaliações qualitativa e quantitativa do sistema.

Em (STURM; PEPPE; PLOOG, 2016), os autores propuseram o desenvolvimento de uma aplicação para auxiliar no aprendizado da identificação das emoções básicas. O jogo proposto explorou situações cotidianas por meio da utilização de *emoticon*. Para avaliar a ferramenta, três profissionais utilizaram o jogo com os indivíduos com TEA para explorar situações que proporcionasse o aprendizado das emoções básicas. Durante as avaliações, correções e ajustes foram levantados para que o jogo adaptasse as características sócio comportamentais dos indivíduos com TEA. Os autores mostraram que aplicações com essas funcionalidades podem contribuir com o tratamento de padrões anormais de atenção, incluindo o reconhecimento e expressão das emoções. De forma semelhante ao trabalho de (TSANGOURI et al., 2016) também não ocorreu uma análise quantitativa do sistema.

Em (MALINVERNI et al., 2017), os autores propuseram uma abordagem para desenvolver jogos com fins terapêuticos. Os jogos criados exploravam as experiências de especialistas, contribuições de crianças e *designers*. Nesse trabalho, um jogo baseado no dispositivo *Kinect* foi construído para crianças com nível severo do TEA. Os experimentos foram realizados em um grupo de 10 crianças com TEA. Os bons resultados ocorreram com o uso de elementos de mecânica inseridos nos jogos, os quais se demonstraram atraentes para as crianças. Os autores identificaram conceitos de *design* eficazes e importantes para futuras pesquisas com jogos para crianças com TEA. No entanto, destacaram que o estudo envolveu apenas crianças com TEA com idade de 10 a 11 anos, e não ocorreu análise do comportamento das crianças durante a aplicação do jogo. Além disso, não foi avaliada a capacidade de transferência das novas habilidades adquiridas para o comportamento cotidiano, em ambientes domésticos.

Em (FRIDENSON-HAYO et al., 2017), os autores desenvolveram um jogo denominado *Emotiplay* para auxiliar no aperfeiçoamento das habilidades emocionais em crianças com TEA. O jogo foi projetado para ser executado em ambientes *desktops* permitindo assim o fácil acesso por meio de um navegador com tecnologias HTML5, CSS e JavaScript. Esse jogo tem animações que retratam cenas cotidianas do convívio social e questionários para serem respondidos a respeito das emoções. Por meio dessa estratégia, os participantes podem expressar as capacidades de reconhecer as emoções. Dentre essas atividades propostas, o usuário deve efetuar o reconhecimento das emoções por meio das expressões faciais, dos gestos, da fala, e outras características socio-comportamentais. Para avaliar a aplicação, indivíduos com TEA em diferentes países foram utilizados. A avaliação em diferentes países permitiu avaliar o desempenho da aplicação do jogo em diferentes culturas e contextos sócio comportamentais. Foram obtidos resultados quantitativos como a taxa de acerto dos participantes por meio da análise estatística dos resultados durante a execução dos experimentos. Esses resultados demonstraram a evolução do vocabulário emocional dos pacientes e a redução dos sinais típicos do TEA.

No trabalho de (DAPOGNY et al., 2018), os autores desenvolveram uma aplicação para auxiliar os indivíduos com TEA a desenvolver quatro emoções básicas: neutro, alegria, tristeza e raiva, por meio de técnicas de aprendizagem de máquina. As imagens em vídeos foram convertidas para níveis de cinza e aplicado o detector de face Viola e Jones da biblioteca OpenCV. Então, pontos foram traçados nessa região e 49 características foram extraídas para classificação com algoritmo florestas randômicas. A classificação dessas emoções foi avaliada em uma base de imagens públicas, bem como em ambiente real em pessoas de duas cidades distintas. Os resultados para a classificação dessas emoções foram satisfatórios no processo de reconhecimento de emoções. Para tornar o jogo atrativo, os autores desenvolveram cenários envolvendo situações sociais em que era necessário que os usuários expressassem as emoções solicitadas. Uma avaliação qualitativa não foi realizada e autores recomendam que essa avaliação seja realizada no futuro.

No trabalho proposto por (ALMEIDA et al., 2019), os autores desenvolveram um jogo para auxiliar crianças com TEA a reconhecer expressões faciais associadas apenas a quatro emoções

básicas: alegria, tristeza, raiva e surpresa. O jogo foi desenvolvido usando a estratégia *Role-Playing Game* (RPG) com a intenção de tornar o jogo com requisitos de habilidades sociais e não tão competitivo. O jogo foi construído com interfaces 2D desenvolvidas para manter a atenção, controle de acesso e registro para permitir o monitoramento da progressão. Para a avaliação da aplicação, uma equipe multidisciplinar composta por cinco especialistas - pedagogia, psicologia, psicopedagogia e desenvolvedores de jogos - acompanharam os testes. Os experimentos foram realizados com 10 crianças com TEA e 28 crianças com desenvolvimento neuro-típico, separadas em grupos controle e experimental. Com base nas respostas de questionários de usabilidade, os autores destacam que houve relevante aceitação da aplicação, mas o tempo de exposição das crianças com TEA deve ser aprimorado para que ocorra um reconhecimento das expressões faciais mais rápido.

Em (PAVEZ et al., 2021) os autores desenvolveram um aplicativo que simula um espelho com uma *webcam*, em que redes neurais convolucionais são empregadas para analisar as imagens que são capturadas por uma câmera e compará-las com a que o indivíduo com TEA deve realizar para detectar cinco emoções básicas. Para o processo de treinamento do espelho inteligente, os autores avaliaram a solução em dois modelos de redes neurais convolucionais a VGG 16 e ResNet50. A estrutura da aplicação foi desenvolvida no formato de uma API responsável por receber as imagens dos indivíduos e classificá-las em tempo real. Os resultados mais expressivos foi para o modelo VGG 16 na base de imagens CK+. A ferramenta foi avaliada com pessoas com autismo e monitorada por especialistas. Nos experimentos, os profissionais puderam avaliar o índice de aceitação e diversas informações de usabilidade da ferramenta. A ferramenta possibilitou a realização das sessões de tratamento durante o período de isolamento social. Como restrições, o método proposto apresenta a necessidade de um hardware específico e configurações adequadas, o que para certas pessoas pode ser um obstáculo para uso no dia a dia. Assim, é relatado pelos autores que a abordagem deve ser aprimorada em relação à qualidade das interfaces.

## 3.2 Considerações Finais

A partir dos estudos apresentados nessa seção é possível observar a relevância do tema de pesquisas. Essas pesquisas se empenharam em melhorar o diagnóstico, tratamento e a qualidade de vida dos indivíduos portadores de TEA com técnicas computacionais. Esses trabalhos exploram desde a detecção das emoções através das expressões faciais, como em (DAPOGNY et al., 2018), até a criação de jogos e aplicações dinâmicas, que utilizam recursos tecnológicos como: realidade virtual, realidade aumentada e jogos (CHEN; LEE; LIN, 2016; CHEN et al., 2019; MAZUREK; ENGELHARDT; CLARK, 2015).

A maioria desses trabalhos investiga estratégias de forma individual ao paciente. Não há aspectos e características relevantes para os profissionais da área da saúde em relação ao aprimoramento do paciente ou evolução em sessões de terapia. Grande parte desses trabalhos foram

voltados para a detecção e classificação das emoções através das expressões faciais. Em (DAPOGNY et al., 2018), algumas características foram exploradas para contribuir com o especialista. No entanto, esse estudo não possui uma avaliação quantitativa do impacto e contribuições dessa proposta para o aprimoramento do portador com TEA.

Nos trabalhos de (CHEN; LEE; LIN, 2016; CHEN; LEE; LIN, 2015) são citadas a ausência de resultados quantitativos, que demonstrem a evolução dos indivíduos com TEA no decorrer das intervenções. Já em (MALINVERNI et al., 2017) foi demonstrada a relevância em adaptar, testar e avaliar essas soluções em ambientes e, especialmente, ambientes domésticos, sem a presença de um especialista. Ainda nos trabalhos de (MAZUREK; ENGELHARDT; CLARK, 2015; ALMEIDA et al., 2019), os autores destacam a necessidade de adaptar o contexto ao tempo necessário para cada indivíduo desenvolver as habilidades e tratar as questões da heterogeneidade dos participantes.

Com base nesses pontos destacados na literatura, essa pesquisa tem como proposta o desenvolvimento de uma plataforma dinâmica, que possa abranger diversas metodologias (aprendizado baseado em jogos, jogos sérios, etc) e recursos computacionais para contribuir na evolução do quadro clínico de indivíduos com TEA. Essa plataforma deverá considerar as características de cada usuário, permitindo personalizar de acordo com as necessidades do usuário. Pretende-se ainda ter uma plataforma que atenda a heterogeneidade do indivíduo e auxilie os profissionais da saúde com dados da evolução do quadro clínico de cada indivíduo. Essa ferramenta ainda contribuirá com o aperfeiçoamento de técnicas computacionais para a detecção e classificação das emoções.



---

## Metodologia

Neste capítulo é apresentada uma plataforma computacional para auxiliar o aprimoramento de habilidades relacionadas as emoções para as pessoas com TEA por meio de um jogo com conceitos investigados de jogos sérios, redes neurais convolucionais e métodos de processamento de imagens digitais.

### 4.1 Projeto do Jogo

Para o desenvolvimento do jogo e elementos básicos da representação das emoções foram utilizadas as recomendações propostas por (NOVAK, 2011). De acordo com os autores desse trabalho, um jogo deve ter etapas de desenvolvimento como: conceito, pré-produção, protótipo, produção e pós-produção. Cada fase é responsável por explorar habilidades a serem desenvolvidas com um jogo. Nas subseções são apresentadas as etapas no desenvolvimento desse jogo.

#### 4.1.1 Conceito

A etapa conceito começa com a decisão de planejamento do projeto (também conhecida como fase de pré-produção). A etapa conceito do jogo foi explorada com o objetivo de propor uma ferramenta voltada para os indivíduos com TEA e foram considerados:

- ❑ **Público alvo:** Pessoas com TEA, profissionais da saúde e pais ou responsáveis;
- ❑ **Recursos necessários:** Computadores, placas gráficas e aparelhos *smartphones*;
- ❑ **Plataforma:** *Web* e dispositivos móveis com sistema operacional *Android*.

O objetivo do desenvolvimento do conceito foi decidir sobre o conteúdo do jogo e transmitir a ideia para outras pessoas, por escrito, de forma simples e clara.

### 4.1.2 Pré-produção

Nessa etapa foram definidos os personagens, cenários e enredo do jogo. Os personagens foram nomeados como Ana, Juninho e Michelzinho. Além disso, foram definidos os enredos para que em cada cenário fosse possível trabalhar as emoções básicas como:

- ❑ **Cenário da emoção alegria:** O personagem Ana ou Juninho ganha uma bicicleta de sua mãe;
- ❑ **Cenário da emoção tristeza:** O personagem Ana ou Juninho na praia, compra um sorvete, mas deixa cair antes de comer;
- ❑ **Cenário da emoção raiva:** Os personagens Ana ou Juninho estão com raiva pois não consegue terminar de montar um quebra-cabeça;
- ❑ **Cenário da emoção surpresa:** O personagem Ana ou Juninho está caminhando pelo parque quando começa a chover;
- ❑ **Cenário da emoção desgosto:** O personagem Ana ou Juninho percebe que a rua está com lixos e ratos;
- ❑ **Cenário da emoção medo:** Os personagens Ana e Juninho ficam com medo pois veem um fantasma.

Após a definição dos cenários e enredos, outras animações foram desenvolvidas com a finalidade de tornar as intervenções mais lúdicas e naturais. Na Figura 17, a modelagem de alguns cenários elaborados de forma digital são apresentados: (a) é mostrado a animação da emoção alegria, (b) desgosto, (c) raiva e em (d) surpresa, para ensinar a interpretação das emoções em indivíduos com TEA.

No desenvolvimento desses cenários foram utilizadas as recomendações dos especialistas, pacientes e experiência de *designers* de software para a sua construção, seguindo a abordagem proposta em (MALINVERNI et al., 2017). O principal objetivo é tornar a aplicação mais atraente por meio de técnicas de elicitación e fusão durante a fase de modelagem do software.

### 4.1.3 Protótipo

Nessa etapa, foi desenvolvido um protótipo do jogo, de baixa fidelidade, geralmente, baseado em papel, usando cartões, placas, ladrilhos e/ou miniaturas, do jogo, para avaliá-lo, com objetivo de garantir que a "mecânica" de jogo estivesse adequada e atraente para o público. Com esse protótipo foi possível estabelecer detalhes para a produção da realidade. Na Figura 18 é mostrado o material desenvolvido para a exploração dos requisitos para produção do jogo durante a fase de modelagem.

O desenvolvimento dessa etapa foi realizada juntamente com profissionais da área de saúde, da educação e pais de crianças com TEA. A cada evolução no protótipo o mesmo passava por

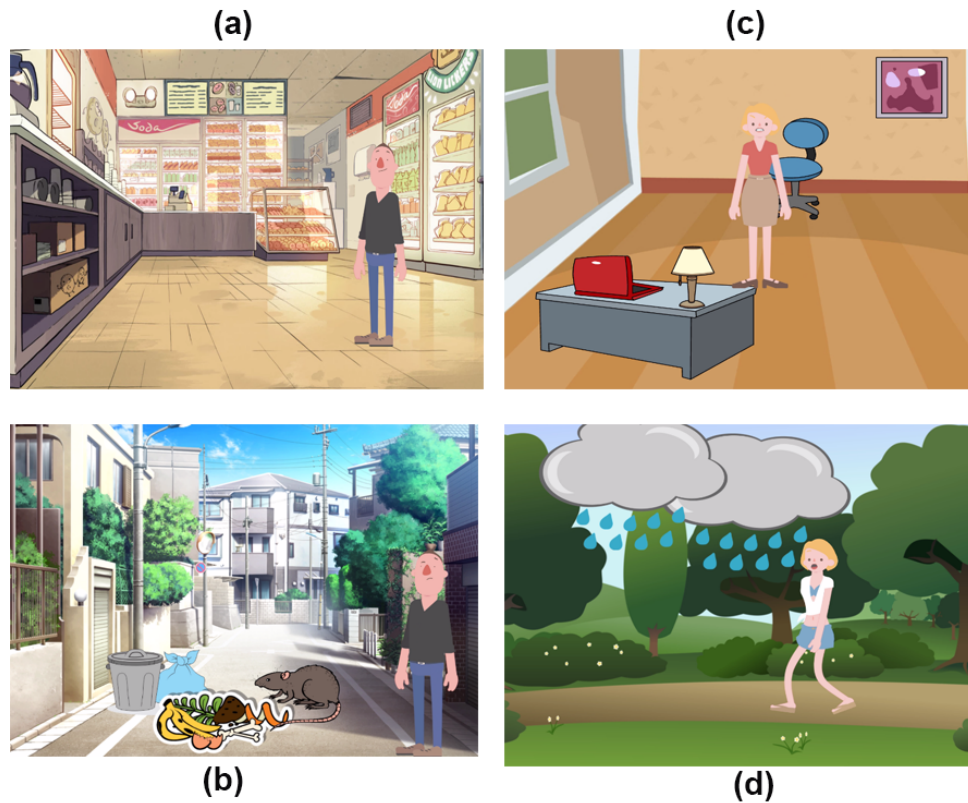


Figura 17 – Animações desenvolvidas para ensinar a interpretação das emoções em indivíduos com TEA: (a) alegria, (b) desgosto, (c) raiva e (d) surpresa.

avaliações nas quais eram sugeridas modificações e atualizações, as quais foram implementadas na versão final do jogo.

#### 4.1.4 Produção

Essa etapa é composta por uma série de passos fundamentais para o desenvolvimento do jogo. Os passos analisados nessa fase foram:

- ☐ Codificação dos protótipos;
- ☐ Desenvolvimento de conteúdo;
- ☐ Textos no idioma do jogo;
- ☐ Especificação da narrativa do jogo;
- ☐ Interface para os usuários;
- ☐ Compatibilidade de *Hardware* e *Software*;
- ☐ Compatibilidade de interface;
- ☐ Artes (desenhos, modelagem 3D) e sons;
- ☐ Manual.



Figura 18 – Tabuleiro desenvolvido durante a produção do jogo.

#### 4.1.5 Modelagem Computacional

Para a modelagem computacional foram utilizadas ferramentas como a plataforma da Unity (UNITY, 2018) e o *software* Blender (BLENDER, 2018), uma ferramenta de criação 3D gratuita e de código aberto. O processo de modelagem 3D envolve alguns conceitos importantes para a criação dos modelos computacionais em 3D. Esses conceitos estão descritos:

- ☐ Coordenadas e Vetores: Tem a função de especificar a localização dos objetos e da cena;
- ☐ Formas 3D: Objetos;
- ☐ Materiais e Texturas: Na renderização volumétrica;
- ☐ Detecção de Colisão: Algoritmos para identificar intersecções de dois corpos sólidos presentes na cena;
- ☐ Objetos e Componentes: Elementos como os personagens, botões, etc;
- ☐ Cenários: Ambiente com características como, luz, som, etc.

Para o desenvolvimento do personagem Juninho foram utilizados um total de 23.680 polígonos e 46.166 triângulos com objetivo de representar um indivíduo com altura de 1,63 metros. Na Figura 19 é demonstrado o personagem modelado em 3D com os vértices sob duas perspectivas.

Na primeira etapa de modelagem cada um dos personagens foi devidamente projetado e modelado em 3D, seguindo todos os procedimentos de modelagem computacional, conforme demonstrado na Figura 20.

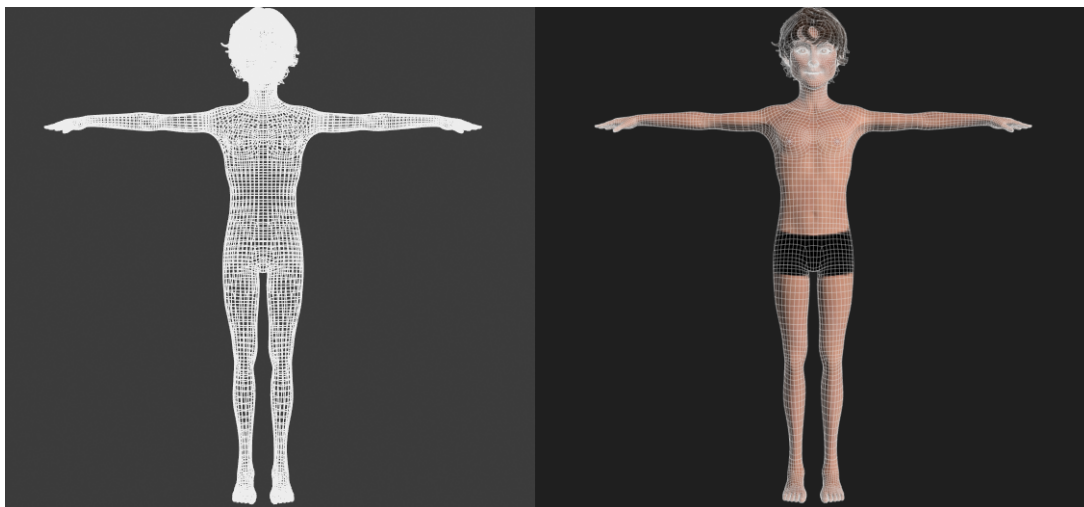


Figura 19 – Processo de modelagem de personagem em 3D para o personagem Juninho.

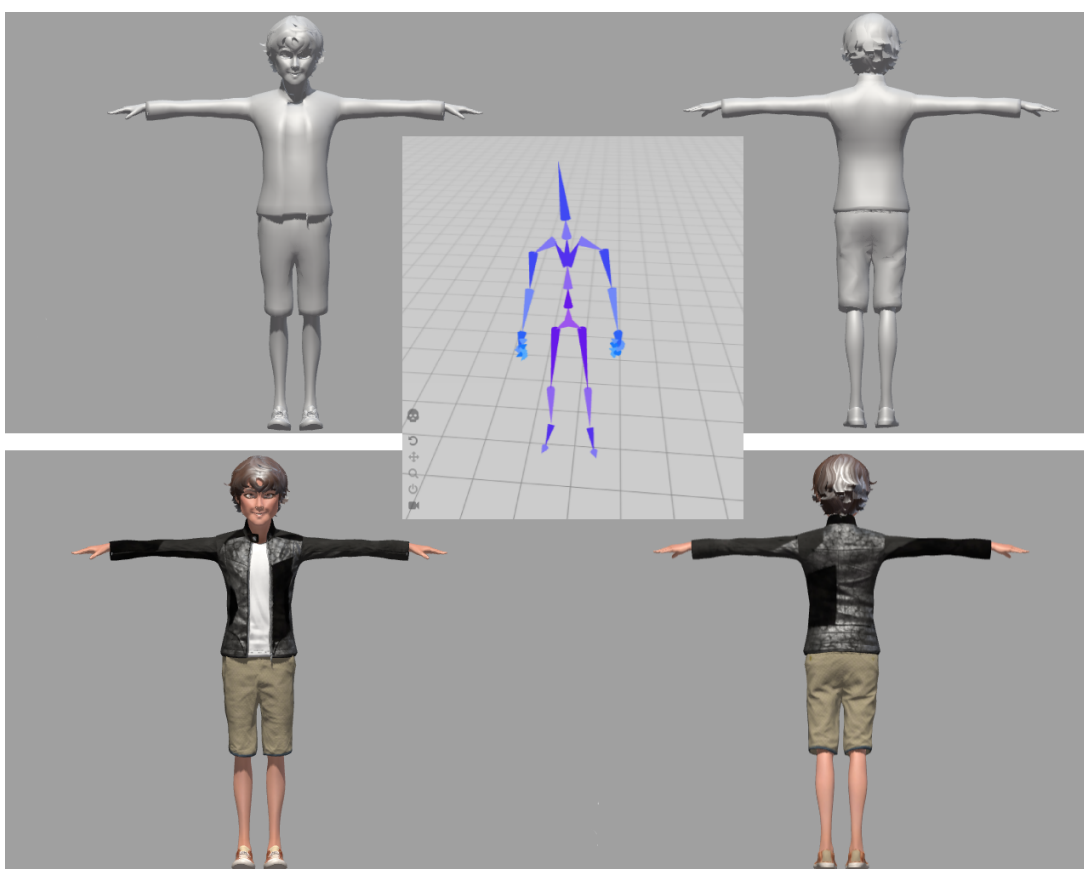


Figura 20 – Desenvolvimento, estruturação e modelagem em 3D do personagem.

Os personagens, Juninho, Ana e Michelzinho, são responsáveis por interagir e demonstrar para os usuários as expressões faciais com o máximo de fidelidade possível, assim como ocorrem nas expressões faciais em seres humanos. Nessa fase de elaboração dos personagens foi usado o software *Makehuman* (MAKEHUMAN, 2021) para a modelagem dessas expressões nos personagens em 3D. O uso dessa ferramenta permitiu a construção dos personagens mais realistas e autênticos às expressões faciais humanas. Na Figura 21 é apresentada uma sequência de emoções de um dos personagens desenvolvido em 3D.

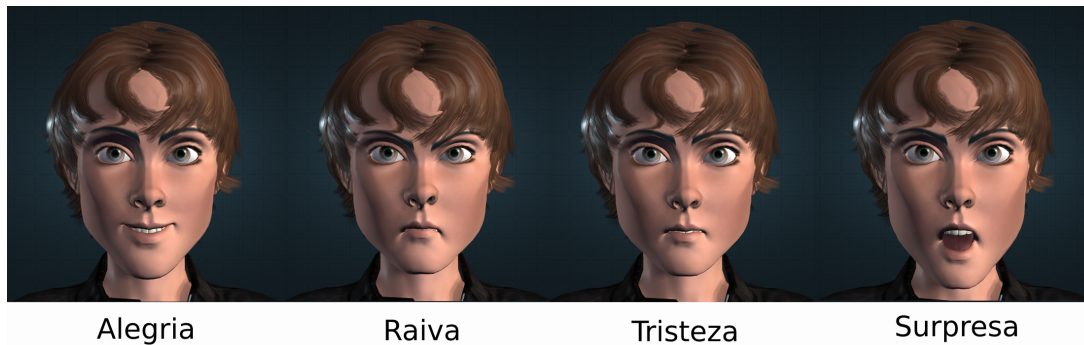


Figura 21 – Personagem desenvolvido em 3D expressando as emoções: alegria, raiva, tristeza e surpresa com base nos pontos chaves propostos por (EKMAN; ROSENBERG, 2005).

As animações dos personagens foram desenvolvidas a partir da análise dos vídeos da base (PANTIC et al., 2005). Essa base de domínio pública possui 2.900 vídeos de um conjunto de 75 pessoas expressando cada uma das emoções de acordo com a teoria *Facial Action Coding System* (FACS) (FRIESEN; EKMAN, 1984). Cada pessoa da base usa de 40 a 45 segundos para expressar cada uma das emoções.

## 4.2 Jogo para detecção de emoções baseadas em expressões faciais

Na Figura 22 é descrita a aplicação denominada “*Face Emotions*”, um jogo sério gratuito, desenvolvido para aprimorar as habilidades emocionais a partir do reconhecimento de expressões faciais de indivíduos com TEA. O jogo proposto tem como objetivo apresentar os personagens (modelos) expressando emoções e com auxílio de um tutor (outro personagem) auxiliar no treinamento das habilidades relacionadas aos gestos da expressão de uma determinada emoção. Esse trabalho foi desenvolvido utilizando o HTML, na versão 5.2 (MOZILLA, 2018), JavaScript (MOZILLA, 2021) a linguagem de programação Python (PYTHON, 2021), versão 3.8. Nesse jogo não foi utilizado nenhum motor predefinido disponível na literatura de forma a permitir a personalização de acordo com as características do público envolvido. O equipamento para realização do processamento e armazenamento dos dados foi um computador com sistema



operacional Linux com 8GB de RAM, 1TB e 8GB de GPU. Os vídeos foram capturados com 30 quadros por segundos em modelo de cores RGB.

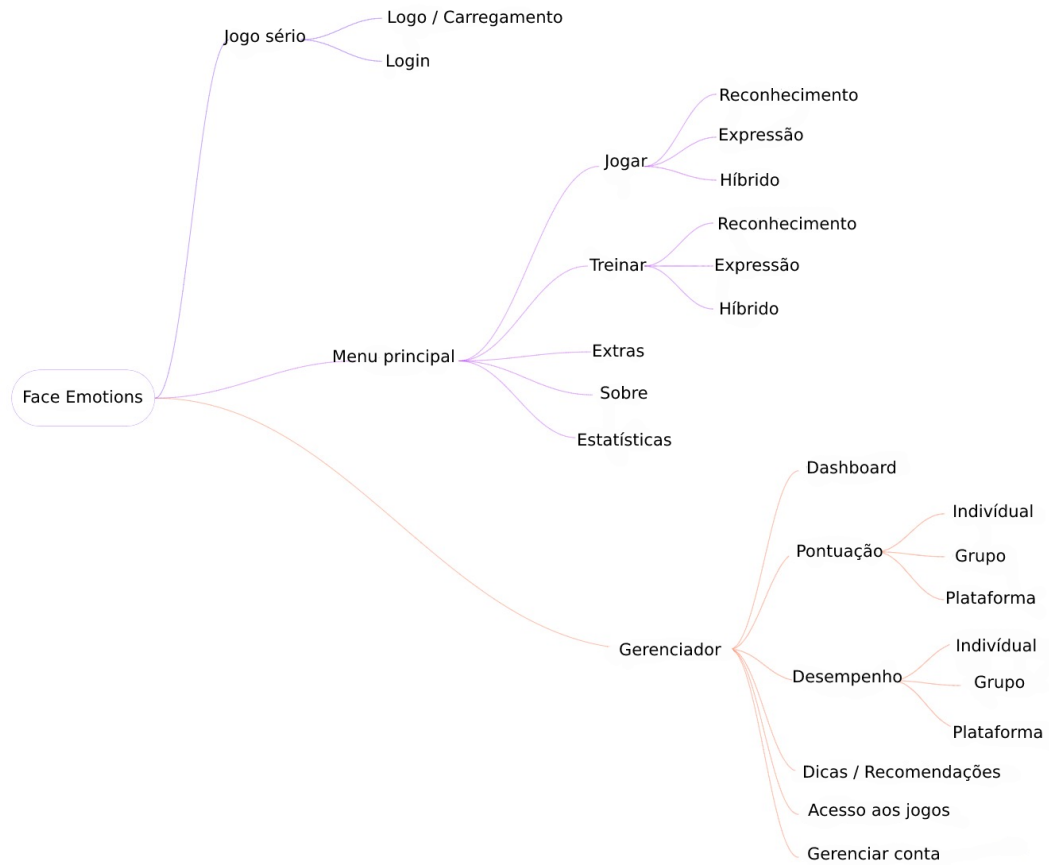


Figura 22 – Mapa mental da plataforma do jogo *Face Emotions*.

Na Figura 23 são apresentados os principais módulos da ferramenta *Face Emotions*: módulo SG, módulo de detecção e classificação e módulo de dados, o qual é usado para o armazenamento e processamento das informações capturadas de cada usuário.

No módulo SG é apresentado ao usuário as instruções iniciais a respeito do jogo. O usuário seleciona um personagem, denominado tutor, que auxiliará durante o processo de aprimoramento das habilidades. Em seguida, a etapa de avaliação do conhecimento das habilidades. Após essa etapa, o usuário recebe novas instruções para expressar as emoções seguindo as informações do tutor. As expressões faciais realizadas pelo usuário são capturadas pela *webcam* e as imagens são enviadas ao módulo de detecção e reconhecimento das emoções. Nessa etapa, são extraídas as informações de regiões de interesse na face do usuário. As informações dos pontos chave da face (do inglês, *facial keypoints* - FKs), o Histograma de Gradientes Orientados (do inglês, *histogram of oriented gradients* - HOG) e características aprendidas pelas camadas iniciais de um modelo CNN são empregadas para detecção e reconhecimento da emoção. Os dados obtidos nesse processo são armazenados para avaliação do especialista. Nessa fase, as informações capturadas são: o nível de atenção, quantidade de acertos, quantidade de erros e o tempo necessário para expressar a emoção (seção parâmetros extras na Figura 23). Uma in-

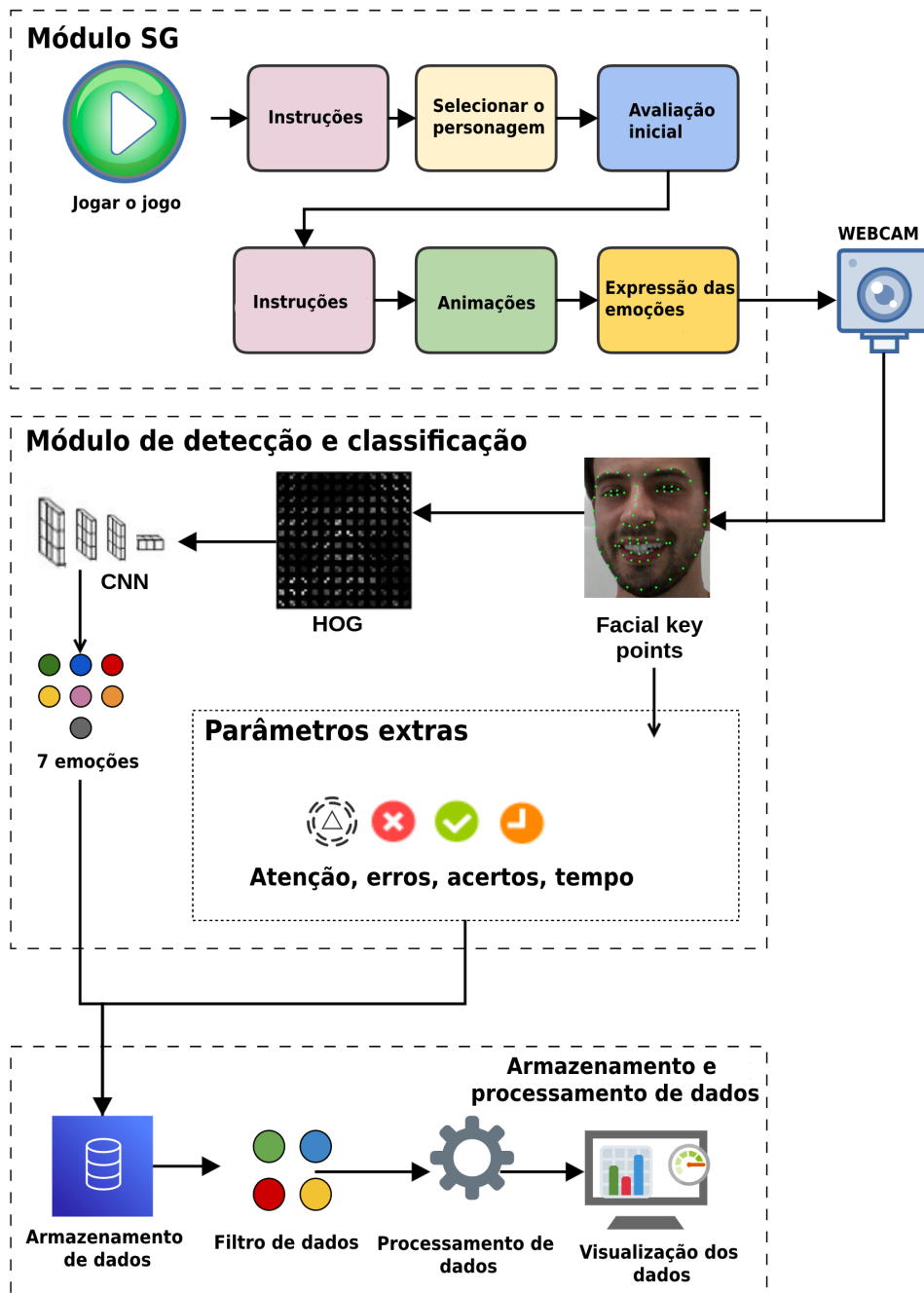


Figura 23 – Fluxograma das etapas da ferramenta proposta para aprimoramento das habilidades das emoções.



terface gráfica permite ao especialista analisar esses dados obtidos em cada sessão de forma a acompanhar o aprimoramento dessas habilidades nos indivíduos com TEA.



Figura 24 – A interface apresenta a personagem Ana escolhida para auxiliar no processo de aprimoramento das emoções.

Na Figura 24 é apresentado um dos personagens, Ana, que auxiliará durante o aprimoramento das habilidades relacionadas as seis emoções básicas. Além disso, os usuários podem modificar os acessórios do personagem (ver Figura 25). O usuário pode escolher os itens: roupas, boné, chapéu, óculos, etc. Essa etapa permite a interação com usuário de forma que usuário possa ter maior motivação e personalização conforme recomendado por (YUSOFF, 2010) na aplicação.

Na parte inicial do jogo, o usuário é submetido a uma avaliação do seu conhecimento sobre emoções. Então, o usuário deve assistir a seis vídeos com uma pessoa representando cada uma das emoções durante o intervalo de tempo de um minuto. Os vídeos com as emoções básicas são apresentados de forma aleatória. Ao final de cada vídeo, três imagens com emoções diferentes são mostradas para o usuário escolher a emoção que estava representada no vídeo, conforme demonstrado na Figura 26. Durante o intervalo de cinco minutos, os participantes escolhem a imagem que corresponde à emoção mostrada no vídeo.

Após essa avaliação, o usuário irá jogar seguindo o caminho no mundo das emoções conforme demonstrado na Figura 27. Em cada parte, uma das emoções básicas será apresentada e o personagem irá auxiliar o usuário na expressão da emoção.

O usuário executará os movimentos para reprodução da emoção e a câmera irá capturar o vídeo para que o algoritmo de detecção e reconhecimento possa analisar a emoção apresentada.



Figura 25 – Interface para personalização do personagem com opções de roupas e outros acessórios (boné, chapéu e óculos).

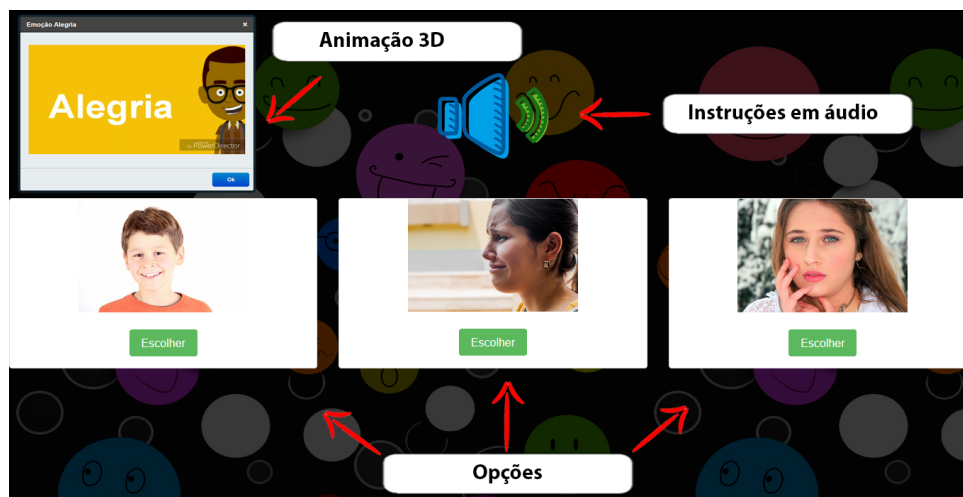


Figura 26 – Interface do jogo para avaliar as competências do conhecimento da emoção.

Para cada região da face representada adequadamente da emoção requerida, a barra de estado é preenchida. Após o preenchimento da barra, o sistema permite ao usuário avançar para a próxima fase em que uma nova emoção será apresentada. Na Figura 28 é ilustrada a interface do jogo com a emoção "alvo" requerida (na parte superior direita em bordas amarela) e na parte superior esquerda, com borda na cor roxa, tem-se o personagem representando a emoção baseada nas informações obtidas do usuário no centro da tela. Os pontos marcados na imagem no centro da tela permitem que o usuário possa acompanhar os pontos para a representação da emoção. As informações como tempo gasto para representação, rastreamento ocular das regiões da tela que foram visualizadas pelo usuário, os músculos ativados da face, os números de acertos e erros são armazenados. Essas informações são empregadas para futura análise do especialista.

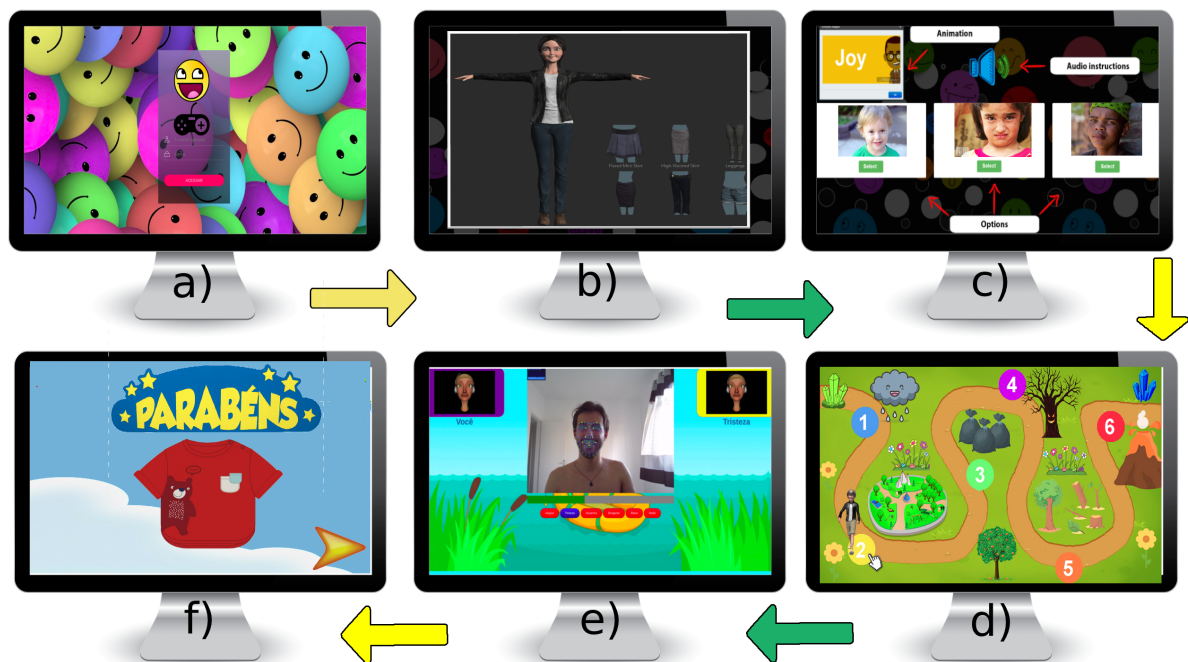


Figura 27 – *Gameplay* do SG proposto: (a) identificação do usuário (ID); escolha de um dos personagens; (c) avaliação das habilidades sobre as emoções básicas; (d) cenário com 6 diferentes interações sobre as emoções; (e) o usuário pode melhorar as habilidades emocionais; e (f) a cada emoção expressa corretamente, um item surpresa é liberado para o personagem que pode avançar no mapa do jogo.

O usuário pode jogar as fases do *Face Emotions* e finalizar a partida quando chega no final do caminho das emoções. Caso queira jogar novamente, ele inicia uma nova partida no jogo.



Figura 28 – Interface de treinamento de competências emocionais baseado nas expressões faciais e funcionalidades.

Quando o usuário apresenta uma sequência de erros na expressão de uma emoção, são apresentadas dicas com o objetivo de ajudar o jogador a representar corretamente essa emoção durante uma partida deste jogo. Contudo, quando o participante não consegue representar a emoção, a ferramenta não permite que o jogador avance para a fase seguinte. O participante é

direcionado para a fase de treino com imagens e animações. Na fase de treino, o tutor expressa a emoção e um som (áudio) descreve o nome da emoção. Há também uma imagem de uma pessoa mostrando essa emoção (ver Figura 29).

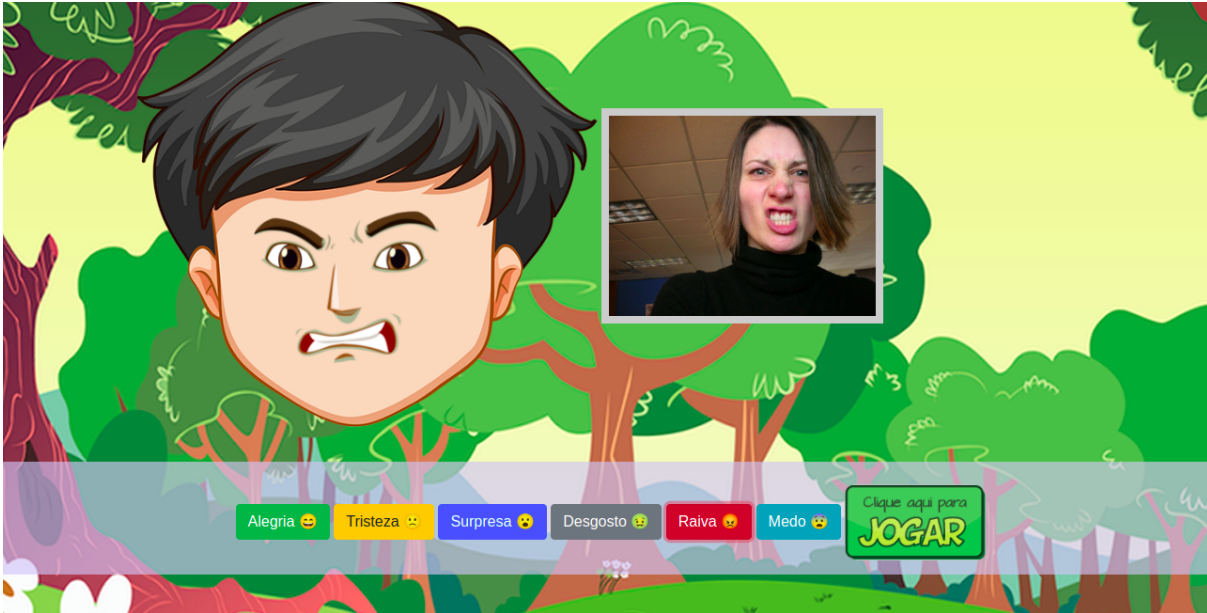


Figura 29 – Interface de treinamento para reconhecimento de emoções e capacidades de expressão.

As seis emoções são apresentadas na fase de treinamento e esse processo é repetido duas vezes para cada emoção. Após essa etapa, é apresentada uma mensagem perguntando se deseja continuar o treinamento. Neste processo, a informação de treinamento é armazenada para os especialistas para as tomadas de decisões do tratamento.

#### 4.2.1 Módulo Detecção e Classificação

Nesse módulo, as expressões da face do usuário são capturadas em vídeo por uma câmera. Esse vídeo capturado no modelo de cores RGB é convertido para a escala de níveis de cinza. Esse processo permite diminuir a escala de cores presente nos quadros conforme relatados, nos estudos propostos por (CANAL et al., 2022; SÁNCHEZ et al., 2011; SHAN; GONG; MCOWAN, 2009).

Esses quadros são repassados ao algoritmo proposto por Viola-Jones para a detecção de características das regiões da face do usuário (VIOLA; JONES, 2004). Essa técnica de detecção permite analisar a aparência dos objetos e busca “aprender” características usando um conjunto de imagens (VIOLA; JONES, 2004; GRACIOSO et al., 2013). Então, inicialmente, é necessário realizar um cálculo em todos os pixels do rosto, e o uso de uma imagem integral permite realizar esse cálculo rapidamente. Nessa etapa,  $I(x,y)$  representa o valor na imagem integral e  $i(x,y)$  representa a imagem de entrada dada por:

$$I(x,y) = i(x,y) + I(x-1,y) + I(x,y-1) - I(x-1,y-1) \quad (2)$$

Com as regiões da face, o método permite extrair as características *Haar-like* em que informações de recursos de bordas, recursos de linha e recursos centralizados podem ser capturadas (ver Figura 30). Após obtenção das características, uma classificação baseada no algoritmo Adaboost, em uma estrutura em árvore denominada cascata de classificadores (ver Figura 31), é empregada (FREUND; SCHAPIRE, 1997). Quando as amostras são aceitas pelo algoritmo, a região é detectada conforme representado pela Figura 32. Nesta fase, as regiões da face são detectadas como sobrancelhas, olhos, lábios, nariz e mandíbula.

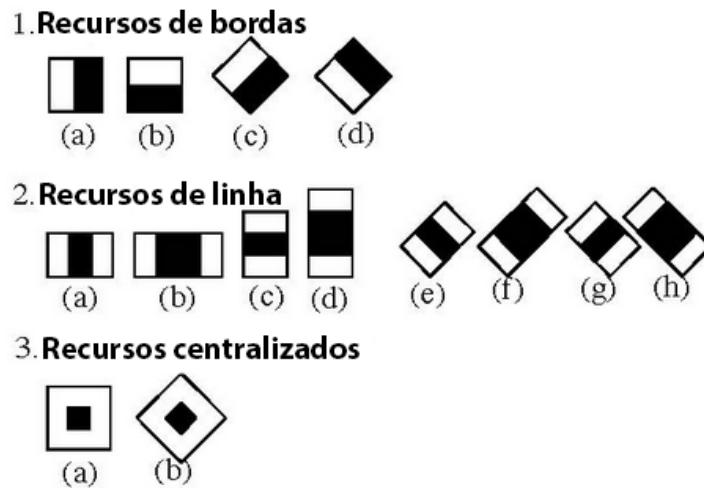


Figura 30 – Informações extraídas com a técnica *Haar-Like*: (a), (b), (c) e (d) são características de borda; (e), (f), (g), (h), (i), (j), (k) e (l) são traços de linha; e (m) e (n) recursos *center surround*. Adaptado de (LIENHART; MAYDT, 2002).

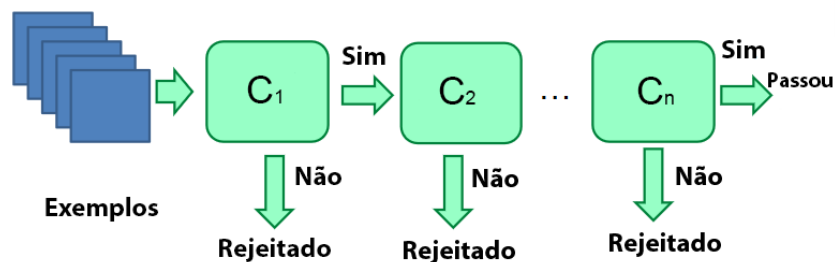


Figura 31 – Exemplo do detector de estruturas baseado no modelo em cascata. Adaptado de (AZCARATE et al., 2005b).

A biblioteca Dlib C++ foi empregada para a obtenção dos pontos faciais (KAZEMI; SULLIVAN, 2014; MUNASINGHE, 2018; CARAMIHALE; POPESCU; ICHIM, 2018). Nessa etapa, a biblioteca Dlib permite capturar 68 pontos da face, os quais são apresentados na Figura 33.

Na face, as regiões referentes aos pontos são:

- ❑ Sobrancelha direita - pontos 17, 18, 19, 20 e 21;
- ❑ Sobrancelha esquerda - pontos 22, 23, 24, 25 e 26;



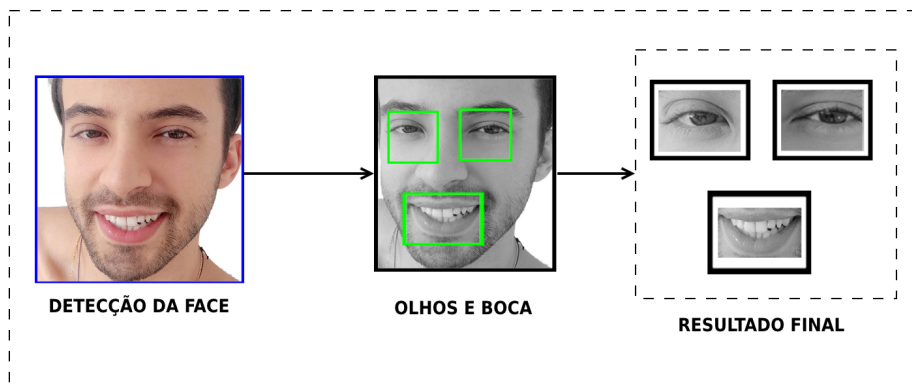


Figura 32 – Etapas das regiões detectadas da face (face, olhos e boca).

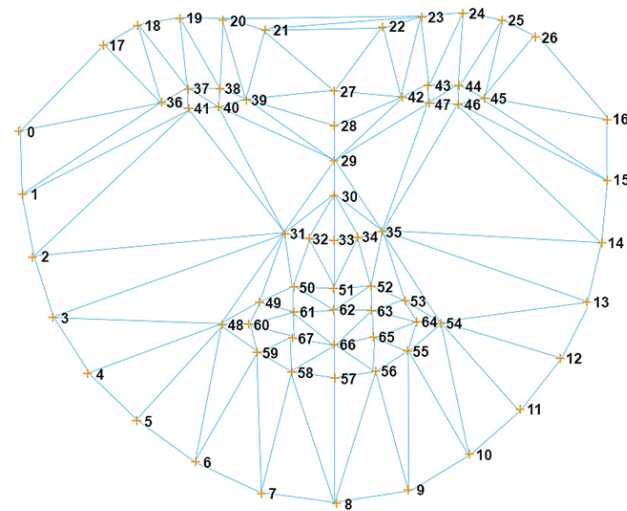


Figura 33 – Marcadores inseridos na imagem do rosto do usuário.

- ❑ Olho direito - pontos 36, 37, 38, 39, 40 e 41;
- ❑ Olho esquerdo - pontos 42, 43, 44, 45, 46 e 47;
- ❑ Nariz - pontos 27, 28, 29, 30, 31, 32, 33, 34 e 35;

Região da boca:

- ❑ Lábio externo superior - pontos 48, 49, 50, 51, 52, 53, 54 e 54;
- ❑ Lábio interno superior - pontos 60, 61, 62, 63 e 64;
- ❑ Lábio interno inferior - pontos 60, 67, 66, 65 e 64;
- ❑ Lábio externo inferior - pontos 48, 59, 58, 57, 56, 55 e 54.

Com o uso da biblioteca *Dlib*, o primeiro conjunto de coordenadas localizado é responsáveis por demarcar todo o contorno na face do usuário, conforme demonstrado na Figura 34(a). Essa

região da face é importante por estar presente em todas as expressões faciais na detecção das emoções básicas (ver Figura 34). O segundo conjunto é responsável por detectar e mapear os olhos, direito e esquerdo, do usuário conforme apresentado na Figura 35. Um exemplo dos marcadores nas regiões das sobrancelhas direita e esquerda na face do usuário é apresentado na Figura 36. Essa região também está presente em parte das expressões faciais, para a detecção das emoções.

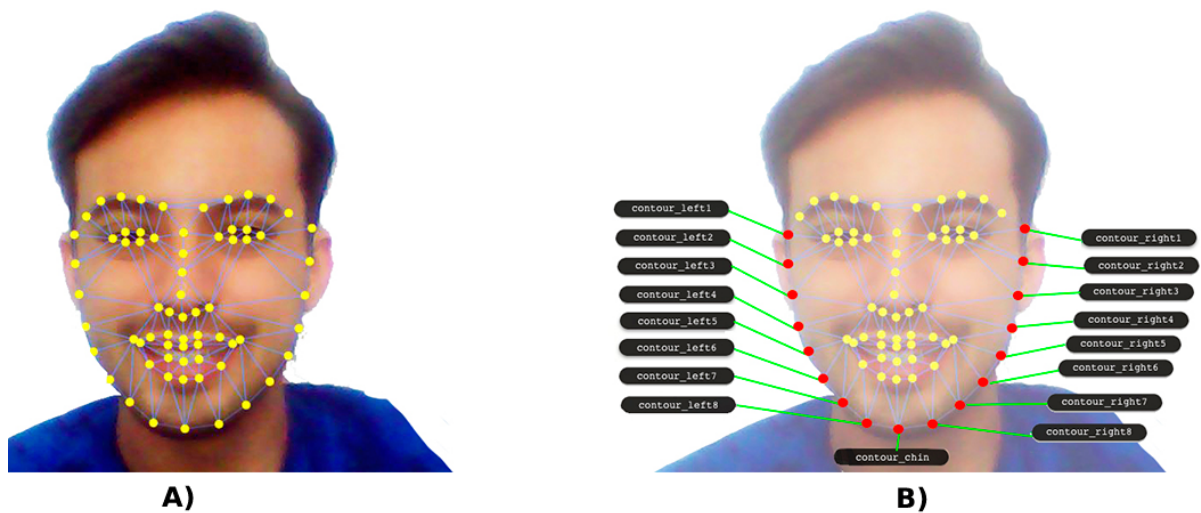


Figura 34 – Demonstração das coordenadas dos pontos: a) pontos marcados em amarelo do contorno da face do usuário; b) pontos que representam o lado direito e esquerdo da face do usuário.



Figura 35 – Exemplo das coordenadas de pontos para detecção do contorno dos olhos direito (a) e esquerdo (b) do usuário.

O quarto conjunto de coordenadas é responsável pelas regiões da boca e nariz presentes na face do usuário. Essas estruturas são importantes para detectar seus movimentos durante representação das expressões faciais, conforme demonstrado na Figura 37.

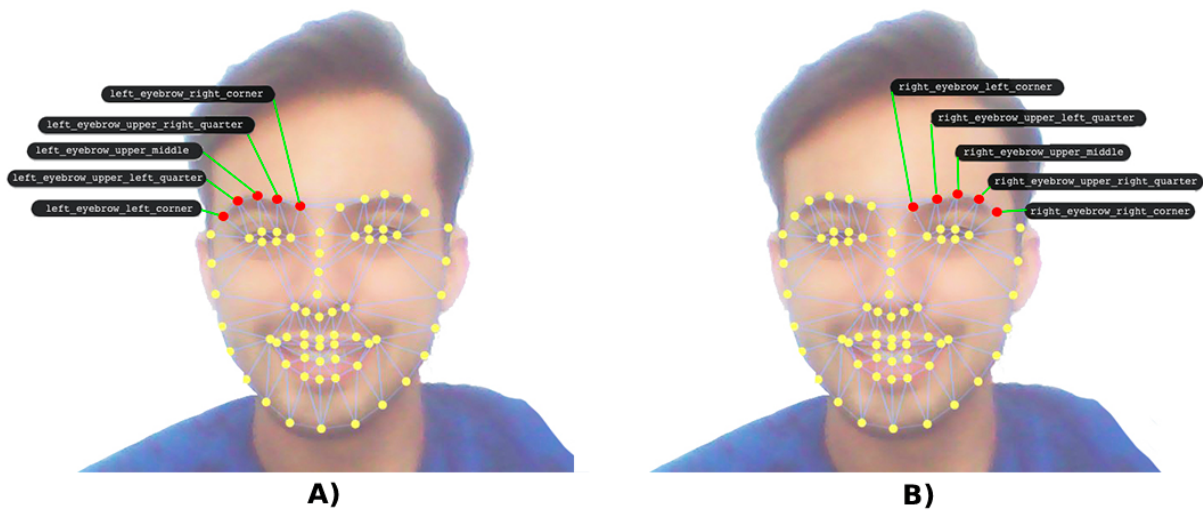


Figura 36 – Pontos responsáveis pelo contorno das sobrancelhas: (a) lado direito do usuário e (b) lado esquerdo do usuário.

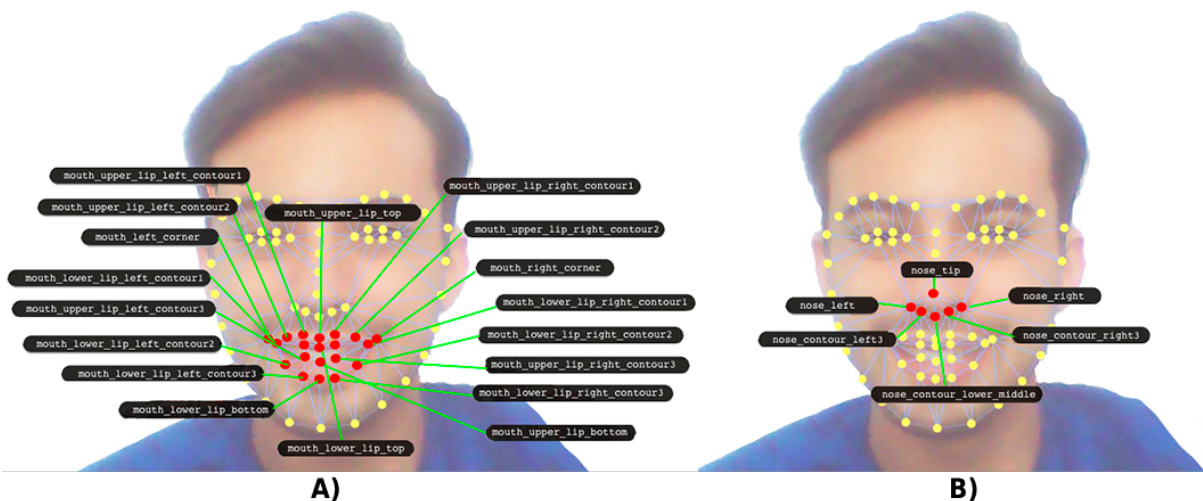


Figura 37 – Demonstração das coordenadas de pontos responsáveis pelo contorno de regiões: (a) boca e (b) nariz.

Esse conjunto de 68 pontos foi selecionado baseado na teoria FACS (EKMAN; FRIESEN, 1978) para representar os pontos faciais. Esses foram determinados de forma empírica para análise das regiões marcadas. Os pontos faciais foram armazenados em um vetor de características a ser analisado na etapa de predição e teste da classificação das emoções. Em seguida, foram calculadas as distâncias Euclidianas entre os pontos. Devido a movimentação dos músculos das regiões da face pelo usuário, o algoritmo de rastreamento do fluxo óptico proposto por (FLEET; WEISS, 2006) foi aplicado para estimar o movimento nos quadros seguintes do vídeo.

As regiões detectadas (do inglês, *regions of interest* - ROIs) na face são ainda avaliadas para extração de características com o método de histogramas orientados a gradiente (do inglês, *histogram of oriented gradients* - HOG) e características aprendidas por meio de um modelo de rede neural convolucional. O método HOG é um descritor de características que permite



determinar as ocorrências de orientação de gradiente em porções localizadas de uma imagem com a finalidade de extrair informações de objetos (GONZALEZ; WOODS, 2018). Os gradientes são calculados em uma imagem por bloco, em que o bloco é considerado uma grade de pixels constituídos a partir da magnitude  $g$  e direção  $\Theta$  da mudança nas intensidades dos pixels (MOHAMMED; MELHUM, 2020). Na Equação 3, os termos  $g_x$  e  $g_y$  são, respectivamente, os componentes, horizontal e vertical, da mudança da intensidade do pixel. O descritor foi calculado sobre blocos (*kernel*) com dimensões de 8x8 pixels.

$$\Theta = \arctan \frac{g_y}{g_x} \quad (3)$$

$$g = \sqrt{g_x^2 + g_y^2} \quad (4)$$

As informações são obtidas a partir do gradiente da imagem contendo a magnitude e direção de cada pixel usando as Equações 3 e 4. O tamanho da ROI usada foi 48 x 48 pixels da face analisada. Os valores para cada pixel em um bloco de 8 x 8 pixels são quantizados em 9 divisões (*bins*) usando os parâmetros obtidos pelo ângulo direcional do gradiente e soma da magnitude dos pixels.

As mesmas ROIs são utilizadas como entrada para uma arquitetura CNN baseada na arquitetura da MobileNet. A MobileNet é um modelo de arquitetura desenvolvida para classificação de imagens que exige menos recursos computacionais, podendo ser empregadas em dispositivos móveis e equipamentos sem placas gráficas com GPU (HOWARD et al., 2017; SADIK; ANWAR; REZA, 2021).

Esse modelo usa camadas convoluções separáveis em profundidade (do inglês, *depthwise separable convolution* - DSC) e hiper-parâmetros denominados multiplicadores de largura e de resolução. Essa técnica permite delimitar os recursos computacionais como latência e tamanho, seguindo as características da aplicação a ser implementada (HOWARD et al., 2017; SADIK; ANWAR; REZA, 2021). As convoluções separáveis em profundidade tem como principal característica dividir uma etapa de convolução tradicional em duas operações, a primeira sendo em profundidade (do inglês, *depthwise convolution*) e a segunda uma convolução de tamanho 1x1 chamada convolução pontual (do inglês, *pointwise convolution*). Esse processo permite reduzir a quantidade de parâmetros das camadas convolucionais. Na Figura 38 é mostrada a diferença entre um filtro de convolução tradicional e um filtro DSC, onde o termo  $D_k$  é o tamanho do *kernel*,  $M$  é o número de canais de entrada e  $N$  é o número de canais de saída.

A arquitetura de uma MobileNet possui 28 camadas das quais apenas a primeira é uma convolução tradicional, as demais são baseadas nas convoluções de profundidade e pontual. Nesse modelo há uma etapa de *batch normalization* e uma operação de ativação após operações de convolução. O *batch normalization* é uma técnica empregada para padronizar automaticamente as entradas e permite contribuir na etapa de treinamento do modelo, podendo até melhorar o desempenho por meio de uma ação de regularização. A função de ativação tem como função transformar a entrada. A função de ativação usada foi a linear retificada ou *Rectified Linear*

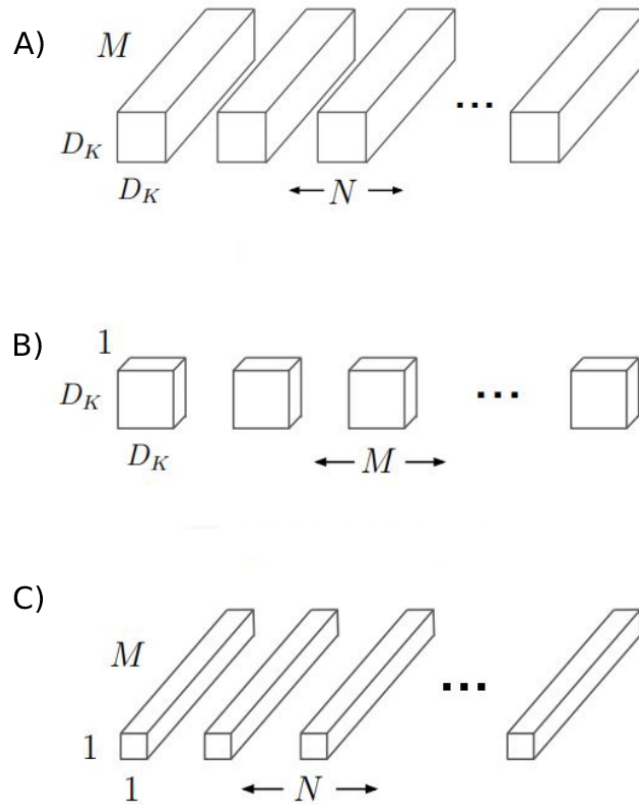


Figura 38 – Representação dos filtros convolucionais, em que em (A) representa um método tradicional. As operações em dois estágios em que a convolução em profundidade dada em (B) e a convolução ponto a ponto dada em (C) para criar um filtro. **Fonte:** (HOWARD et al., 2017)

*Unit* (ReLU) (HOWARD et al., 2017). As camadas denominadas *pooling* têm como função simplificar as informações na saída da camada convolucional. Nessa etapa, foi empregada a operação de *max-pooling*, no qual apenas o maior valor da unidade é passado para a saída. Essa sumarização de dados serve para diminuir a quantidade de pesos a serem aprendidos e também evitar *overfitting*. Nas etapas finais foram empregadas o *flattening* e a camada softmax para a classificação (HOWARD et al., 2017).

No modelo proposto foi realizada uma adaptação em que três camadas totalmente conectadas foram inseridas conforme demonstrado na Figura 39. As características obtidas pelas camadas convolucionais da rede CNN foram agregadas as características obtidas com as informações extraídas com as características obtidas por *face keypoints* e HOG nomeados como descritores *handcrafts* para avaliação na camada de classificação. Segundo os autores em (BASHA et al., 2020), as técnicas *handcrafts* podem auxiliar a CNN na capacidade de aprender os recursos específicos e assim aprimorar os resultados.

Para o desenvolvimento da CNN foi utilizado a linguagem de programação Python com TensorFlow, uma biblioteca de código aberto projetada para aprendizado de máquina que pode ser aplicado em diversas áreas (GOOGLE, 2021). Durante o treinamento foi necessário em

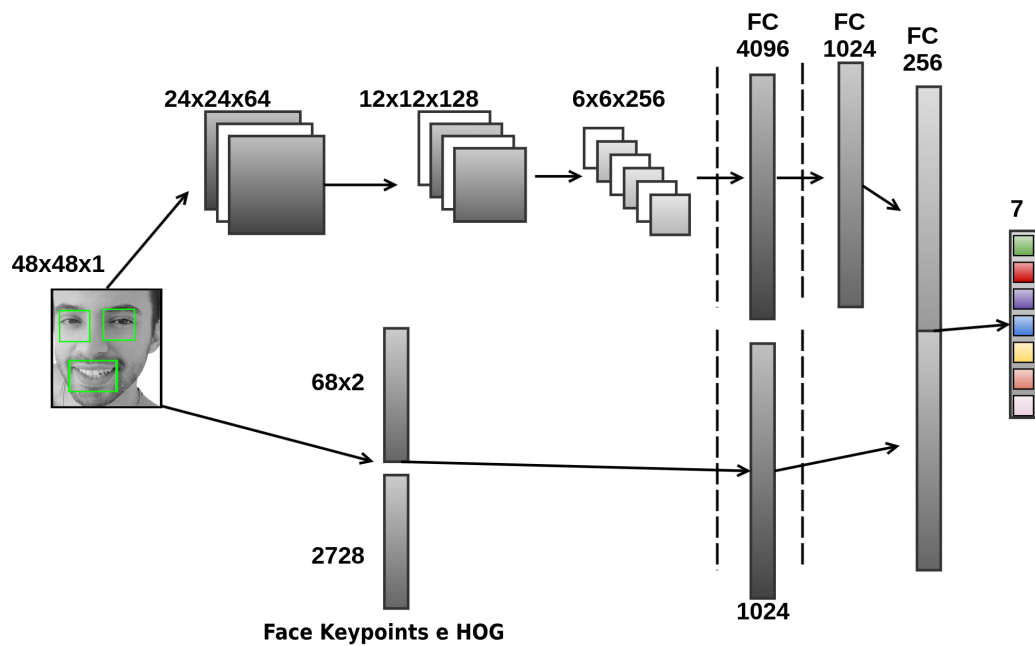


Figura 39 – Modelo proposto para obtenção das informações da face para classificação da emoção.

torno de 2 horas e 30 minutos para o treinamento do modelo proposto.

#### 4.2.2 Módulo Dados

Essa ferramenta possui ainda um módulo com uma interface que permite analisar os dados que foram capturados nos processos de aprimoramento das habilidades das emoções (ver Figura 40). O especialista pode usar esses dados para acompanhar a evolução em diversas fases das sessões de tratamento de um indivíduo. Por meio desse recurso é possível avaliar a quantidade de sessões, as emoções em que os usuários tiveram mais dificuldades, tempo empregado para expressar as emoções e músculos ativados em cada uma das emoções.

Nesse módulo todas as informações capturadas são armazenadas em um banco de dados relacional utilizando o sistema gerenciador de banco de dados MYSQL (MYSQL, 2018). Essas informações são armazenadas e podem ser consultadas posteriormente para avaliação do histórico de evolução do indivíduo.

Na Figura 41 são apresentados as tabelas e os seus respectivos campos desenvolvidos para o módulo de dados da aplicação. Para o módulo de dados foi desenvolvido um total de cinco tabelas de banco de dados, são elas: a tabela paciente, responsável por armazenar os dados do mesmo bem como o número identificador dos usuários na plataforma, aplicadores com a finalidade de registrar os dados dos especialistas envolvidos na intervenção. Uma tabela para armazenar os movimentos dos músculos da face durante a aplicação para a análise de dados e por último uma tabela responsável por armazenar os erros dos pacientes durante as aplicações.

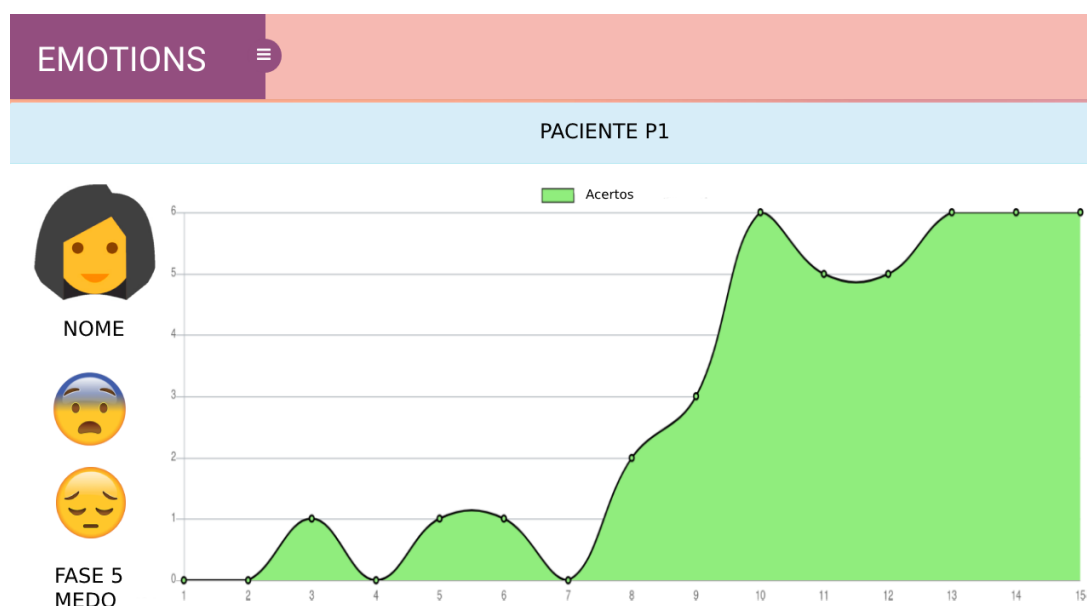


Figura 40 – Interface de (*dashbord*) com as informações de desempenho do usuário fornecido ao especialista.

<b>pushsist_autismo face_erro</b> cod : int(11) emocao : varchar(100) paciente : varchar(100)	<b>pushsist_autismo face_emocao3</b> cod : int(11) data : varchar(100) alegria : varchar(100) tristeza : varchar(100) raiva : varchar(100) medo : varchar(100) desgosto : varchar(100) surpresa : varchar(100) emocao : varchar(100) paciente : varchar(100) smile : varchar(100) inerbrow : varchar(100) browrise : varchar(100) nose : varchar(100) uperlip : varchar(100) lipcorner : varchar(100) chinrise : varchar(100) lippuker : varchar(100) lippress : varchar(100) lipsuck : varchar(100) mouthopen : varchar(100) smike : varchar(100) eyeclose : varchar(100) atention : varchar(100)
<b>pushsist_autismo aplicadores</b> cod : int(11) nome : varchar(100) cpf : varchar(100) telefone : varchar(100) email : varchar(100) foto : varchar(100) senha : varchar(100)	
<b>pushsist_autismo pacientes</b> cod : int(11) medico : varchar(100) nome : varchar(100) responsavel : varchar(100) email : varchar(100) telefone : varchar(100) idade : varchar(100) nivel : varchar(100) level : varchar(100) sexo : varchar(100)	

Figura 41 – Tabelas e seus respectivos campos desenvolvidos para o módulo de dados.

### 4.3 Métricas de Avaliação

As métricas empregadas para avaliação do desempenho da ferramenta foram a a revocação, precisão, acurácia e o F1-Score. A medida revocação permite avaliar o número de instâncias positivas que foram corretamente classificadas. Também foram utilizados os valores da matriz de confusão para o cálculo da precisão, o que pode representar o valor preditivo positivo. A acurácia permite mostrar o desempenho geral do modelo que foi classificado corretamente entre os grupos. Já a medida  $F1 - score$  é definida pela Equação 5.

$$F_1 = 2 * \frac{\text{Precisão} \times \text{Revocação}}{\text{Precisão} + \text{Revocação}} \quad (5)$$

A medida F1-score é necessária quando se deseja buscar um equilíbrio entre as métricas precisão e revocação, que se adaptam à solução como pode ser observado nos estudos de (VALSTAR; PANTIC, 2012a; VALSTAR et al., 2017; ACEVEDO et al., 2017).

Para avaliar o desempenho da aplicação na capacidade de auxiliar indivíduos com TEA de desenvolver e aperfeiçoar as habilidades emocionais foram recrutados de acordo as recomendações e orientações obtidas pelo comitê de ética em pesquisa da Universidade Federal de Uberlândia, sob o número CAAE: 82555417 000005152. Neste trabalho, dois cenários com diferentes características para os indivíduos com TEA foram investigados.

Para realizar a avaliação dos participantes foi aplicado o método *multiple baseline design*, o qual permite analisar por meio de características e configurações, o período anterior e posterior do tratamento de indivíduos. Esse modelo é utilizado em pesquisas médicas, psicológicas e biológicas, especialmente para os indivíduos com TEA, conforme relatados nos estudos (CHEN; LEE; LIN, 2016; APPLE et al., 2005; INGERSOLL et al., 2005; CHEN; LEE; LIN, 2015). Neste trabalho, esse método foi empregado com objetivo de investigar as intervenções para o aprendizado de emoções nesses indivíduos. Esse método é dividido em três fases: a) a referência (baseline), que permite coletar informações sobre o conhecimento dos indivíduos; b) a intervenção, que possibilita obter os dados de uso da ferramenta; e c) a manutenção, fase realizada após a intervenção para reavaliar o desempenho dos participantes.

No estágio de referência são coletados informações sobre o conhecimento dos indivíduos. Cada sessão consiste em assistir os seis vídeos com uma pessoa representando cada uma das emoções básicas durante o intervalo de tempo de um minuto. Os vídeos com as emoções básicas são apresentados de forma aleatória. Ao final de cada vídeo, três imagens com emoções diferentes são mostradas para que o usuário escolha a emoção apresentada no vídeo. Durante o intervalo de cinco minutos, os participantes devem escolher a imagem que corresponde à emoção mostrada no vídeo. As sessões dessa fase ocorrem uma vez por dia num intervalo de cinco vezes na semana durante aproximadamente 30 minutos.

Na primeira sessão da fase de intervenção, os indivíduos são orientados pelo especialista sobre SG, onde são apresentadas as animações com um personagem 3D expressando a emoção alvo. Informações sobre a pontuação e a funcionalidade da interface também são fornecidas

aos participantes. Os participantes são orientados a assistir os vídeos e reproduzir as emoções. Usando uma *webcam*, mímicas faciais inferidas pelo usuário são exibidas pelo personagem. O usuário tem cinco minutos para representar cada emoção corretamente. O usuário tem como observar a representação da expressão da emoção por meio de uma barra de progresso. Quando o usuário expressa corretamente a emoção, o SG fornece uma recompensa pelo sucesso na representação. Essa etapa foi realizada uma vez por dia da semana, durante cinco vezes por semana, em uma sessão de aproximadamente 30 minutos. Os vídeos com as emoções básicas são apresentados de forma aleatória. Os personagens usados na sessão de intervenção são diferentes daqueles usados nas fases de referência e manutenção.

As sessões de manutenção são conduzidas quatro semanas após a fase de intervenção para demonstrar as habilidades emocionais dos indivíduos. Durante a avaliação, todos os participantes assistem aos vídeos com as emoções apresentadas de forma aleatória. Os participantes escolhem a imagem que corresponde à emoção mostrada no vídeo. Essas sessões ocorrem uma vez por dia cinco vezes numa semana com duração de aproximadamente 30 minutos.

---

## Resultados e Discussão

Esse capítulo apresenta os resultados dos experimentos realizados em dois estágios do estudo: a Seção 5.1 mostra os experimentos e a avaliação das técnicas empregadas para o reconhecimento das emoções, a partir das expressões faciais em bases de domínio público, com faces de diversas pessoas. Na segunda parte, na Seção 5.2, é apresentado os experimentos, as contribuições e limitações da aplicação das ferramentas desenvolvidas para auxiliar no desenvolvimento das habilidades emocionais em indivíduos com TEA, em dois diferentes ambientes de investigação (educacional e clínico).

### 5.1 Avaliação dos Algoritmos de Detecção e Reconhecimento de Emoções em Bancos de Dados de Domínio Público

Para avaliação do método de detecção e reconhecimento das emoções foram utilizados quatro bases de dados de domínio público. A base de imagens denominada CK+ (KANADE; COHN, 2005), possui imagens de pessoas expressando as emoções básicas, as quais são divididas em: alegria (324), tristeza (253), raiva (183), surpresa (328), desgosto (182) e medo (182). Essa base foi empregada em trabalhos na literatura para avaliação de algoritmos de detecção e reconhecimento de emoções conforme demonstrados em diversos estudos (VALSTAR; PANTIC, 2012b; GHIMIRE et al., 2015; ACEVEDO et al., 2017). Essa base de dados é composta por vídeos/imagens de mais de 200 adultos com idades entre 18 e 50 anos de indivíduos euro-americanos e afro-americanos. As imagens foram feitas em períodos, onde o processo de captura da emoção começa com o rosto em posição neutro e capta a transição para a emoção exigida. Essas imagens foram capturadas em tons de cinza ou RGB com resolução de  $640 \times 490$  ou  $640 \times 480$  pixels.

A segunda base investigada foi a *Facial Emotion Recognition* (FER 2013) (GOODFELLOW et al., 2013), que possui um total de 35.887 imagens de faces em tons de cinza com resolução  $48 \times 48$  pixels, divididas nas seguintes categorias: raiva (4593), (547) desgosto, (5121) medo, (8989) alegria, (6077) tristeza, (4002) surpresa e (6198) neutra. As faces foram capturadas

de maneira que ficassem centralizadas e que ocupassem a mesma proporção de espaço em cada imagem. As imagens não estão distribuídas em estados temporais sendo o estado final da representação da emoção para cada uma.

A terceira base investigada foi a *Real-world Affective Faces Database* (RAF-DB) (LI; DENG, 2019). Essa é uma base de imagens de expressões faciais com aproximadamente 30.000 imagens faciais obtidas por meio da internet em padrão de cores RGB, redimensionadas para a escala 640x480 pixels. Desse conjunto de imagens, o autor disponibilizou um total de 12.271 imagens para estudos. As imagens foram divididas em: surpresa (1290), medo (281), desgosto (717), alegria (4772), tristeza (1982), raiva (705) e neutra (2524). Uma característica importante dessa base é que as imagens nesse banco de dados possuem grande variabilidade na idade dos participantes. Além disso, variações em relação etnia, poses da cabeça, condições de iluminação, oclusões, por exemplo, óculos, pêlos na face ou auto-oclusão, operações de pós-processamento, como filtros e efeitos especiais, etc.

A quarta base empregada nesse estudo foi a *MMI Facial Expression Database* (MMI) (VALSTAR; PANTIC, 2010). Essa base consiste em mais de 2.900 amostras com alta resolução de 640x480 pixels em modelo de cores RGB. Nessa base apenas 174 amostras continham as anotações necessárias para a classificação das emoções. As imagens utilizadas estão distribuídas da seguinte maneira: alegria (35), desgosto (25), medo (24), raiva (27), surpresa (35) e tristeza (28).

Nas Figuras 42(a), 42(b), 42(c) e 42(d) são apresentados exemplos de indivíduos expressando emoções das bases de dados utilizadas CK+, FER2013, RAF e MMI, respectivamente. As imagens presentes nessas bases possuem um número relevante de distribuição das características relacionadas a gênero, idade e etnia dos indivíduos. Essas características contribuem para avaliação da robustez e generalização dos algoritmos sobre os diversos contextos.

O modelo proposto foi investigado sobre as base de dados com propósito de analisar o comportamento dos algoritmos no processo de detecção e reconhecimento das emoções. Nesse experimento uma investigação foi realizada em relação a associação do grupo de características extraídas empregadas para classificação das emoções. A Tabela 3 apresenta os resultados obtidos com a métrica acurácia média para as emoções com os diversos bancos de imagens. Os resultados mostram que a agregação dos descritores contribuíram para elevar os valores das métricas. A composição entre os descritores aprendidos dados pelo modelo CNN e os descritores *handcraft* permitiram um aumento no desempenho do método proposto para as diversas bases de imagens investigadas.

Os valores presentes na Tabela 3 mostram que o método proposto baseado em modelos CNNs obteve melhor desempenho no banco de imagens CK+ em todas as combinações investigadas (83,3%, 91,2% e 98,8%). Com isso, essa base foi selecionada para comparações e análises do método proposto.

Nas Figuras 43 e 44 são apresentados os gráficos de aprendizagem da rede, com as métricas de acurácia e função de perda (*loss*), respectivamente, para o modelo CNN investigado. É



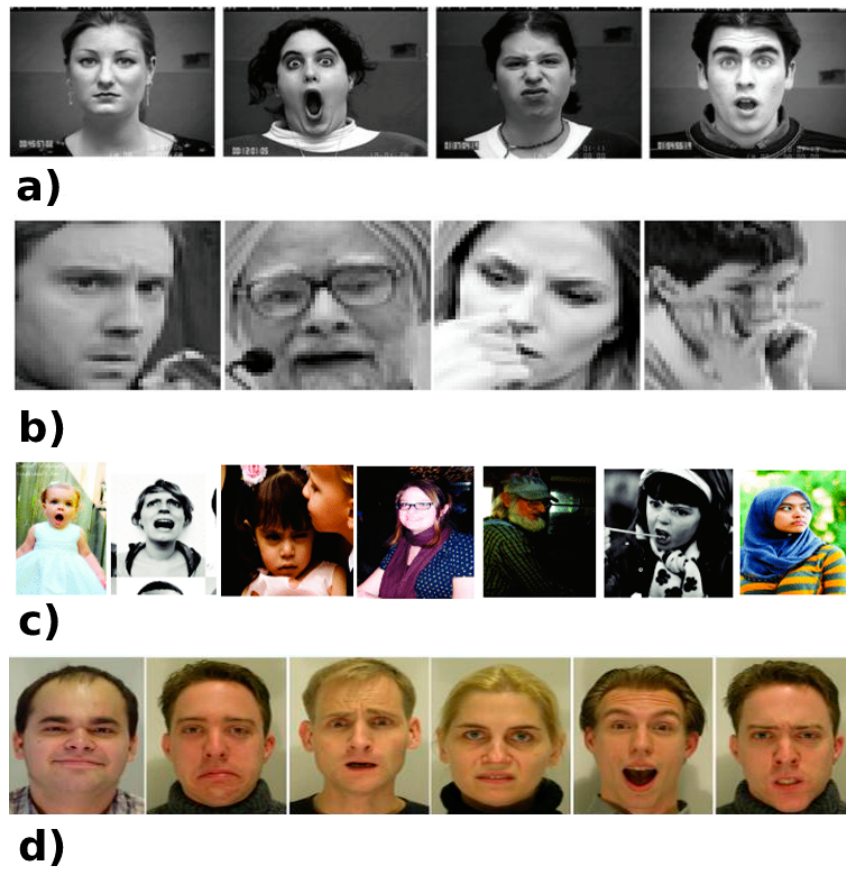


Figura 42 – Exemplo de imagens de pessoas expressando emoções para as bases: (a) CK+, (b) FER2013, (c) RAF e (d) MMI.

Tabela 3 – Resultados obtidos pelo método proposto nas bases: CK+, FER2013, RAF-DB e MMI.

Modelos Propostos	Bancos de Dados			
	CK+	FER2013	RAF-DB	MMI
CNN	83,3%	71,3%	82,6%	74,5%
CNN + <i>Facial keypoints</i>	91,2%	73,1%	85,65%	77,2%
CNN + <i>Facial keypoints</i> + HOG	<b>98,8%</b>	78,8%	88,47%	79,87%

Tabela 4 – Resultados dos experimentos avaliando os parâmetros da CNN.

Épocas	Dropout	FC	Acurácia (%)
500	1	2	85,15
500	2	3	86,23
600	1	2	86,54
600	2	3	89,33
700	1	2	90,47
700	2	3	94,58
800	1	2	95,95
800	2	3	97,68
900	1	2	97,50
900	2	3	97,33
1000	1	2	98,23
<b>1000</b>	<b>2</b>	<b>3</b>	<b>98,80</b>
1000	3	2	98,77
1000	3	3	98,51

Tabela 5 – Resultados obtidos nos experimentos para diferentes modelos de CNN: Modelo proposto, MobileNet, e Inception v3 na base CK+.

Modelo	Acurácia	Sensibilidade	Especificidade
Modelo proposto	98,80%	98,70%	99,75%
MobileNet	81,82%	77,81%	96,39%
Inception v3	91,33%	90,81%	98,38%

possível observar que as taxas de acurácia e de loss se mantiveram estáveis em treinamento e teste.

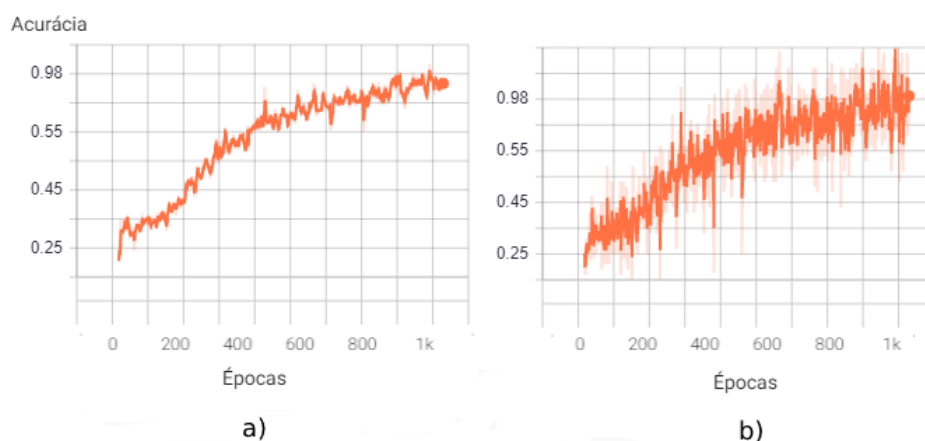


Figura 43 – Valor da acurácia em relação ao número de épocas empregada para avaliação: (a) treinamento e (b) teste.

Para a definição do modelo da CNN em conjunto com os descritores *handcraft*, os parâmetros como: o número de camadas *Fully Connected* (FC) e o número de *dropouts* foram investigados de forma empírica em cada uma das bases citadas anteriormente. A Tabela 4 mostra os

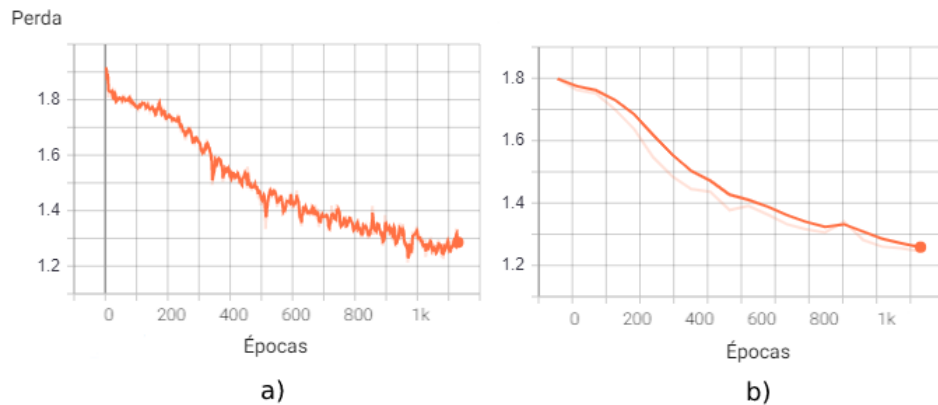


Figura 44 – Valor da minimização de erro (*Loss*) do modelo proposto: (a) treinamento e (b) teste.

resultados obtidos com os modelos CNN e abordagens *handcraft* com as variações relacionadas a investigação de parâmetros para determinar o melhor valor de parâmetros das redes.

Observando os resultados apresentados na Tabela 4, a melhor configuração obtida para a CNN foi de 98.80%, de acurácia com 1000 épocas, duas camadas *dropout* e três camadas totalmente conectadas (FC). Na Tabela 5 é apresentado o modelo proposto em relação a comparação com outras arquiteturas clássicas na literatura para classificação de dados. Nesse caso pode ser observado que o modelo proposto foi 7,5% superior ao Inception e 16,98% melhor que o modelo tradicional da MobileNet.

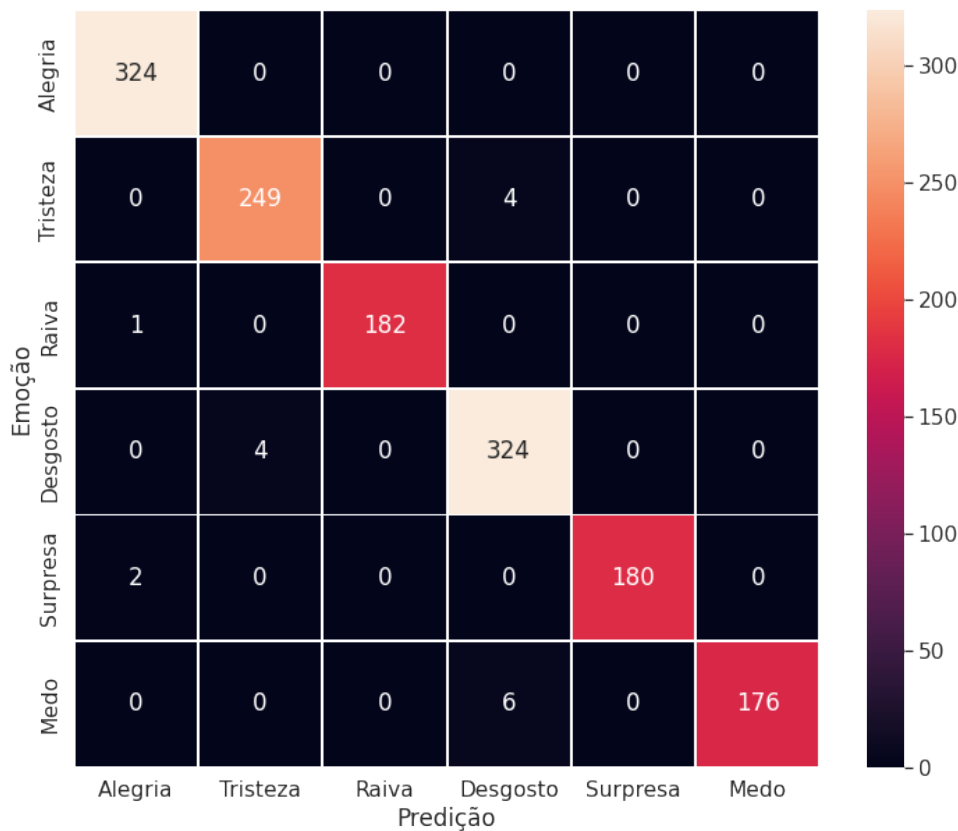
As métricas quantitativas para a avaliação da etapa de detecção e reconhecimento para cada uma das emoções são apresentados na Tabela 6 para a base de imagens CK+. Esses resultados mostram que as medidas inferiores ocorreram com as emoções tristeza, medo e desgosto. Existem emoções em que parte dos sistemas presentes na literatura apresentam dificuldades de classificação semelhantes ao método proposto. Isso foi observado nos estudos propostos por Jain et al. (JAIN; SHAMSOLMOALI; SEHDEV, 2019) e Wu e Lin (WU; LIN, 2018). De maneira geral, o método proposto apresenta resultados promissores permitindo sua utilização como parte do jogo SG para a etapa de detecção e reconhecimento de emoções durante o processo de aprimoramento de habilidades relacionadas à emoção.

Tabela 6 – Resultados obtidos com as métricas acurácia, precisão, especificidade e F1-Score para cada uma das emoções básicas com o modelo proposto na base de imagens CK+.

Emoção	Acurácia	Sensibilidade	Especificidade	F1-Score
Alegria	100,0%	100,0%	99,0%	1,00%
Tristeza	98,0%	98,0%	98,0%	0,98%
Raiva	100,0%	99,0%	100,0%	1,00%
Desgosto	98,0%	99,0%	97,0%	0,98%
Surpresa	99,0%	99,0%	100,0%	0,99%
Medo	98,0%	97,0%	100,0%	0,98%

De forma geral, assim como a métrica acurácia que proporcionou resultados médio de 98,8%, a métrica precisão responsável pela determinação dos valores preditos positivos resultou numa média de 98,66%. Os valores da métrica sensibilidade, ou seja, a fração das instâncias relevantes que foram recuperadas foi de 99,00%. A medida F1-Score é a média ponderada de precisão e revocação. Portanto, essa pontuação considera os falsos-positivos e os falsos-negativos, essa métrica obteve um valor médio de 0,9883%. Intuitivamente, o F1-Score é geralmente mais importante que a precisão, especialmente quando são explorados dados que apresentam desbalanceamento entre as classes investigadas. Como acontece com o banco empregado nesta fase, em que tem um número não balanceado de amostras para as emoções básicas.

Na Figura 45 é apresentado uma matriz de confusão dos resultados obtidos nos experimentos para a base CK+ de imagens. Na matriz é possível observar novamente que as emoções com maior número de erros são: tristeza, desgosto e o medo.



Acurácia média: 98.8%:

Figura 45 – Matriz de confusão do modelo proposto em amostra obtida do conjunto de dados CK+.

Para demonstrar de maneira visual as decisões tomadas sobre o modelo proposto aqui é apresentado o *Gradient-weighted Class Activation Mapping* (Grad-CAM). Esse método usa os gradientes, fluindo para a camada convolucional final produzindo um mapa de localização destacando as regiões importantes na imagem para a classificação (GILDENBLAT; CONTRIBUTORS, 2021). Na Figura 46 é apresentado a aplicação do Grad-CAM na emoção surpresa

Tabela 7 – Comparação de métodos usados pelos trabalhos relacionados com a base de imagens CK+.

<b>Autores</b>	<b>Extração de característica</b>	<b>Classificador</b>	<b>Acurácia</b>
Bilkhu et al. (BILKHU; GUPTA; SRIVASTAVA, 2019)	Regressão em cascata	SVM	89,00%
Wu and Lin (WU; LIN, 2018)	Mapeamento de recursos adaptável	Softmax	89,84%
Noor et al. (NOOR et al., 2020)	HOG	SVM	90,79%
Jain et al. (JAIN; SHAMSOLMOALI; SEHDEV, 2019)	Recursos de aprendizagem	Softmax	93,24%
Ding et al. (DING; ZHOU; CHELLAPPA, 2017)	Pixel a Pixel	DCNN	96,80%
Zhang et al. (ZHANG et al., 2018)	Pixel a Pixel	LSTM	99,00%
Método proposto	CNN + <i>Facial Keypoints</i> + HOG	RNA	98,80%

em cada uma das bases utilizadas nesse trabalho. Nota-se na Figura 46 que as regiões da boca é responsável pela tomada de decisão em relação a classificação da emoção.

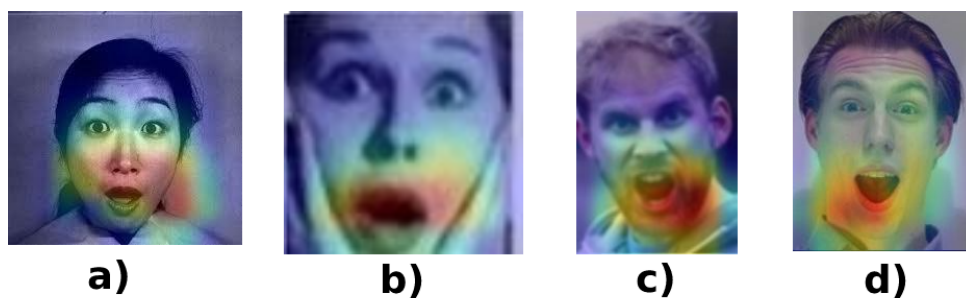


Figura 46 – Representação do Grad-CAM aplicado na emoção surpresa nas bases (a) CK+, (b) FER2013, (c) RAF-DB e (d) MMI.

Uma visão geral dos resultados obtidos com o método proposto em relação a outras abordagens no contexto de classificação de imagens do conjunto de dados CK+ é apresentada na Tabela 7.

## 5.2 Avaliando o SG Face Emotions com os Indivíduos

A segunda etapa da investigação teve como objetivo avaliar o desempenho dos participantes no uso do SG no processo de reconhecimento e aprimoramento das habilidades relacionadas as emoções. Essa etapa foi dividida em dois diferentes cenários realizados com adultos e crianças.

### 5.2.1 Cenário 1: Associação dos Deficientes da Liberdade

Essa avaliação foi empregada em dois grupos de participantes (indivíduos sem TEA e usuários com TEA) nas três fases de investigação para analisar o comportamento dos participantes. Do total de oito indivíduos, quatro possuem diagnóstico de TEA, sendo dois do sexo feminino e dois do sexo masculino e com idades entre 18 a 25 anos e quatro não apresentavam nenhum quadro relacionado ao autismo. Os quatro indivíduos com diagnóstico de TEA foram selecionados na Associação dos Deficientes da Liberdade (APDL), localizada em Uberlândia, Minas Gerais, Brasil. Os dados foram coletados após aprovação do Comitê de Ética em Pesquisa da

Tabela 8 – Dados dos participantes do cenário 1 investigado.

	P1	P2	P3	P4	P5	P6	P7	P8
<b>Idade</b>	18	19	21	25	22	20	23	25
<b>Sexo</b>	M	F	F	M	M	M	F	F
<b>Nível TEA</b>	1	2	3	2	—	—	—	—

UFU-MG (Parecer nº 82555417 0 0000 5152). A Tabela 8 mostra os detalhes a respeito de cada um dos participantes envolvidos nessa avaliação.

A Tabela 9 apresenta as taxas corretas para cada etapa do teste de linha de base múltipla. Os participantes apresentaram desempenho melhores nas fases de intervenção e manutenção em relação ao estágio inicial (linha de base). Observe que todos os participantes mostraram um avanço em relação às pontuações obtidas no estágio de linha de base. Na fase de manutenção, o aumento mais significativo foi observado para o indivíduo P3, que tinha um desempenho de 56,66% e melhorou para 86,66%, destaca-se ainda que esse indivíduo tem o grau mais elevado do TEA. Na fase de manutenção, os participantes mantiveram as habilidades que haviam adquirido na fase de intervenção, e o público obteve taxa de acerto superior a 86%. Na Tabela 9 pode ser observado que para os participantes sem TEA a taxa de acertos são consideradas alta devido as características desses pacientes.

Tabela 9 – Taxa de acerto (%) para os indivíduos investigados nesse cenário.

<b>Participante</b>	<b>Taxa de acerto</b>		
	<b>Referência</b>	<b>Intervenção</b>	<b>Manutenção</b>
P1	56,66	86,66	96,66
P2	30,00	53,33	86,66
P3	40,00	60,00	96,66
P4	53,33	73,33	96,66
P5	100	100	100
P6	100	100	100
P7	96,66	100	100
P8	100	100	100

Na Tabela 10 é exibido o número de vezes que o SG alertou o usuário devido à perda de foco em cada uma das etapas de avaliação (*eye-tracking*) para todos os participantes.

A Tabela 11 traz o erro total para cada emoção demonstrado apenas para os indivíduos com TEA. Essa tabela traz informações apenas dos indivíduos com TEA devido ao desempenho desse público em relação as taxas de acertos. Também é importante destacar que o indivíduo P2, que teve um desempenho mais baixo na fase de manutenção (Tabela 9) resultou em valores maiores em relação ao processo de alerta de atenção como o número de erros em cada emoção (Tabelas 10 e 11). Essas emoções tristeza, medo e desgosto são significativamente menos reconhecidas mesmo em usuários típicos, como relatado anteriormente em outros trabalhos na literatura (MANCINI et al., 2018).

Tabela 10 – Ocorrência do número de alerta que o aplicativo gerou para o usuário devido à falta de atenção.

Usuário	Alerta de atenção		
	Referência	Intervenção	Manutenção
P1	4	3	1
P2	7	4	2
P3	6	3	1
P4	7	3	1
P5	0	0	0
P6	0	0	0
P7	0	0	0
P8	0	0	0

Tabela 11 – Total de erros em cada etapa para os participantes com TEA.

	Fase	Emoções						Soma
		Alegria	Tristeza	Raiva	Surpresa	Desgosto	Medo	
<b>P1</b>	Referência	0	4	2	0	3	4	13
	Intervenção	0	1	1	0	1	1	4
	Manutenção	0	0	0	0	0	1	1
<b>P2</b>	Referência	0	5	5	3	3	5	21
	Intervenção	0	4	1	2	3	4	14
	Manutenção	0	1	2	0	0	1	4
<b>P3</b>	Referência	0	4	3	3	3	5	18
	Intervenção	0	3	2	0	2	5	10
	Manutenção	0	1	0	0	0	0	1
<b>P4</b>	Referência	0	4	2	1	3	4	14
	Intervenção	0	4	0	0	1	3	8
	Manutenção	0	0	0	0	0	1	1
<b>Soma</b>		0	31	18	9	19	34	109

Baseado no desempenho dos participantes desse cenário, a Figura 47 apresenta respectivamente, as avaliações e taxas de acertos para cada uma das sessões dos participantes P1 até P8. Na Figura 47, pode-se observar que o participante P2 elevou sua pontuação em cada etapa da avaliação. Para os participantes P5 a P8, conforme observado na Tabela 9 mostra que as pontuações não sofreram alterações, uma vez que estes usuários não tinham restrições quanto à representação das emoções. O experimento com os indivíduos típicos foi importante para analisar se o SG tinha alguma limitação em particular ou se havia necessidade de mudanças nas animações.

Na fase de referência ilustrada na Figura 47, os especialistas observaram que o foco dos participantes com TEA não estavam direcionados para a parte principal do SG; esses usuários normalmente se concentravam nas outras partes da interface SG. Essa informação é também comprovada pela Tabela 10 com a taxa de alerta de atenção. Os especialistas relataram que alguns participantes foram capazes de expressar o nome correto da emoção, mas não foram

capazes de usar seus músculos faciais para representar a emoção. Além disso, os especialistas observaram que os indivíduos com TEA demonstraram interesse no SG e nas tecnologias através de diálogos estabelecidos durante as sessões, como descrito por (CONSTANTIN et al., 2017). Nas primeiras sessões da fase de base, o tempo definido para os usuários escolherem a emoção foi de três minutos, mas o ajuste foi necessário porque a maioria dos participantes tinha medo de escolher a emoção incorreta ou de exceder o tempo estimado. Em seguida, as instruções foram revisadas novamente pelo especialista e o tempo de escolha foi ajustado para cinco minutos. Os resultados apresentados aqui foram baseados nessa adaptação do tempo. Esta etapa contribuiu para uma maior compreensão da representação das emoções e permitiu o uso da ferramenta. Durante a fase de intervenção, os especialistas guiaram os participantes a se concentrarem nas animações, a fim de manter sua atenção nas expressões faciais e nas regiões faciais utilizadas para representar as emoções. Os especialistas relataram que os usuários estavam interessados no jogo porque o personagem escolhido era capaz de reproduzir suas mímicas faciais. Esta característica e o uso de uma recompensa no final da fase motivou os usuários durante a melhoria da habilidade emocional. Na fase de manutenção, os usuários estavam motivados e comentaram com especialistas que não confundiriam as emoções básicas. No final desta fase, eles comentaram que gostaria de ter acesso ao jogo em plataformas abertas na internet.

Na Figura 48 é apresentado os resultados obtidos durante a intervenção pelos participantes sem TEA P5, P6, P7 e P8 respectivamente. Observando os resultados da Figura 48 é possível observar que todos os participantes com exceção do P7 que errou uma emoção, foram capazes de acertar todas as emoções durante a intervenção.

### 5.2.2 Cenário 2: Clínica de Terapia Ocupacional

Para a segunda etapa de avaliação, uma parceria com uma clínica de terapia ocupacional foi realizado para avaliação com o indivíduo com TEA. Foi selecionado um grupo de oito indivíduos com diagnóstico de TEA para avaliar o desempenho do trabalho proposto em ambiente de clínica. Os indivíduos com diagnóstico de TEA foram selecionados no hospital universitário da Universidade Federal de Uberlândia (UFU), localizado na cidade de Uberlândia, Minas Gerais, Brasil. A coleta de dados foi realizada após aprovação do Comitê de Ética em Pesquisa da UFU (Parecer nº 82555417 0 0000 5152). Essa etapa da investigação teve como objetivo avaliar o desempenho dos indivíduos com TEA por meio do jogo sério. A Tabela 12 mostra os detalhes a respeito de cada um dos participantes envolvidos nessa avaliação.

Tabela 12 – Dados dos participantes do cenário 2 investigado.

	<b>P9</b>	<b>P10</b>	<b>P11</b>	<b>P12</b>	<b>P13</b>	<b>P14</b>	<b>P15</b>	<b>P16</b>
<b>Idade</b>	12	9	6	7	10	6	8	11
<b>Sexo</b>	M	F	F	M	M	M	F	F
<b>Nível TEA</b>	1	2	3	2	1	2	3	1



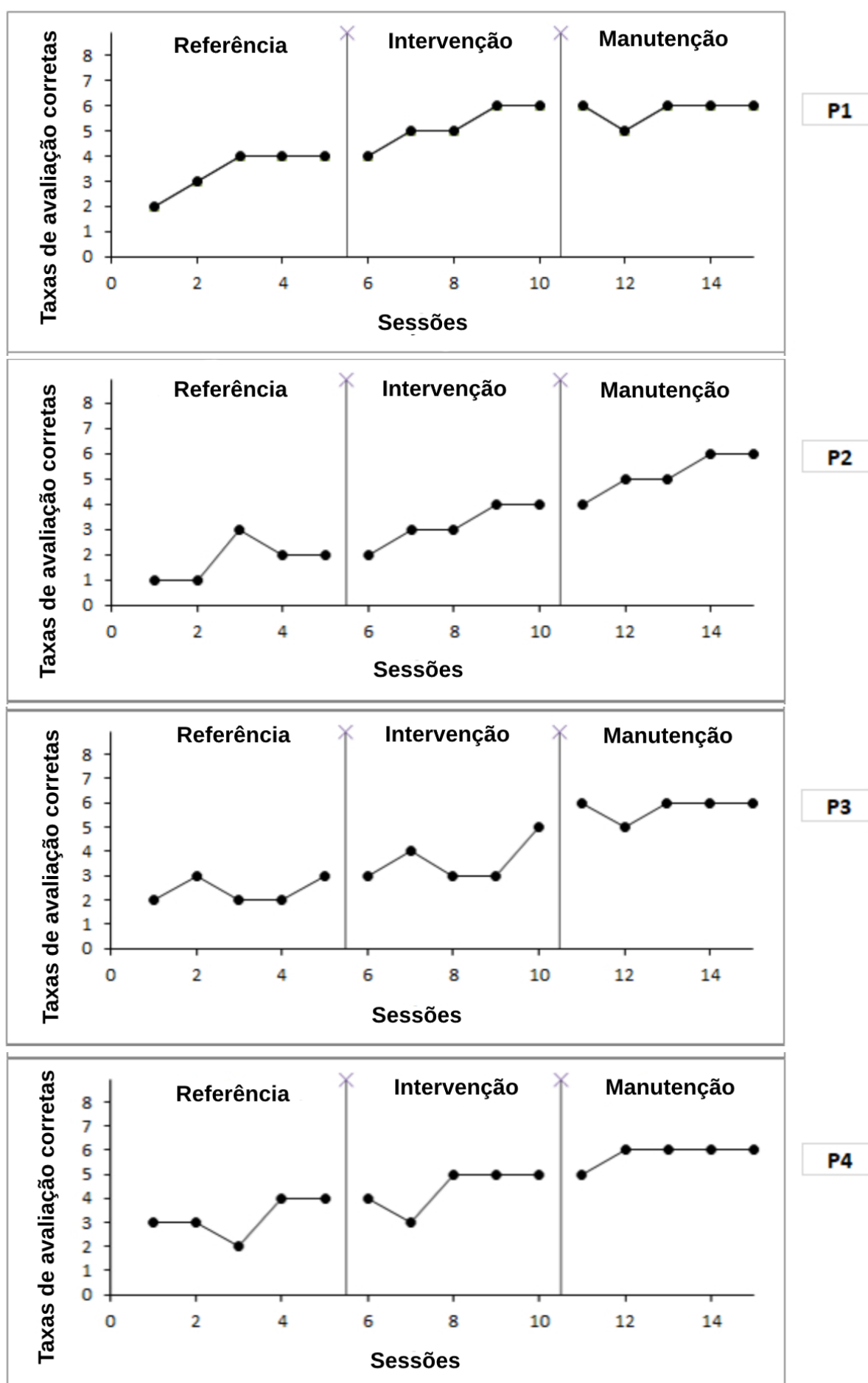


Figura 47 – Taxa de avaliações corretas dos participantes com TEA para as três fases.

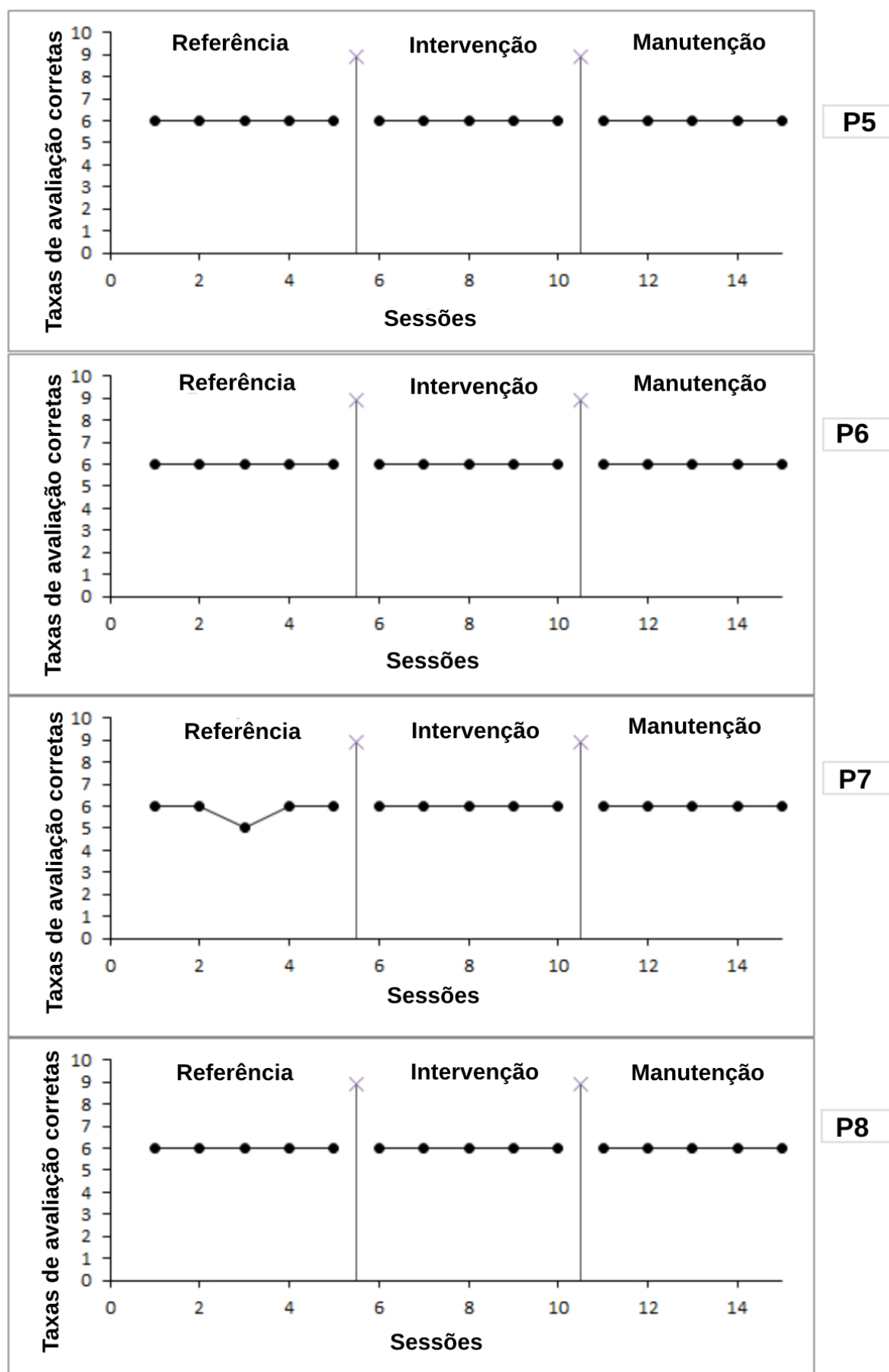


Figura 48 – Taxa de avaliações corretas dos participantes P5, P6, P7 e P8 sem TEA para as três fases.

Foi realizada uma visita a clínica e a ferramenta foi apresentada aos especialistas. Em seguida, os especialistas foram treinados para utilizar a ferramenta e as dúvidas sobre sua usabilidade foram esclarecidas. Uma reunião com os pais também ocorreu para que o jogo fosse apresentado e as dúvidas também foram esclarecidas. Para avaliar as habilidades cognitivas, sociais e de comunicação dos participantes, foram realizadas entrevistas com os pais e os especialistas. Os participantes selecionados com TEA tinham entre 6 e 12 anos, sendo quatro meninos e quatro meninas. O SG foi avaliado com crianças com diferentes níveis de intensidade (do nível 1 - nível 3) do espectro do TEA. Durante a aplicação do método de linha de base (fases de linha de base, intervenção e manutenção), um psicólogo acompanhou os participantes para apoiar o uso da ferramenta. Para esse grupo, a aplicação do instrumento ocorreu de forma individualizada com psicólogos em sessões de terapia. Para a aplicação do SG foi utilizado um computador *desktop* com *webcam*.

A Tabela 13 apresenta as taxas de acerto para cada estágio para o projeto de linha de base múltipla. É possível observar que os participantes apresentaram um avanço nos resultados (pontuação) para a identificação das emoções em relação ao estágio inicial. Esses resultados mostram um avanço com esse segundo cenário de investigação. O aumento mais significativo foi observado para o indivíduo P14.

Tabela 13 – Resultados (%) para os indivíduos com TEA.

Participantes	Taxa de acertos		
	Referência	Intervenção	Manutenção
P9	63,33	80,00	83,33
P10	30,00	70,00	93,33
P11	63,33	90,00	100,0
P12	60,00	73,33	93,33
P13	20,00	36,66	93,33
P14	6,66	36,66	96,66
P15	23,33	46,66	93,33
P16	6,66	46,66	80,00

A Tabela 14 mostra o número de vezes que o SG alertou o usuário devido à perda de foco em cada uma das etapas de avaliação (eye-tracking). Essa tabela mostra que todos os participantes tiveram maiores índices de perda de foco na fase de linha de base e que esses valores diminuíram durante as sessões. Observa-se também que P14 apresentou maiores índices de desatenção, o que pode estar relacionado ao baixo índice de desempenho na fase inicial (referência).

A Tabela 15 traz o erro total para cada emoção demonstrado apenas para os indivíduos com TEA. Na fase referência, os especialistas observaram que os participantes de forma semelhante ao Cenário 1 não apresentavam atenção direcionada a parte principal da interface do jogo (ver Tabela 14). Isso mostra que independente da idade e contexto, o problema inicial foi detectado em ambos os cenários. Normalmente, seu foco era direcionado para outras partes da interface SG.

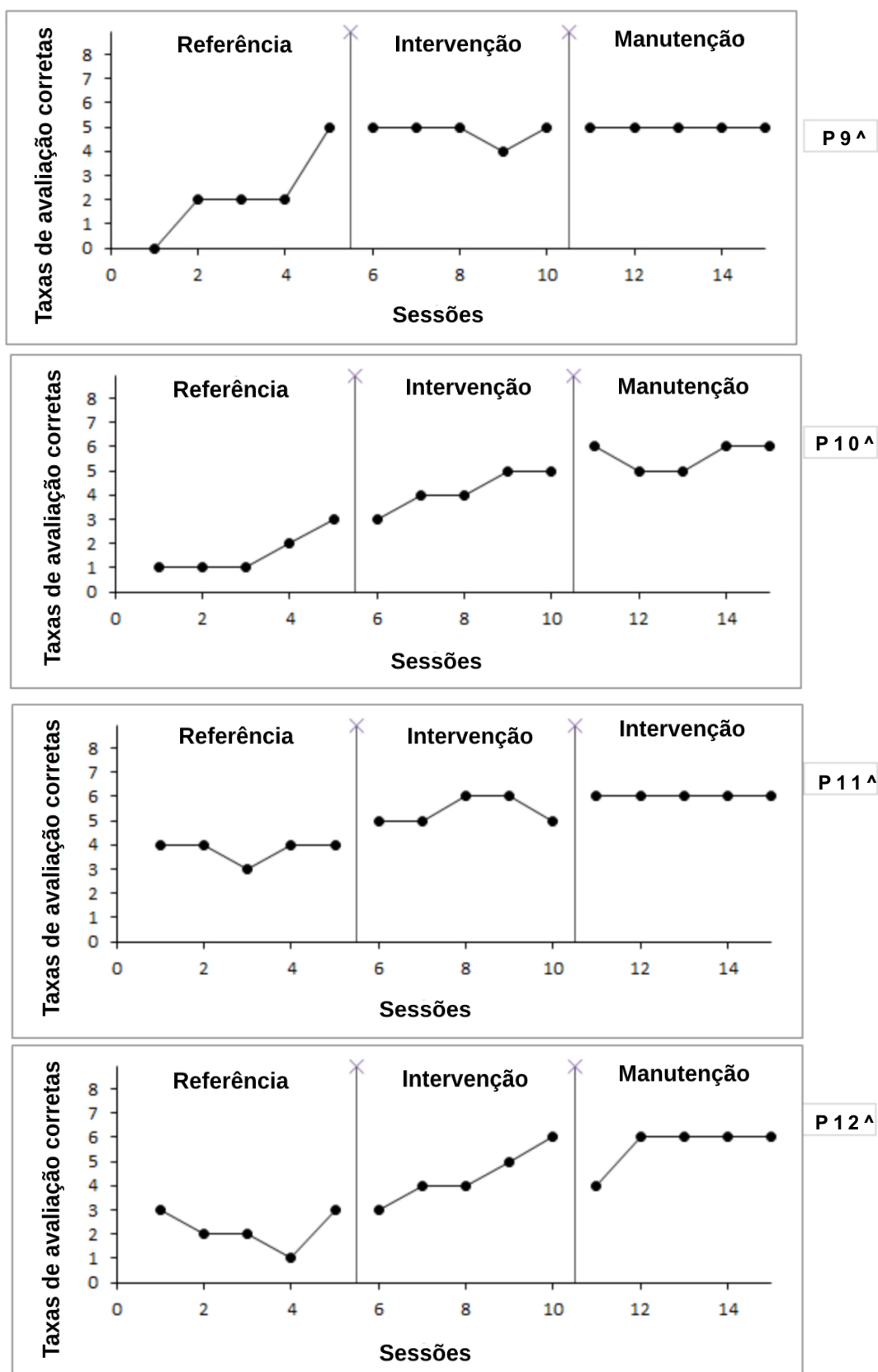


Figura 49 – Taxa de avaliações corretas dos participantes P9, P10, P11 e P12 com TEA para as três fases.

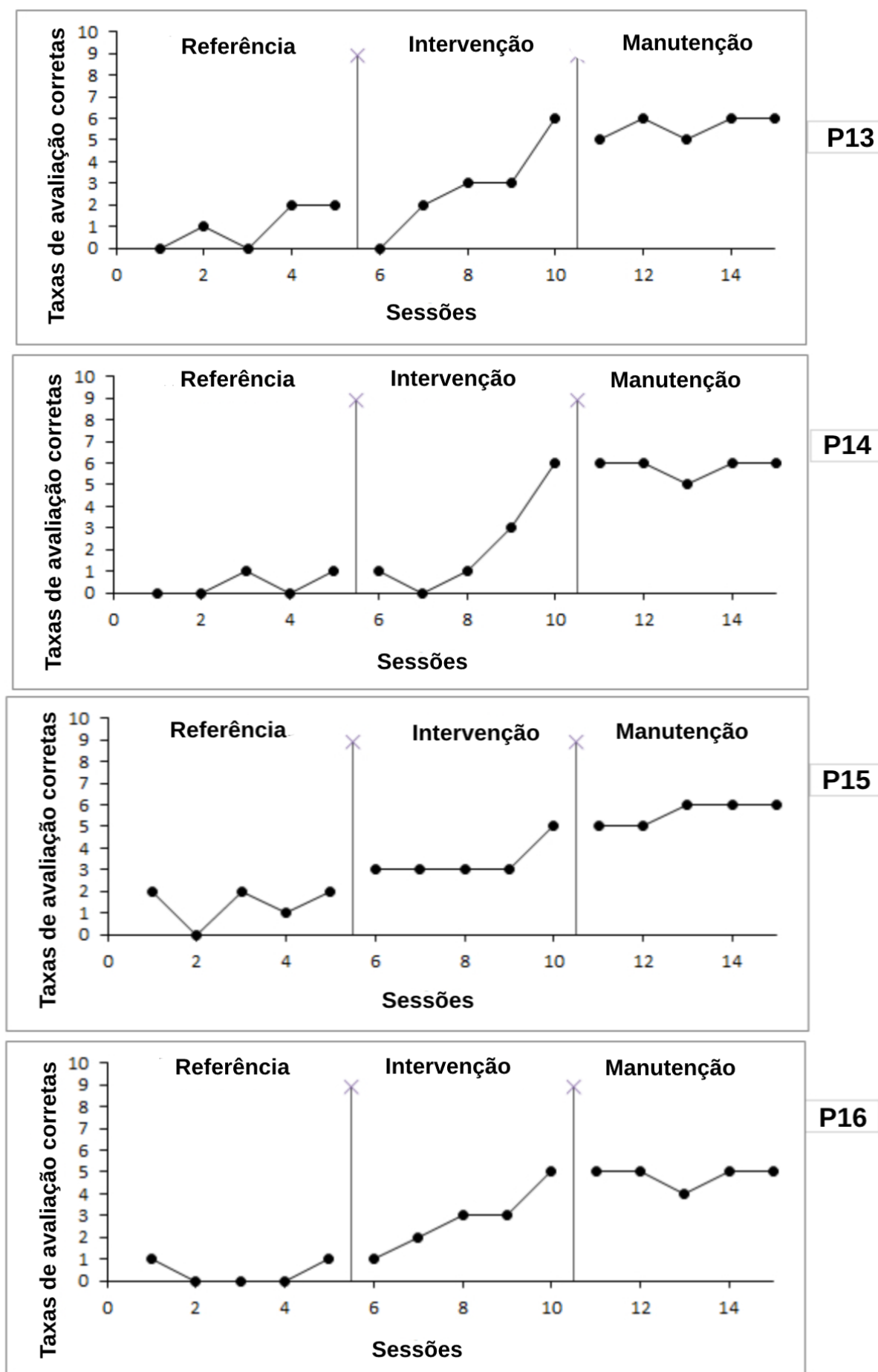


Figura 50 – Taxa de avaliações corretas do participante P13, P14, P15 e P16 com TEA para as três fases.

Tabela 14 – Ocorrência do número de alerta que o aplicativo gerou para o usuário devido à falta de atenção.

Usuário	Alerta de atenção		
	Referência	Intervenção	Manutenção
P9	3	2	1
P10	2	1	0
P11	1	1	0
P12	2	0	0
P13	5	3	0
P14	7	5	1
P15	2	1	1
P16	4	1	0

Tabela 15 – Total de erros em cada etapa para os participantes com TEA.

	Fase	Emoções						Soma
		Alegria	Tristeza	Raiva	Surpresa	Desgosto	Medo	
<b>P9</b>	Referência	0	3	1	0	2	5	11
	Intervenção	0	1	0	0	2	3	6
	Manutenção	0	2	1	0	0	2	5
<b>P10</b>	Referência	2	4	3	3	2	7	21
	Intervenção	0	3	1	0	0	5	9
	Manutenção	0	1	0	0	0	1	2
<b>P11</b>	Referência	1	3	2	1	1	3	11
	Intervenção	0	1	0	0	1	1	3
	Manutenção	0	0	0	0	0	0	0
<b>P12</b>	Referência	1	2	2	2	1	4	12
	Intervenção	0	2	0	1	2	3	8
	Manutenção	0	0	0	0	0	2	2
<b>P13</b>	Referência	1	8	2	2	3	8	24
	Intervenção	1	7	1	2	2	6	19
	Manutenção	0	0	0	0	0	2	2
<b>P14</b>	Referência	3	6	2	3	4	9	27
	Intervenção	1	5	1	3	2	7	19
	Manutenção	0	1	0	0	0	1	2
<b>P15</b>	Referência	0	5	2	2	5	9	23
	Intervenção	0	4	0	0	2	8	14
	Manutenção	0	0	0	0	1	1	2
<b>P16</b>	Referência	2	5	5	5	4	7	28
	Intervenção	0	4	2	2	3	5	16
	Manutenção	0	1	0	0	2	3	6
<b>Soma</b>		12	68	25	26	39	102	272

Na Figura 50 são apresentados os resultados da taxa de avaliação dos acertos para os participantes P9 até P16. As fases da avaliação do projeto de linha de base múltipla com as sessões contribuíram para o aprimoramento das representações e também para o treinamento de usabilidade da ferramenta. Na fase de intervenção (ver Figura 50), os especialistas orientaram os participantes a se concentrarem nas animações enquanto mantinham a atenção nas expressões faciais para representar as emoções. Os participantes P14 e P16 mostraram uma melhora na pontuação de identificação de emoção quando comparados a fase inicial. É possível observar pelos resultados que alguns participantes tiveram uma melhora nas taxas de acerto a cada sessão. Essas informações são observadas principalmente na fase de intervenção. Na fase de manutenção, todos os participantes mostram um aprimoramento das habilidades com boa pontuação em relação às fases anteriores. De forma semelhante ao Cenário 1, observa-se uma evolução dos participantes em relação ao aprimoramento das habilidades relacionadas a detecção e reconhecimento das emoções.





---

## Conclusão

Este capítulo apresenta as conclusões desta pesquisa incluindo suas contribuições ao estado da arte e as limitações do sistema proposto. Com isso, são descritos os trabalhos futuros que objetivam aprimorar o método deste trabalho e avaliá-lo em sua versão atual e aprimorada para usuários com TEA.

### 6.1 Conclusões

Esse trabalho apresenta como contribuição principal a construção de uma ferramenta capaz de auxiliar indivíduos com TEA a desenvolver ou aprimorar as habilidades em detectar e expressar as emoções básicas durante as interações sociais. A utilização de tecnologias com a metodologia de jogos sérios permitiu fornecer aos usuários um processo interativo por meio das técnicas de jogos. Isso possibilitou explorar de maneira mais agradável o ensino das emoções básicas: alegria, tristeza, raiva, medo, desgosto e surpresa.

Na etapa de detecção e reconhecimento das emoções, denominado módulo de detecção e classificação, o sistema proposto apresentou uma aproximação para reconhecimento das emoções em tempo real a partir da associação dos descritores aprendidos no modelo de rede CNN, *Facial keypoints* e HOG. Essa etapa foi avaliada sobre diversas bases de domínio pública sobre as emoções básica de pessoas sem TEA. O ambiente em imagens de domínio pública mostrou que o sistema tem importantes resultados no processo de detecção e reconhecimento das emoções. Ressalta-se que na base Ck+ esses resultados foram melhores aos demais conjuntos. Isso aconteceu por diversas características do processo de aquisição dessas imagens nas diversas bases investigadas. Esse módulo ainda forneceu dados (tempo gasto, erros na representação, foco de atenção, etc) que podem contribuir com o especialista no processo de acompanhamento do indivíduo durante o tratamento.

O jogo sério desenvolvido permitiu ao usuário ter uma interação, pois a ferramenta permite escolhas em relação ao personagens do jogo. Também tem como critério motivador recompensas após cada representação correta da emoção nas fases do jogo durante uma partida. O jogo tem cenários em um contexto do mundo das emoções, onde as habilidades relacionadas a cada

emoção básica pode ser investigada e aprimorada pelo usuário. Os personagens foram modelados levando em consideração as características e recomendação dos psicólogos para o público alvo. O software ainda tem um sistema para gerenciamento dos dados, onde o especialista pode acompanhar a cada sessão os avanços e dificuldades de cada indivíduo. Esses dados pode gerar relatórios em relação ao desempenho desse usuário. Com esses dados é possível ter interação direta sobre o tratamento assim como acompanhar as evoluções a cada sessão.

A ferramenta com todos os módulos foram usados na investigação de dois cenários com características diferentes entre as idades dos indivíduos assim como também o ambiente de aplicação (escola e clínica). Os cenários mostraram que a ferramenta traz uma contribuição no aprimoramento das habilidades das emoções, pois a aplicação do método *multiple baseline design*, que analisa por meio de características e configurações, diferentes períodos do tratamento de indivíduo mostrou resultados promissores na evolução dos usuários durante sua avaliação.

Uma versão parcial da ferramenta foi disponibilizada para aprimoramento das habilidades de detecção e reconhecimento de emoções sem informações de identificação do usuário, apenas capturando os dados da partida. O jogo está disponível para aparelhos celulares da plataforma Android e para computadores em geral no site: <https://michelzinho.pushsistemas.com.br>. Nos tempos atuais de isolamento social devido à Covid-19, essa ferramenta vem sendo utilizada como uma boa alternativa para entretenimento e ao mesmo tempo explora aspectos sociais e emocionais do público geral. Até fevereiro de 2022 já foram realizadas mais de 559 mil partidas usando essa versão.

## 6.2 Trabalhos Futuros

Como limitação o trabalho tem os problemas com o reconhecimento das micro-expressões e questões como o desbalanceamento das bases de imagens com a finalidade de aperfeiçoar e aprimorar os métodos de detecção e classificação das emoções, principalmente, as emoções onde os resultados foram menos expressivos nas avaliações das bases. Em trabalhos futuros torna-se necessário aprimoramento do método para tratar as microexpressões, assim como investigação de algoritmos de agrupamento de dados para auxiliar os especialistas na determinação de padrões de comportamento dos usuários, permitindo que o sistema possa fornecer estratégias de direcionamento do tratamento durante as sessões de intervenção. Um outro ponto a ser explorado são os sinais de eletroencefalograma (EEG) que podem ser investigados com o objetivo de aperfeiçoar a análise possibilitando ter mais informações para uma tomada de decisão do especialista.

Outra característica que pretendemos investigar nos trabalhos futuros é avaliar a interferência das relações pessoais dos indivíduos como, por exemplo, familiaridade com tecnologias digitais, incentivo familiar, etc e como elas podem influenciar de maneira direta ou indireta nos resultados das intervenções com a utilização de jogos sérios.

## 6.3 Produções Bibliográficas Obtidas no Doutorado

Os seguintes artigos foram publicados em conferências locais e internacionais, assim como em periódicos relevantes, como resultados obtidos da pesquisa desenvolvida no curso de doutoramento.

- ❑ DANTAS, A. C.; NASCIMENTO, M. Z. do. Recognition of emotions for people with autism: An approach to improve skills. International Journal of Computer Games Technology, Hindawi, 2022.
- ❑ DANTAS, A. C.; NASCIMENTO, M. Z. do. Face Emotions: Improving Emotional Skills in Individuals with Autism. Aceito em 09/03/2022 na revista Multimedia Tools and Applications.
- ❑ DATAS, A. C; NASCIMENTO, M Z. do. Michelzinho: Jogo sério para o ensino de habilidades emocionais em pessoas com autismo ou deficiência intelectual. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2019. v. 30, n. 1, p. 644.
- ❑ DANTAS, A. C. et al. Recomendação de estratégias pedagógicas através de emoções, perfis de personalidade e inteligências múltiplas utilizando raciocínio baseado em casos. In: Brazilian Symposium on Computers in Education (Simpósio Brasileiro de Informática na Educação-SBIE). 2018. v. 29, n. 1, p. 1213.
- ❑ DANTAS, A. C.; NEVES, Leandro Alves; MARTINS, A. S.; NASCIMENTO, M.Z. Application for Emotion Recognition Through Facial Expressions for the Visually Disability. In: XIV Workshop de Visão Computacional, 2018, Ilhéus. XIV Workshop de Visão Computacional, 2018. p. 45-49.

Esse trabalho também possibilitou a criação de produtos (*softwares*), o qual foi registrado como:

- ❑ DANTAS, A. C.; Z. NASCIMENTO, M. MICHELZINHO - EMOÇÕES E AUTISMO. 2019. Patente: Programa de Computador. Número do registro: 512019000746-8, data de registro: 18/04/2019, título: "MICHELZINHO - EMOÇÕES E AUTISMO", Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial.
- ❑ DANTAS, A. C.; Z. NASCIMENTO, M. EmotionFace - Emoções e Autismo. 2020. Patente: Programa de Computador. Número do registro: 512020002346-0, data de registro: 28/10/2020, título: "EmotionFace - Emoções e Autismo", Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial.

- ❑ DANTAS, A. C.; Z. NASCIMENTO, M. ; SATOKATA, ALESSANDRA ; BORGES, JOÃO; FARIA, PAULO R. ; OLIVEIRA, PEDRO . HISTOLOGIA E BIOLOGIA CELULAR PRÁTICA - HistoBioCelP. 2019. Patente: Programa de Computador. Número do registro: 512019001424-3, data de registro: 05/07/2019, título: "HISTOLOGIA E BIOLOGIA CELULAR PRÁTICA - HistoBioCelP", Instituição de registro: INPI - Instituto Nacional da Propriedade Industrial.

Esse trabalho possibilitou também a publicação de capítulos de livros tais como os listados a seguir:

- ❑ DANTAS, ADILMAR COELHO; Melo, Sara Luzia de ; Prado, Núbia Figueira ; Fernandes, Márcia Aparecida ; Takahashi, Eduardo Koky ; NASCIMENTO, MARCELO ZANCHETTA DO . XQUESTION: UM APLICATIVO DE PERGUNTAS E RESPOSTAS PARA DECISÕES ESTRATÉGICAS DO PROFESSOR DURANTE UMA AULA. Informática Aplicada à Educação 2. 2ed.: Atena Editora, 2019, p. 1-12.
- ❑ DANTAS, ADILMAR COELHO; Melo, Sara Luzia de ; Xavier, Michel Santos ; Guimarães, Guilherme Brilhante ; Santos, Ananda Roberta dos ; Torres, Heidie da Silva ; Gonçalves, Celso André de Souza Barros ; NASCIMENTO, MARCELO ZANCHETTA DO . ALFABETA: UM JOGO COM REALIDADE VIRTUAL PARA AUXILIAR A ALFABETIZAÇÃO E O APRENDIZADO DA GRAFIA CORRETA DE PALAVRAS. Informática Aplicada à Educação 2. 2ed.: Atena Editora, 2019, p. 90-98.

Durante o desenvolvimento deste doutoramento, foram também publicadas algumas pesquisas colaborativas, no campo de pesquisa. Esses artigos estão listados abaixo:

- ❑ DANTAS, ADILMAR C.; XAVIER, MICHEL S. ; GUIMARÃES, GUILHERME B. ; TORRES, HEIDIE S. ; TAKAHASHI, EDUARDO K. ; MELO, SARA L. ; NASCIMENTO, MARCELO Z. DO . Diquinha: Aplicativo móvel para auxiliar no processo de aprendizado não formal em museus. In: Workshops do Congresso Brasileiro de Informática na Educação, 2020, Brasil. Anais dos Workshops do IX Congresso Brasileiro de Informática na Educação, p. 95.
- ❑ DANTAS, ADILMAR COELHO; MELO, SARA DE ; NEVES, LEANDRO ; MILESSI, TAYNARA ; NASCIMENTO, MARCELO ZANCHETTA DO . Michelzinho: Jogo sério para o ensino de habilidades emocionais em pessoas com autismo ou deficiência intelectual. In: XXX Simpósio Brasileiro de Informática na Educação (Brazilian Symposium on Computers in Education), 2019, Brasília. Anais do XXX Simpósio Brasileiro de Informática na Educação. p. 644.
- ❑ PEREIRA JÚNIOR, CLEON ; DANTAS, ADILMAR COELHO ; ABREU, ALINE DE SOUZA LIMA ; REIS, MATHEUS ; MELO, SARA DE ; NASCIMENTO, MARCELO

- ZANCHETTA DO ; DORÇA, FABIANO ; FERNANDES, MARCIA . Personalização das interações de um agente conversacional utilizando emoções e perfis de personalidade. In: XXX Simpósio Brasileiro de Informática na Educação (Brazilian Symposium on Computers in Education), 2019, Brasília. Anais do XXX Simpósio Brasileiro de Informática na Educação, p. 1092.
- ❑ GUIMARÃES, GUILHERME ; DANTAS, ADILMAR COELHO ; XAVIER, MICHEL ; TORRES, HEIDIE ; LIMA, KAYRO ; MELO, SARA DE ; NASCIMENTO, MARCELO ZANCHETTA DO ; KAGIMURA, RICARDO . Aplicativo com Realidade Aumentada para auxiliar no Processo de Ensino e Aprendizagem de Fundamentos de Mecânica Clássica. In: VIII Congresso Brasileiro de Informática na Educação, 2019, Brasília. Anais dos Workshops do VIII Congresso Brasileiro de Informática na Educação, p. 1204.
- ❑ DANTAS, ADILMAR COELHO; TORRES, HEIDIE ; FARIA, ISABELA DE PAULA ; ARAÚJO, WANDERSON ; GUIMARÃES, GUILHERME ; MACHADO, GUILHERME ; ALMEIDA, ADRIANNE VERAS DE ; NASCIMENTO, MARCELO ZANCHETTA DO . AstroBot: Um chatbot com inteligência artificial para auxiliar no processo de ensino e aprendizagem de física. In: VIII Congresso Brasileiro de Informática na Educação, 2019, Brasília. Anais dos Workshops do VIII Congresso Brasileiro de Informática na Educação, p. 1196.
- ❑ DANTAS, ADILMAR COELHO; MELO, SARA DE ; FERNANDES, MARCIA ; LIMA, LUCIANO ; NASCIMENTO, MARCELO ZANCHETTA DO . Recomendação de estratégias pedagógicas através de emoções, perfis de personalidade e inteligências múltiplas utilizando raciocínio baseado em casos. In: XXIX Simpósio Brasileiro de Informática na Educação (Brazilian Symposium on Computers in Education), 2018, Fortaleza, p. 1213.
- ❑ DANTAS, ADILMAR COELHO; MELO, SARA DE ; XAVIER, MICHEL ; GUIMARÃES, GUILHERME ; SANTOS, ANANDA; TORRES, HEIDIE; GONÇALVES, CELSO; NASCIMENTO, MARCELO Z. AlfaBeta: Um jogo com realidade virtual para auxiliar na alfabetização e no aprendizado da grafia correta de palavras. In: VII Congresso Brasileiro de Informática na Educação, 2018, Fortaleza, p. 180.
- ❑ GUIMARÃES, GUILHERME; DANTAS, ADILMAR COELHO; PRATES, EMANUEL PEREIRA, BRENDA; XAVIER, MICHEL; NASCIMENTO, MARCELO Z DO; TAKAHASHI, EDUARDO. Tabela Periódica com Realidade Aumentada Aplicada no Processo de Ensino e Aprendizagem de Química. In: VII Congresso Brasileiro de Informática na Educação, 2018, Fortaleza, 2018. p. 187.
- ❑ C. DANTAS, A.; B. GOUVEIA, E. ; C. MELO, M. ; A. GONÇALVES, C. ; Z. NASCIMENTO, M. . INTERFACE MÓVEL PARA O RECONHECIMENTO DE EMOÇÕES ATRAVÉS DE EXPRESSÕES FACIAIS EM DEFICIENTES VISUAIS. In: XI Simpósio de Engenharia Biomédica, 2018, Uberlândia.

- ❑ DANTAS, ADILMAR COELHO; MELO, SARA DE ; PRADO, NÚBIA ; FERNANDES, MARCIA ; TAKAHASHI, EDUARDO ; NASCIMENTO, MARCELO Z. XQUESTION: Um aplicativo de perguntas e respostas para decisões estratégicas do professor durante uma aula. In: VI Congresso Brasileiro de Informática na Educação, 2017, Recife, p. 160.
- ❑ DANTAS, ADILMAR COELHO; BORGES, JOÃO ; OLIVEIRA, PEDRO ; SATOKATA, ALESSANDRA ; NASCIMENTO, MARCELO ZANCHETTA DO ; FARIA, PAULO R. Desenvolvimento de Aplicativo para ensino de biologia celular e histologia no âmbito universitário através de dispositivos móveis. In: XXIX Simpósio Brasileiro de Informática na Educação (Brazilian Symposium on Computers in Education), 2018, Fortaleza, p. 1704.
- ❑ DANTAS, ADILMAR COELHO; NASCIMENTO, MARCELO ZANCHETTA DO. In: EDUCOM MAIS: Recomendação pelo Ministério da Educação. In: Ministério da Educação <<https://educommais.educacao.rs.gov.br/odas/michelzinhoemocoese-autismo44818>>, 2021.

---

## Referências

- ACEVEDO, D. et al. A simple geometric-based descriptor for facial expression recognition. In: IEEE. **Automatic Face & Gesture Recognition (FG 2017), 2017 12th IEEE International Conference on**. [S.l.], 2017. p. 802–808.
- AJURIAGUERRA, J. d.; MARCELLI, D. Manual de psicopatologia infantil. In: **Manual de psicopatologia infantil**. [S.l.: s.n.], 1986.
- ALMEIDA, L. M. et al. Altriras: A computer game for training children with autism spectrum disorder in the recognition of basic emotions. **International Journal of Computer Games Technology**, Hindawi, v. 2019, 2019.
- APPLE, A. L. et al. Effects of video modeling alone and with self-management on compliment-giving behaviors of children with high-functioning asd. **Journal of Positive Behavior Interventions**, Sage Publications Sage CA: Los Angeles, CA, v. 7, n. 1, p. 33–46, 2005.
- ASSOCIATION, A. P. et al. **DSM-5: Manual diagnóstico e estatístico de transtornos mentais**. [S.l.]: Artmed Editora, 2014.
- ASTINGTON, J. W.; GOPNIK, A. Knowing you've changed your mind: Children's understanding of representational change. **Developing theories of mind**, p. 193–206, 1988.
- AZCARATE et al. **Automatic facial emotion recognition**. 2005.
- AZCARATE, A. et al. Automatic facial emotion recognition. **Universiteit van Amsterdam**, Amsterdam, NL, 2005. Disponível em: <<http://staff.science.uva.nl/~rvalenti/projects/mmis/Automatic%20Facial%20Emotion%20Recognition.pdf>>.
- BAPTISTA, C. R.; BOSA, C. **Autismo e educação-reflexões e propostas de intervenção**. [S.l.]: Artmed Editora, 2002.
- BARON-COHEN, S. The development of a theory of mind in autism: deviance and delay? **Psychiatric Clinics**, Elsevier, v. 14, n. 1, p. 33–51, 1991.
- BARON-COHEN, S. Theory of mind and autism: A review. **International review of research in mental retardation**, v. 23, n. 23, p. 169–184, 2001.
- BARON-COHEN, S. et al. Does the autistic child have a “theory of mind”. **Cognition**, v. 21, n. 1, p. 37–46, 1985.

- BARON-COHEN, S. et al. Empathizing and systemizing in autism spectrum conditions. **Handbook of autism and pervasive developmental disorders**, John Wiley & Sons Inc. New Jersey, v. 1, p. 628–639, 2005.
- BASHA, S. S. et al. Impact of fully connected layers on performance of convolutional neural networks for image classification. **Neurocomputing**, Elsevier, v. 378, p. 112–119, 2020.
- BASTANFARD, A. et al. A novel multimedia educational speech therapy system for hearing impaired children. In: SPRINGER. **Pacific-Rim Conference on Multimedia**. [S.l.], 2010. p. 705–715.
- BASTANFARD, A.; TAKAHASHI, H.; NAKAJIMA, M. Toward e-appearance of human face and hair by age, expression and rejuvenation. In: IEEE. **2004 International conference on Cyberworlds**. [S.l.], 2004. p. 306–311.
- BEGEER, S. et al. Emotional competence in children with autism: Diagnostic criteria and empirical evidence. **Developmental Review**, Elsevier, v. 28, n. 3, p. 342–369, 2008.
- BENGIO, Y. et al. Learning deep architectures for ai. **Foundations and trends® in Machine Learning**, Now Publishers, Inc., v. 2, n. 1, p. 1–127, 2009.
- BENKHERRAT, M. Digital exchange communication system for children and youths with autism. In: SPRINGER. **Proceedings of SAI Intelligent Systems Conference**. [S.l.], 2018. p. 487–495.
- BILKHU, M. S.; GUPTA, S.; SRIVASTAVA, V. K. Emotion classification from facial expressions using cascaded regression trees and svm. In: **Computational Intelligence: Theories, Applications and Future Directions-Volume II**. [S.l.]: Springer, 2019. p. 585–594.
- BLACK, D. W.; GRANT, J. E. **DSM-5 TM guidebook the essential companion to the Diagnostic and statistical manual of mental disorders, fifth edition**. American Psychiatric Publishing, 2014. Disponível em: <<http://dx.doi.org/10.1176/appi.books.9780890425596>>.
- BLENDER. **Blender About@ONLINE**. 2018. Disponível em: <<https://www.blender.org/about/>>.
- BOLTE, S. et al. The development and evaluation of a computer-based program to test and to teach the recognition of facial affect. Taylor & Francis, n. sup2, p. 61–68, 2002.
- BOSA, C. Autismo: atuais interpretações para antigas observações. **Autismo e educação: reflexões e propostas de intervenção**, Artmed Porto Alegre, p. 21–39, 2002.
- BOUCENNA, S. et al. Interactive technologies for autistic children: A review. **Cognitive Computation**, Springer, v. 6, n. 4, p. 722–740, 2014.
- BRUNELLI, R.; POGGIO, T. Face recognition: Features versus templates. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 15, n. 10, p. 1042–1052, 1993.
- BRYANT, F. B.; YARNOLD, P. R. Principal-components analysis and exploratory and confirmatory factor analysis. American Psychological Association, 1995.
- CANAL, F. Z. et al. A survey on facial emotion recognition techniques: A state-of-the-art literature review. **Information Sciences**, Elsevier, v. 582, p. 593–617, 2022.



CARAMIHALE, T.; POPESCU, D.; ICHIM, L. Emotion classification using a tensorflow generative adversarial network implementation. **Symmetry**, Multidisciplinary Digital Publishing Institute, v. 10, n. 9, p. 414, 2018.

CHANG, C.-Y.; HUANG, Y.-C. A subject-dependent facial expression recognition system. In: **Emotion Recognition: A Pattern Analysis Approach**. [S.l.: s.n.], 2014. p. 89–112.

CHEN, C.-H.; LEE, I.-J.; LIN, L.-Y. Augmented reality-based self-facial modeling to promote the emotional expression and social skills of adolescents with autism spectrum disorders. **Research in developmental disabilities**, Elsevier, v. 36, p. 396–403, 2015.

CHEN, C.-H. C.; LEE, I.-J.; LIN, L.-Y. Augmented reality-based video-modeling storybook of nonverbal facial cues for children with autism spectrum disorder to improve their perceptions and judgments of facial expressions and emotions. **Computers in Human Behavior**, Elsevier, v. 55, p. 477–485, 2016.

CHEN, J. et al. A pilot study on evaluating children with autism spectrum disorder using computer games. **Computers in Human Behavior**, Elsevier, v. 90, p. 204–214, 2019.

COHN, J. F. et al. Feature-point tracking by optical flow discriminates subtle differences in facial expression. In: IEEE. **Automatic Face and Gesture Recognition, 1998. Proceedings. Third IEEE International Conference On**. [S.l.], 1998. p. 396–401.

COLEMAN, T. E.; MONEY, A. G. Student-centred digital game-based learning: a conceptual framework and survey of the state of the art. **Higher Education**, Springer, v. 79, n. 3, p. 415–457, 2020.

CONSTANTIN, A. et al. Designing computer-based rewards with and for children with autism spectrum disorder and/or intellectual disability. **Computers in Human Behavior**, Elsevier, v. 75, p. 404–414, 2017.

CONTROL, C. D.; PREVENTION. **Autism Spectrum Disorder (ASD)@ONLINE**. 2018. Disponível em: <<https://www.cdc.gov/ncbddd/autism/data.html>>.

DACHAPALLY, P. R. Facial emotion detection using convolutional neural networks and representational autoencoder units. **arXiv preprint arXiv:1706.01509**, 2017.

DAPOGNY, A. et al. Jemime: a serious game to teach children with asd how to adequately produce facial expressions. In: IEEE. **2018 13th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2018)**. [S.l.], 2018. p. 723–730.

DAPRETTO, M. et al. Understanding emotions in others: mirror neuron dysfunction in children with autism spectrum disorders. **Nature neuroscience**, Nature Publishing Group, v. 9, n. 1, p. 28, 2006.

DARWIN, C. **The expression of the emotions in man and animals**. [S.l.]: University of Chicago press, 1965. v. 526.

DARWIN, C.; GARCIA, L. d. S. L. **A expressão das emoções no homem e nos animais**. [S.l.]: Companhia das Letras, 2000.

DENG, L.; YU, D. et al. Deep learning: methods and applications. **Foundations and Trends® in Signal Processing**, Now Publishers, Inc., v. 7, n. 3–4, p. 197–387, 2014.

- DIDEHBANI, N. et al. Virtual reality social cognition training for children with high functioning autism. **Computers in Human Behavior**, Elsevier, v. 62, p. 703–711, 2016.
- DING, H.; ZHOU, S. K.; CHELLAPPA, R. Facenet2expnet: Regularizing a deep face recognition net for expression recognition. In: IEEE. **2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)**. [S.l.], 2017. p. 118–126.
- EKMAN, P.; FRIESEN, W. Facial action coding system. **Alto: Consulting Psychologist Press**, Palo, 1978.
- EKMAN, P.; FRIESEN, W. V.; HAGER, J. C. The facial action coding system. **Research Nexus eBook**, Salt Lake City, UT, 2002.
- EKMAN, P.; ROSENBERG, E. **What the Face Reveals: Basic and Applied Studies of Spontaneous Expression Using the Facial Action Coding System (FACS)**. New York, USA: Oxford – University Press, 2005. 12 - 17 p.
- ELSHAHAWEY, M.; ABOELNAGA, K.; SHARAF, N. Codaroutine: A serious game for introducing sequential programming concepts to children with autism. In: IEEE. **2020 IEEE Global Engineering Education Conference (EDUCON)**. [S.l.], 2020. p. 1862–1867.
- FLEET, D.; WEISS, Y. Optical flow estimation. In: **Handbook of mathematical models in computer vision**. [S.l.]: Springer, 2006. p. 237–257.
- FREUND, Y.; SCHAPIRE, R. E. A decision-theoretic generalization of on-line learning and an application to boosting. **Journal of computer and system sciences**, Elsevier, v. 55, n. 1, p. 119–139, 1997.
- FRIDENSON-HAYO, S. et al. ‘emotiplay’: a serious game for learning about emotions in children with autism: results of a cross-cultural evaluation. **European child & adolescent psychiatry**, Springer, v. 26, n. 8, p. 979–992, 2017.
- FRIESEN, W. V.; EKMAN, P. Emfacs-7. **Manual não publicado**, 1984.
- GARCIA, L. S. L. A expressão das emoções no homem e nos animais. In: **Traduzido por: Leon de Souza Lobo Garcia – A Origem das Espécies, publicado por Charles Darwin – 1859**. São Paulo, Brasil: Companhia de Bolso, 2009. p. 33–55. ISSN 978-85-3591-398-9. [Http://www.letras.ufmg.br/arquivos/matte/bib/darwin.pdf](http://www.letras.ufmg.br/arquivos/matte/bib/darwin.pdf).
- GARRIS, R.; AHLERS, R.; DRISKELL, J. E. Games, motivation, and learning: A research and practice model. **Simulation & gaming**, Sage Publications Sage CA: Thousand Oaks, CA, v. 33, n. 4, p. 441–467, 2002.
- GHIMIRE, D. et al. Recognition of facial expressions based on tracking and selection of discriminative geometric features. **Int. J. Multimed. Ubiquitous Eng**, v. 10, n. 3, p. 35–44, 2015.
- GILDENBLAT, J.; CONTRIBUTORS. **PyTorch library for CAM methods**. [S.l.]: GitHub, 2021. <<https://github.com/jacobgil/pytorch-grad-cam>>.
- GONZALEZ, R.; WOODS, R. **Digital Image Processing**. Pearson, 2018. ISBN 9780133356724. Disponível em: <<https://books.google.com.br/books?id=0F05vgAACAAJ>>.
- GOODFELLOW, I. J. et al. Challenges in representation learning: A report on three machine learning contests. In: SPRINGER. **International conference on neural information processing**. [S.l.], 2013. p. 117–124.

GOOGLE. **Uma plataforma completa de código aberto para machine learning**. [S.l.]: Web, 2021. <<https://www.tensorflow.org>>.

GRACIOSO, A. C. N. da R. et al. Emotion recognition system using open web platform. In: IEEE. **Security Technology (ICCST), 2013 47th International Carnahan Conference on**. [S.l.], 2013. p. 1–5.

GREENSPAN, S. I.; WIEDER, S. **Engaging autism: Using the floortime approach to help children relate, communicate, and think**. [S.l.]: Da Capo Lifelong Books, 2006.

GROSSARD, C. et al. Serious games to teach social interactions and emotions to individuals with autism spectrum disorders (asd). **Computers & Education**, Elsevier, v. 113, p. 195–211, 2017.

GROSSARD, C. et al. Teaching facial expression production in autism: The serious game je-mime. **Creative Education**, Scientific Research Publishing, v. 10, n. 11, p. 2347, 2019.

GROSSMAN, R. B.; EDELSON, L. R.; TAGER-FLUSBERG, H. Emotional facial and vocal expressions during story retelling by children and adolescents with high-functioning autism. **Journal of Speech, Language, and Hearing Research**, ASHA, 2013.

HEWITT, S. Compreender o autismo-estratégias para alunos com autismo nas escolas regulares. **Porto: Porto Editora**, p. 8, 2006.

HINTON, G. E.; OSINDERO, S.; TEH, Y.-W. A fast learning algorithm for deep belief nets. **Neural computation**, MIT Press, v. 18, n. 7, p. 1527–1554, 2006.

HOWARD, A. G. et al. Mobilenets: Efficient convolutional neural networks for mobile vision applications. **arXiv preprint arXiv:1704.04861**, 2017.

HSU, R.-L.; ABDEL-MOTTALEB, M.; JAIN, A. K. Face detection in color images. **IEEE transactions on pattern analysis and machine intelligence**, IEEE, v. 24, n. 5, p. 696–706, 2002.

HWANG, G.-J. et al. Development of a personalized educational computer game based on students' learning styles. **Educational Technology Research and Development**, Springer, v. 60, n. 4, p. 623–638, 2012.

INGERSOLL, B. et al. The effects of a developmental, social—pragmatic language intervention on rate of expressive language production in young children with autistic spectrum disorders. **Focus on Autism and Other Developmental Disabilities**, Sage Publications Sage CA: Los Angeles, CA, v. 20, n. 4, p. 213–222, 2005.

JAIN, D. K.; SHAMSOLMOALI, P.; SEHDEV, P. Extended deep neural network for facial emotion recognition. **Pattern Recognition Letters**, Elsevier, v. 120, p. 69–74, 2019.

JOU, G. I. d.; SPERB, T. M. Teoria da mente: diferentes abordagens. **Psicologia: reflexão e crítica**, SciELO Brasil, v. 12, n. 2, p. 287–306, 1999.

KAMP-BECKER, I. et al. Categorical and dimensional structure of autism spectrum disorders: The nosologic validity of asperger syndrome. **Journal of autism and developmental disorders**, Springer, v. 40, n. 8, p. 921–929, 2010.

KANADE, T. C.; COHN, J. Au-coded facial expression database. 2005.

- KASS, M.; WITKIN, A.; TERZOPOULOS, D. Snakes: Active contour models. **International journal of computer vision**, Springer, v. 1, n. 4, p. 321–331, 1988.
- KAZEMI, V.; SULLIVAN, J. One millisecond face alignment with an ensemble of regression trees. In: **Proceedings of the IEEE conference on computer vision and pattern recognition**. [S.l.: s.n.], 2014. p. 1867–1874.
- KETKAR, N. et al. **Deep Learning with Python**. [S.l.]: Springer, 2017. v. 1.
- KIM, Y.; LEE, H.; PROVOST, E. M. Deep learning for robust feature generation in audiovisual emotion recognition. In: IEEE. **2013 IEEE international conference on acoustics, speech and signal processing**. [S.l.], 2013. p. 3687–3691.
- KLIN, A. Autismo e síndrome de asperger: uma visão geral autism and asperger syndrome: an overview. **Rev Bras Psiquiatr**, SciELO Brasil, v. 28, n. Supl I, p. S3–11, 2006.
- KONAR, A.; CHAKRABORTY, A. **Emotion Recognition: A Pattern Analysis Approach**. Wiley, 2014. ISBN 9781118910610. Disponível em: <<http://books.google.com.br/books?id=-JLVBQAAQBAJ>>.
- KOSUDA, S. et al. Decision tree sensitivity analysis for cost-effectiveness of chest fdg-pet in patients with a pulmonary tumor (non-small cell carcinoma). **Kaku igaku. The Japanese journal of nuclear medicine**, v. 35, n. 6, p. 395–404, 1998.
- LI, K.-H. et al. The effects of applying game-based learning to webcam motion sensor games for autistic students' sensory integration training. **Turkish Online Journal of Educational Technology-TOJET**, ERIC, v. 11, n. 4, p. 451–459, 2012.
- LI, S.; DENG, W. Reliable crowdsourcing and deep locality-preserving learning for unconstrained facial expression recognition. **IEEE Transactions on Image Processing**, IEEE, v. 28, n. 1, p. 356–370, 2019.
- LIENHART, R.; MAYDT, J. An extended set of haar-like features for rapid object detection. In: IEEE. **Image Processing. 2002. Proceedings. 2002 International Conference on**. [S.l.], 2002. v. 1, p. I–I.
- LIMA, A. M. O. d. et al. Analysis of softwares for emotion recognition in children and teenagers with autism spectrum disorder. **Revista CEFAC**, SciELO Brasil, v. 21, n. 1, 2019.
- MAENNER, M. J. et al. Prevalence of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, united states, 2016. **MMWR Surveillance Summaries**, Centers for Disease Control and Prevention, v. 69, n. 4, p. 1, 2020.
- MAKEHUMAN. **What is MakeHuman?** [S.l.]: Web, 2021. <[http://www.makehumancommunity.org/wiki/Documentation:What\\_is\\_MakeHuman%3F](http://www.makehumancommunity.org/wiki/Documentation:What_is_MakeHuman%3F)>.
- MALINVERNI, L. et al. An inclusive design approach for developing video games for children with autism spectrum disorder. **Computers in Human Behavior**, Elsevier, v. 71, p. 535–549, 2017.
- MANCINI, G. et al. Recognition of facial emotional expressions among italian pre-adolescents, and their affective reactions. **Frontiers in psychology**, Frontiers, v. 9, p. 1303, 2018.

- MANTA, O. et al. A three-module proposed solution to improve cognitive and social skills of students with attention deficit disorder (add) and high functioning autism (hfa) innovative technological advancements for students with neurodevelopmental disorders. In: **Proceedings of the 13th ACM International Conference on Pervasive Technologies Related to Assistive Environments**. [S.l.: s.n.], 2020. p. 1–7.
- MARSCHNER, S.; SHIRLEY, P. **Fundamentals of computer graphics**. [S.l.]: CRC Press, 2015.
- MAZUREK, M. O.; ENGELHARDT, C. R.; CLARK, K. E. Video games from the perspective of adults with autism spectrum disorder. **Computers in Human Behavior**, Elsevier, v. 51, p. 122–130, 2015.
- METRI, P.; GHORPADE, J.; BUTALIA, A. Facial emotion recognition using context based multimodal approach. **International Journal of Emerging Sciences**, Springfield Publishing Corporation, v. 2, n. 1, p. 171–183, 2012.
- MOHAMMED, G. M.; MELHUM, A. I. Implementation of hog feature extraction with tuned parameters for human face detection. **International Journal of Machine Learning and Computing** 10.5, 2020.
- MONSALVE, E.; WERNECK, V.; LEITE, J. Evolución de un juego educacional de ingeniería de software a través de técnicas de elicitación de requisitos. In: **Proceedings of XIII Workshop on Requirements Engineering (WER'2010)**, Cuenca, Ecuador. [S.l.: s.n.], 2010. p. 12–23.
- MOTA, J. **Reconhecimento emocional em crianças com perturbação do espectro do autismo: Um estudo exploratório com o videojogo “Lets face It!”**. Dissertação (Mestrado) — Tese de Mestrado não publicada). Universidade do Porto, Faculdade de Psicologia e de Ciências da Educação, Porto, Portugal, 2012.
- MOZILLA. **HTML5@ONLINE**. 2018. Disponível em: <<https://developer.mozilla.org/en-US/docs/Glossary/HTML5>>.
- MOZILLA. **JavaScript Tutoriais**. [S.l.]: Web, 2021. <<https://developer.mozilla.org/pt-BR/docs/Web/JavaScript>>.
- MUNASINGHE, M. Facial expression recognition using facial landmarks and random forest classifier. In: IEEE. **2018 IEEE/ACIS 17th International Conference on Computer and Information Science (ICIS)**. [S.l.], 2018. p. 423–427.
- MUÑOZ, P. d. O. L. **Rastreamento de olhar e reconhecimento de emoções em crianças com transtorno do espectro autístico**. Tese (Doutorado) — Universidade de São Paulo, 2018.
- MYSQL. **Why MySQL@ONLINE**. 2018. Disponível em: <<https://www.mysql.com/why-mysql/>>.
- NASS, C.; BRAVE, S. Emotion in human-computer interaction. In: **The human-computer interaction handbook**. [S.l.]: CRC Press, 2007. p. 94–109.
- NG, H.-W. et al. Deep learning for emotion recognition on small datasets using transfer learning. In: **Proceedings of the 2015 ACM on international conference on multimodal interaction**. [S.l.: s.n.], 2015. p. 443–449.
- NGUYEN, H. **Gpu gems 3**. [S.l.]: Addison-Wesley Professional, 2007.

- NIELSEN, M. A. **Neural networks and deep learning**. [S.l.]: Determination press San Francisco, CA, USA:, 2015. v. 2018.
- NOOR, J. et al. Facial expression recognition using hand-crafted features and supervised feature encoding. In: IEEE. **2020 International Conference on Electrical, Communication, and Computer Engineering (ICECCE)**. [S.l.], 2020. p. 1–5.
- NOVAK, J. **Game development essentials: an introduction**. [S.l.]: Cengage Learning, 2011.
- OLANOW, C. W.; WATTS, R. L.; KOLLER, W. C. An algorithm (decision tree) for the management of parkinson's disease (2001):: Treatment guidelines. **Neurology**, AAN Enterprises, v. 56, n. suppl 5, p. S1–S88, 2001.
- OMS. **Autism Spectrum Disorders, Key Facts@ONLINE**. 2021. Disponível em: <<https://www.who.int/news-room/fact-sheets/detail/autism-spectrum-disorders>>.
- PADGETT, C.; COTTRELL, G. W. Representing face images for emotion classification. In: **Advances in neural information processing systems**. [S.l.: s.n.], 1997. p. 894–900.
- PANTIC, M. et al. Web-based database for facial expression analysis. In: IEEE. **Multimedia and Expo, 2005. ICME 2005. IEEE International Conference on**. [S.l.], 2005. p. 5–pp.
- PAVEZ, R. et al. Emo-mirror: a proposal to support emotion recognition in children with autism spectrum disorders. **Neural Computing and Applications**, Springer, p. 1–12, 2021.
- PENNINGTON, R. C.; CARPENTER, M. Teaching written expression to students with autism spectrum disorder and complex communication needs. **Topics in Language Disorders**, LWW, v. 39, n. 2, p. 191–207, 2019.
- PERGHER, G. K. et al. Memory, mood and emotion. **Revista de psiquiatria do Rio Grande do Sul**, SciELO Brasil, v. 28, n. 1, p. 61–68, 2006.
- PLIMLEY, L. A. A review of quality of life issues and people with autism spectrum disorders. **British Journal of Learning Disabilities**, Wiley Online Library, v. 35, n. 4, p. 205–213, 2007.
- PREMACK, D.; WOODRUFF, G. Does the chimpanzee have a theory of mind? **Behavioral and brain sciences**, Cambridge University Press, v. 1, n. 4, p. 515–526, 1978.
- PRENSKY, M. Digital game-based learning. **Computers in Entertainment (CIE)**, ACM, v. 1, n. 1, p. 21–21, 2003.
- PYTHON. **Python documentation**. [S.l.]: Web, 2021. <<https://docs.python.org/3/>>.
- QIN, Y. et al. Polygenic risk for autism spectrum disorder affects left amygdala activity and negative emotion in schizophrenia. **Translational psychiatry**, Nature Publishing Group, v. 10, n. 1, p. 1–12, 2020.
- RIBU, K. Teaching computer science to students with asperger's syndrome. **OSLOMET**, Tapir Akademisk Forlag, 2010.
- ROSEBROCK, A. **Deep Learning for Computer Vision with Python: ImageNet Bundle**. [S.l.]: PyImageSearch, 2017.

- RUSSELL, S.; NORVIG, P. **Inteligência artificial**. CAMPUS - RJ, 2004. ISBN 9788535211771. Disponível em: <<https://books.google.com.br/books?id=wBMvAAAACAAJ>>.
- SADIK, R.; ANWAR, S.; REZA, M. L. Autismnet: Recognition of autism spectrum disorder from facial expressions using mobilenet architecture. **International Journal**, v. 10, n. 1, 2021.
- SAGAYARAJ, K.; GOPAL, C. R.; KARTHIKEYAN, S. The efficacy of technology and non-technology based intervention for children with autism spectrum disorder: A meta-analysis. **International journal of Innovative Science and Research Technology**, v. 5, n. 3, p. 863–868, 2020.
- SAHA, C. et al. Facial expressions: A cross-cultural study. **Emotion Recognition: A Pattern Analysis Approach**, Wiley Online Library, p. 69–87, 2015.
- SAHYOUN, F. **Facial Action Coding System@ONLINE**. 2016. Disponível em: <<http://www.fadysahyoun.com/2015/11/27/facial-action-coding-system>>.
- SALMAM, F. Z.; MADANI, A.; KISSI, M. Facial expression recognition using decision trees. In: IEEE. **2016 13th International Conference on Computer Graphics, Imaging and Visualization (CGiV)**. [S.l.], 2016. p. 125–130.
- SÁNCHEZ, A. et al. Differential optical flow applied to automatic facial expression recognition. **Neurocomputing**, Elsevier, v. 74, n. 8, p. 1272–1282, 2011.
- SANDIN, S. et al. The familial risk of autism. **Jama**, American Medical Association, v. 311, n. 17, p. 1770–1777, 2014.
- SCASSELLATI, B.; ADMONI, H.; MATARIĆ, M. Robots for use in autism research. **Annual review of biomedical engineering**, v. 14, 2012.
- SEGUIN, D. et al. Amygdala subnuclei development in adolescents with autism spectrum disorder: Association with social communication and repetitive behaviors. **Brain and Behavior**, Wiley Online Library, 2021.
- SHAN, C. C.; GONG, S.; MCOWAN, P. W. Facial expression recognition based on local binary patterns: A comprehensive study. **Image and vision Computing**, Elsevier, v. 27, n. 6, p. 803–816, 2009.
- SILVA, M. Z. d. L. da; ARTUSO, A. R.; TORTATO, C. S. B. Tecnologias de inclusão no ensino de crianças com tea. **Revista Eletrônica Pesquiseduca**, v. 12, n. 26, p. 157–179, 2020.
- SILVER, M.; OAKES, P. Evaluation of a new computer intervention to teach people with autism or asperger syndrome to recognize and predict emotions in others. **Autism**, Sage Publications Sage CA: Thousand Oaks, CA, v. 5, n. 3, p. 299–316, 2001.
- STRICKROTH, S. et al. Game-based promotion of motivation and attention for socio-emotional training in autism: Exploring the secrets of facial expressions by combining minecraft and a mobile app. **i-com**, De Gruyter, v. 19, n. 1, p. 17–30, 2020.
- STURM, D.; PEPPE, E.; PLOOG, B. emot-ican: Design of an assessment game for emotion recognition in players with autism. In: IEEE. **Serious games and applications for health (segah), 2016 ieee international conference on**. [S.l.], 2016. p. 1–7.

- TANG, J. S. et al. Development and feasibility of mindchip™: A social emotional telehealth intervention for autistic adults. **Journal of autism and developmental disorders**, Springer, v. 51, n. 4, p. 1107–1130, 2021.
- TANG, S.; HANNEGHAN, M.; RHALIBI, A. E. Introduction to games-based learning. **Games Based Learning Advancements for Multi-Sensory Human Computer Interfaces**. New York: IGI Global, 2009.
- THIBES-UEM, P. A.; ALENCAR-UEM, G. A. R. D.; AZEVEDO-UEM, F. C. D. Estratégia de intervenção: Ensinando emoções. 2009.
- THOMSON, P. et al. A computerised guidance tree (decision aid) for hypertension, based on decision analysis: development and preliminary evaluation. **European journal of cardiovascular nursing**, SAGE Publications Sage UK: London, England, v. 5, n. 2, p. 146–149, 2006.
- TREVISAN, D. A.; HOSKYN, M.; BIRMINGHAM, E. Facial expression production in autism: A meta-analysis. **Autism Research**, Wiley Online Library, v. 11, n. 12, p. 1586–1601, 2018.
- TREVISAN, D. F. et al. Aplicativos para intervenção comportamental com estudantes com transtorno do espectro do autismo. **Revista Brasileira de Informática na Educação**, v. 29, p. 1487–1504, 2021.
- TSANGOURI, C. et al. An interactive facial-expression training platform for individuals with autism spectrum disorder. In: IEEE. **MIT Undergraduate Research Technology Conference (URTC)**, 2016 IEEE. [S.l.], 2016. p. 1–3.
- UNITY. **Unity User Manual@ONLINE**. 2018. Disponível em: <<https://docs.unity3d.com/Manual/index.html>>.
- VALSTAR, M.; PANTIC, M. Induced disgust, happiness and surprise: an addition to the mmi facial expression database. In: PARIS, FRANCE. **Proc. 3rd Intern. Workshop on EMOTION (satellite of LREC): Corpora for Research on Emotion and Affect**. [S.l.], 2010. p. 65.
- VALSTAR, M. F.; PANTIC, M. Fully automatic recognition of the temporal phases of facial actions. **IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)**, IEEE, v. 42, n. 1, p. 28–43, 2012.
- VALSTAR, M. F.; PANTIC, M. Fully automatic recognition of the temporal phases of facial actions. **Trans. Sys. Man Cyber. Part B**, IEEE Press, Piscataway, NJ, USA, v. 42, n. 1, p. 28–43, fev. 2012. ISSN 1083-4419. Disponível em: <<http://dx.doi.org/10.1109/TSMCB.2011.2163710>>.
- VALSTAR, M. F. et al. Fera 2017-addressing head pose in the third facial expression recognition and analysis challenge. In: IEEE. **2017 12th IEEE International Conference on Automatic Face & Gesture Recognition (FG 2017)**. [S.l.], 2017. p. 839–847.
- VIOLA, P.; JONES, M. Rapid object detection using a boosted cascade of simple features. In: IEEE. **Computer Vision and Pattern Recognition, 2001. CVPR 2001. Proceedings of the 2001 IEEE Computer Society Conference on**. [S.l.], 2001. v. 1, p. I–I.
- VIOLA, P.; JONES, M. J. Robust real-time face detection. **Int. J. Comput. Vision**, Kluwer Academic Publishers, Hingham, MA, USA, v. 57, n. 2, p. 137–154, maio 2004. ISSN 0920-5691. Disponível em: <<http://dx.doi.org/10.1023/B:VISI.0000013087.49260.fb>>.



- WEITLAUF, A. S. et al. Brief report: Dsm-5 “levels of support:” a comment on discrepant conceptualizations of severity in asd. **Journal of autism and developmental disorders**, Springer, v. 44, n. 2, p. 471–476, 2014.
- WU, B.-F.; LIN, C.-H. Adaptive feature mapping for customizing deep learning based facial expression recognition model. **IEEE access**, IEEE, v. 6, p. 12451–12461, 2018.
- YUSOFF, A. **A conceptual framework for serious games and its validation**. Tese (Doutorado) — University of Southampton, 2010.
- ZABLOTSKY, B.; BRADSHAW, C. P.; STUART, E. A. The association between mental health stress and coping supports in mothers of children with autism spectrum disorders. **Journal of autism and developmental disorders**, Springer, v. 43, n. 6, p. 1380–1393, 2013.
- ZAVAREZ, M. V.; BERRIEL, R. F.; OLIVEIRA-SANTOS, T. Cross-database facial expression recognition based on fine-tuned deep convolutional network. In: IEEE. **2017 30th SIBGRAPI Conference on Graphics, Patterns and Images (SIBGRAPI)**. [S.l.], 2017. p. 405–412.
- ZHANG, Z. et al. From facial expression recognition to interpersonal relation prediction. **Int. J. Comput. Vision**, Kluwer Academic Publishers, USA, v. 126, n. 5, p. 550–569, maio 2018. ISSN 0920-5691. Disponível em: <<https://doi.org/10.1007/s11263-017-1055-1>>.