UNIVERSIDADE DE SÃO PAULO
ESCOLA DE ARTES, CIÊNCIAS E HUMANIDADES
PROGRAMA DE PÓS-GRADUAÇÃO EM SISTEMAS DE INFORMAÇÃO

HENRIQUE SILVA BORTOLETTI

Aquisição de vídeos e processamento de imagens faciais em dispositivos móveis para auxílio ao diagnóstico do Transtorno do Espectro Autista

São Paulo
2022

HENRIQUE SILVA BORTOLETTI

Aquisição de vídeos e processamento de imagens faciais em dispositivos móveis para auxílio ao diagnóstico do Transtorno do Espectro Autista

Dissertação apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação.

Área de concentração: Metodologia e Técnicas da Computação

Orientador(a): Profa. Dra. Fátima de Lourdes dos Santos Nunes Marques

Coorientador(a): Profa. Dra. Ariane Machado Lima

São Paulo

2022

Autorizo a reprodução e divulgação total ou parcial deste trabalho, por qualquer meio convencional ou eletrônico, para fins de estudo e pesquisa, desde que citada a fonte.

Ficha catalográfica elaborada pela Biblioteca da Escola de Artes, Ciências e Humanidades,
com os dados inseridos pelo(a) autor(a)
Brenda Fontes Malheiros de Castro CRB 8-7012; Sandra Tokarevicz CRB 8-4936

Silva Bortoletti, Henrique
Aquisição de vídeos e processamento de imagens faciais em dispositivos móveis para auxílio ao diagnóstico do Transtorno do Espectro Autista / Henrique Silva Bortoletti; orientadora, Fátima de Lourdes dos Santos Nunes Marques; coorientadora, Ariane Machado Lima. -- São Paulo, 2022.
130 p: il.

Dissertacao (Mestrado em Ciencias) - Programa de Pós-Graduação em Sistemas de Informação, Escola de Artes, Ciências e Humanidades, Universidade de São Paulo, 2022.
Versão corrigida

1. Transtorno do espectro autista. 2. Processamento de imagens faciais. 3. Dispositivos móveis. 4. Sistema de auxílio ao diagnóstico. 5. Antropometria. I. Marques, Fátima de Lourdes dos Santos Nunes, orient. II. Lima, Ariane Machado, coorient. III. Título.

Dissertação de autoria de Henrique Silva Bortoletti, sob o título **"Aquisição de vídeos e processamento de imagens faciais em dispositivos móveis para auxílio ao diagnóstico do Transtorno do Espectro Autista"**, apresentada à Escola de Artes, Ciências e Humanidades da Universidade de São Paulo, para obtenção do título de Mestre em Ciências pelo Programa de Pós-graduação em Sistemas de Informação, na área de concentração Metodologia e Técnicas da Computação, aprovada em 11 de Outubro de 2022 pela comissão julgadora constituída pelos doutores:

Prof. Dra. Fátima de Lourdes dos Santos Nunes Marques
Universidade de São Paulo
Presidente

Prof. Dr. David Corrêa Martins Junior
Universidade Federal do ABC

Prof. Dr. Rodrigo da Silva Dias
Universidade de São Paulo

Agradecimentos

Agradeço a Deus. Agradeço aos meus pais por me apoiarem constantemente durante toda a minha vida acadêmica.

Agradeço à minha orientadora, professora Fátima, e minha coorientadora, professora Ariane, pela companhia, pelos conselhos, pela compreensão e por sempre me incentivar a continuar mesmo nos momentos difíceis.

Agradeço aos funcionários e pesquisadores do Instituto de Psiquiatria do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo, por sua imensurável gentileza e pelo apoio durante o desenvolvimento do projeto. Agradeço aos voluntários que participaram dessa pesquisa.

Agradeço ao CNPq - Conselho Nacional de Desenvolvimento Científico e Tecnológico (Processo 309030/2019-6), à FAPESP - Fundação de Amparo à Pesquisa do Estado de São Paulo (Processo 2020/01992-0) e ao Instituto de Ciência e Tecnologia em Medicina Assistida por Computação científica (INCT-MACC) - processo 2014/50889-7, pelo apoio financeiro para a realização dessa pesquisa.

"Viver significa lutar."

(Sêneca)

Resumo

O Transtorno do Espectro Autista (TEA) é uma anomalia do neurodesenvolvimento que afeta principalmente as habilidades de interação social dos indivíduos. Medidas antropométricas extraídas de imagens faciais podem contribuir para auxiliar no diagnóstico do TEA. Um sistema de auxílio ao diagnóstico baseado em tais medidas, em conjunto com técnicas de aprendizado de máquina, foi desenvolvido por meio de uma parceria entre o Laboratório de Aplicações de Informática em Saúde (LApIS), da Escola de Artes, Ciências e Humanidades (EACH) e o Programa de Transtornos do Espectro Autista (PROTEA), do Instituto de Psiquiatria do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (IPq/HCFMUSP). Atualmente o uso de tal sistema requer a presença física do indivíduo no IPq para a coleta das imagens faciais que serão utilizadas para a extração das medidas antropométricas. O presente projeto visa a contribuir para a democratização do acesso ao auxílio ao diagnóstico por meio de uma aplicação para dispositivos móveis, responsável por capturar vídeos faciais contendo imagens com as características adequadas para a extração das medidas antropométricas. A coleta de vídeos pode ser realizada pelo usuário de forma remota a fim de que seja posteriormente analisada pelo sistema desenvolvido. A aplicação é capaz de capturar um vídeo, por meio de um dispositivo móvel, contendo um número pré-estabelecido de imagens com as características adequadas em relação a posicionamento, alinhamento e expressão facial. Em seguida, o vídeo e os dados coletados são transmitidos para um servidor responsável por armazená-los e processá-los. Por meio de uma avaliação de usabilidade foi possível identificar uma dificuldade inicial ao utilizar a aplicação, mas com rápido aprendizado durante o uso da ferramenta.

Palavras-chaves: Transtorno do espectro autista. Processamento de imagens faciais. Dispositivos móveis. Sistema de auxílio ao diagnóstico. Antropometria.

Abstract

Bortoletti, Henrique Silva. **Acquisition of videos and processing of facial images on mobile devices to aid in Diagnosis of Autism Spectrum Disorder**. 2022. 129 f. Dissertation (Master of Science) – School of Arts, Sciences and Humanities, University of São Paulo, São Paulo, 2022.

Autism Spectrum Disorder (ASD) is a neurodevelopmental anomaly that mainly affects the social interaction skills of individuals. Anthropometric measurements extracted from facial images can contribute to aid in the diagnosis of ASD. A diagnostic aid system based on such measurements, in conjunction with machine learning techniques, was developed through a partnership between the Laboratory of Health Informatics Applications (LApIS), of the School of Arts, Sciences and Humanities (EACH) and the Autism Spectrum Disorders Program (PROTEA), of the Institute of Psychiatry, Hospital das Clínicas, Faculty of Medicine, University of São Paulo (IPq/HCFMUSP). Currently, the use of such a system requires the physical presence of the individual at the IPq for the collection of facial images that will be used for the extraction of anthropometric measurements. The present project aims to contribute to the democratization of access to diagnostic aid through an application for mobile devices, responsible for capturing facial videos containing images with the appropriate characteristics for the extraction of anthropometric measurements. The collection of videos can be performed by the user remotely so that it is later analyzed by the developed system. The application is capable of capturing a video, through a mobile device, containing a pre-established number of images with the appropriate characteristics in relation to positioning, alignment and facial expression. Then the video and data collected are transmitted to a server responsible for storing and processing them. Through a usability evaluation, it was possible to identify an initial difficulty when using the application, but with rapid learning during the use of the tool.

Keywords: Autism spectrum disorder. Facial image processing. Mobile device. Diagnostic support system. Anthropometry.

Lista de figuras

Figura 1 – Visão geral das etapas percorridas durante a revisão sistemática. 28
Figura 2 – Categoria dos artigos incluídos de acordo com seus objetivos e áreas. . 30
Figura 3 – Diagrama sumarizado das técnicas de processamento facial utilizadas pelos artigos. 32
Figura 4 – Algoritmos de processamento de imagem facial utilizados por dispositivos móveis. 34
Figura 5 – Bibliotecas mais utilizadas de processamento facial utilizadas pelos artigos. 36
Figura 6 – Métricas mais utilizadas por objetivo dos artigos. 37
Figura 7 – Relação entre as diferentes etapas do projeto. 39
Figura 8 – Pontos de controle da região da boca, que podem possuir as mesmas coordenadas. 42
Figura 9 – Faixa de de referência utilizada durante a coleta de imagens faciais pelo protocolo desenvolvido em (PINHEIROS, 2018). 42
Figura 10 – Exemplo das diferentes localizações faciais pelas bibliotecas. O retângulo indica a área que o algoritmo identificou como sendo a face da pessoa e os pontos dentro do retângulo são os pontos de controle detectados pela biblioteca Dlib. 44
Figura 11 – Pontos de controle da boca utilizados para determinar sua abertura. . . 46
Figura 12 – Arquitetura do processo de captura de imagens adequadas pela aplicação.O módulo destacado na cor verde é responsável pelo 52
Figura 13 – Perspectiva frontal e lateral dos ângulos de Euler sob os eixos $X, Y e Z$. 55
Figura 14 – Pontos de Landmark identificando pontos médios da orelha e cantos internos dos olhos. 56
Figura 15 – Pontos de controle em azul identificam a parte central dos lábios superior e inferior do indivíduo. 57
Figura 16 – Possíveis situações de falha e sucesso durante a captura das imagens faciais (a) Falha em todas as validações durante a captura. (b) Sucesso em todas as validações durante a captura. 58
Figura 17 – Modelo relacional para o armazenamento das imagens adequadas. . . . 61

Figura 18 – Tempo necessário para cada voluntário da área da computação realizar as tarefas. 64
Figura 19 – Fluxograma de tarefas presentes no questionário. 65
Figura 20 – Escala *Likert* para a pergunta Q3(O(a) Senhor(a) teve alguma dificuldade ao realizar a tarefa?) 67
Figura 21 – Interface inicial da aplicação com a opção para cadastro de novo usuário destacada por um contorno vermelho. 70
Figura 22 – Interface da aplicação para cadastro de usuário. 71
Figura 23 – Interface inicial da aplicação desenvolvida. 72
Figura 24 – Interface da aplicação para iniciar o processo de troca de senha. 73
Figura 25 – Interface da aplicação para registrar novo paciente ou iniciar uma nova captura (a) Interface dos pacientes registrados sem nenhum paciente. (b) Interface dos dados do paciente a ser registrado. (c) interface dos pacientes registrados com um paciente selecionado 75
Figura 26 – Processo de captura de vídeo facial do paciente (a) dados relacionados ao paciente necessários para a captura do vídeo facial (b) Captura em andamento. (c) Imagens adequadas encontradas (d) Mensagem de captura realizada com sucesso. 76
Figura 27 – Planilha de texto responsável por gerenciar os parâmetros da aplicação. 78
Figura 28 – Cognição da interface da aplicação pelos voluntários a cada tarefa realizada. 80
Figura 29 – Porcentagem de voluntários cansados e de voluntários não cansados após a realização de cada tarefa por meio da aplicação desenvolvida. . . 80
Figura 30 – Porcentagem de pesquisadores da área da saúde por tarefa que precisaram utilizar o manual de instruções durante o experimento. 82
Figura 31 – Porcentagem de pesquisadores da área da computação por tarefa que precisaram utilizar o manual de instruções durante o experimento. . . . 83
Figura 32 – Dificuldade dos pesquisadores da área da saúde ao realizar determinada tarefa durante o experimento. 83
Figura 33 – Dificuldade dos pesquisadores da área da computação ao realizar determinada tarefa durante o experimento. 84
Figura 34 – Tempo necessário para cada voluntário da área da saúde realizar as tarefas. 85

Figura 35 – Tempo necessário para cada voluntário da área da computação realizar as tarefas. 85
Figura 36 – Indisponibilidade do sistema desenvolvido durante o experimento. . . . 87
Figura 37 – Falhas identificadas de interpretação pelos participantes ao utilizar a aplicação ou ao realizar determinada tarefa durante o experimento. . . 89

Lista de tabelas

Tabela 1 – *Strings* utilizadas em cada base indexada. 26
Tabela 2 – Critérios de inclusão (I) e exclusão (E) 26
Tabela 3 – Critérios de qualidade. 27
Tabela 4 – Número de distâncias entre pares de algoritmos que possuem diferenças significativas. 44
Tabela 5 – Número de imagens por síndrome. 45
Tabela 6 – Descrição dos atributos presentes nas entidades do modelo relacional. . 62
Tabela 7 – Estudos selecionados 105

Lista de abreviaturas e siglas

EACH	Escola de Artes, Ciências e Humanidades
IPq	Instituto de Psiquiatria
HCFMUSP	Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo
TEA	Transtorno do Espectro Autista
STAT	*Screening Tool for Autism in Toddlers and Young Children*
ADOS	*Autism Diagnostic Observation Schedule*
FACS	*Facial Action Coding System*
CNN	*Convolutional Neural Network*
CLNF	*Constrained Local Neural Fields*
HOG	*Histogram of Oriented Gradients*
DNN	*Deep Neural Network*
LFW	*Labeled Faces in the Wild*
LBP	*Local Binary Patterns*
LBPH	*Local Binary Patterns Histogram*
API	*Application Programming Interface*
REST	*Representational State Transfer*
HTTPS	*Hyper Text Transfer Protocol Secure*

Sumário

1 Introdução 16

1.1 Motivação e justificativa 17

1.2 Objetivos 18

1.3 Organização do documento 18

2 Conceitos fundamentais 19

2.1 Transtorno do Espectro Autista 19

2.2 Medidas antropométricas 20

2.3 Detecção facial 21

2.4 Reconhecimento de emoções 22

3 Revisão bibliográfica sistemática 25

3.1 Método 25

3.2 Resultados e discussões 28

3.2.1 Áreas de aplicação 28

3.2.2 Objetivos dos artigos 29

3.2.3 Técnicas utilizadas para o processamento facial em dispositivos móveis 32

3.2.4 Tecnologias, plataformas e bibliotecas 35

3.2.5 Bases de dados e métricas 36

3.3 Considerações finais 38

4 Materiais e métodos 39

4.1 Etapas da pesquisa 39

4.2 Avaliação de bibliotecas de detecção facial 40

4.2.1 Obtenção das distâncias entre os pontos de controle 41

4.2.2 Comparação dos algoritmos utilizados 43

4.3 Identificação de imagens faciais adequadas 45

4.3.1 Base de dados 45

4.3.2 Resultados do experimento 47

4.4 Análise de requisitos 47

4.5 Desenvolvimento do método 49

4.5.1 Tecnologias 50
4.5.2 Arquitetura da proposta 51
4.5.3 Detecção facial 53
4.5.4 Alinhamento facial 54
4.5.5 Reconhecimento de expressão facial 56
4.5.6 Validação dos parâmetros 57
4.5.7 Transmissão de imagem e dados 59
4.5.8 Recorte dos *frames* 60
4.5.9 Armazenamento 60
4.5.10 Extração de características e modelo de classificação 61
4.6 Avaliação da aplicação 62
4.6.1 Questionário 63
4.6.2 Condução dos experimentos 68
4.7 Considerações finais 69

5 Aplicação desenvolvida 70
5.1 Cadastro de usuário 70
5.2 Autenticação 71
5.3 Troca de senha 73
5.4 Novo paciente 74
5.5 Captura de vídeo do paciente 74
5.6 Configuração dos parâmetros 77
5.7 Considerações finais 77

6 Resultados e discussões 79
6.1 Esforço cognitivo 79
6.2 Curva de aprendizado 81
6.3 Falhas identificadas 85
6.3.1 Falhas de disponibilidade 86
6.3.2 Falhas de interpretação da interface ou tarefas 87

7 Conclusões e trabalhos futuros 90
7.1 Trabalhos futuros 91

REFERÊNCIAS 92

Apêndice A – Tabela dos artigos utilizados na revisão sistemática 104

Apêndice B – Questionário de pesquisa. 111

Apêndice C – Manual da aplicação. 112

1 Introdução

A detecção facial é um dos tópicos mais estudados na área de processamento de imagens. Constitui parte fundamental nos processos de alinhamento, reconhecimento facial e identificação de expressões faciais humanas (YANG *et al.*, 2016), como também na extração de pontos de controle, que, entre outras finalidades, podem ser utilizados para medir as distâncias craniofaciais (WEINBERG *et al.*, 2006).

As medidas craniofaciais podem ser utilizadas em várias aplicações, incluindo auxílio ao diagnóstico de anomalias diversas. Uma das áreas que podem ser beneficiadas é o diagnóstico de crianças com o TEA, pois algumas distâncias craniofaciais podem apresentar dimensões diferentes quando comparadas com as medidas extraídas em indivíduos com desenvolvimento típico (RODIER; BRYSON; WELCH, 1997; ALDRIDGE *et al.*, 2011). Dessa forma, é possível auxiliar o diagnóstico de indivíduos com TEA por meio de algoritmos de classificação que utilizam essas medidas como dados de entrada para algoritmos computacionais (PINHEIROS, 2018).

As tarefas de processamento de imagens faciais possuem alguns desafios relacionados à iluminação do ambiente, oclusão e às diversas possibilidades de posicionamento do rosto em uma imagem (ZHANG *et al.*, 2016). Em Pinheiros (2018), foi planejado um ambiente controlado para solucionar tais desafios; no entanto, isso diminui a democratização do acesso ao auxílio diagnóstico de indivíduos com TEA, já que dessa forma é necessário o deslocamento das pessoas até o ambiente controlado. Adicionalmente, isso pode não ser possível em determinadas situações, como na pandemia causada pelo coronavírus (COVID-19), oficialmente declarada pela Organização Mundial da Saúde em março de 2020, que afetou mais de 170 países e levou os governos a adotarem medidas restritivas de circulação de pessoas (ZHANG; HU; JI, 2020).

Portanto, uma aplicação baseada na arquitetura cliente e servidor (MUSLIM; ISLAM, 2017) pode auxiliar a coleta de imagens faciais adequadas para o auxílio ao diagnóstico, de forma remota. Contudo, tal aplicação necessita estabelecer um padrão de imagens faciais capturadas, apresentar interfaces com usabilidade adequada, que promova a fácil compreensão dos diferentes tipos de usuários.

1.1 Motivação e justificativa

Este projeto de pesquisa faz parte de uma colaboração entre o Laboratório de Aplicações de Informática em Saúde (LApIS), da Escola de Artes, Ciências e Humanidades (EACH) e o Programa de Transtornos do Espectro Autista (PROTEA), do Instituto da Psiquiatria do Hospital das Clínicas da Faculdade de Medicina da Universidade de São Paulo (IPq/HCFMUSP).

Nesta parceria está sendo desenvolvido um sistema de auxílio ao diagnóstico automatizado que visa a contribuir para a detecção do TEA. O sistema adquire sinais de rastreamento de olhar e imagens faciais e os processa por meio de técnicas de processamento de imagem e reconhecimento de padrões. Assim, os sinais são utilizados como entrada de um modelo de classificação, que auxilia no diagnóstico de indivíduos com TEA. Em especial, a presente proposta está relacionada com o processamento de medidas antropométricas extraídas de imagens faciais.

Atualmente é necessário o deslocamento do indivíduo até o Instituto de Psiquiatria para a captura das imagens. Durante a coleta de imagens não é realizada uma análise para verificar se possuem as características necessárias para o auxílio ao diagnóstico. A coleta é feita por uma máquina fotográfica, o indivíduo deve estar posicionado corretamente abaixo de uma faixa vermelha, conforme descrito em Pinheiros (2018). Outro desafio durante a coleta presencial é o estresse causado nos indivíduos durante a captura, pois é necessário que estejam na posição específica descrita sem se movimentar durante alguns minutos até a imagem ser capturada. Atualmente a presença da faixa vermelha é indispensável, pois a sua largura e o seu comprimento são utilizados para a normalização das distâncias calculadas.

O desenvolvimento de mecanismos de captura de imagens faciais por dispositivos móveis pode contribuir para a expansão do acesso aos procedimentos de auxílio ao diagnóstico, ao mesmo tempo, contribui para a ampliação da base de dados de pesquisa, disponibilizando mais instâncias para o treinamento do modelo de classificação. Ainda, será possível identificar a imagem adequada sem a necessidade do indivíduo permanecer imóvel ou a necessidade de uma faixa de referência. Com isso, é possível reduzir um eventual estresse durante a captura e eliminar as imagens faciais inadequadas capturadas. Por

fim, o material coletado também poderá ser utilizado para apoiar o diagnóstico de outras patologias.

1.2 Objetivos

O objetivo principal deste trabalho é desenvolver e validar uma abordagem para adquirir e processar em tempo real imagens faciais em dispositivos móveis, garantindo a adequação para aplicação em sistemas de auxílio ao diagnóstico de TEA.

Para alcançar o objetivo principal são estabelecidos os seguintes objetivos específicos:

- definir parâmetros para capturar as imagens apropriadas para a análise, considerando os diferentes sistemas operacionais presentes nos dispositivos móveis;
- definir e implementar técnicas de processamento de imagens para verificar a adequação das imagens coletadas;
- definir e implementar uma forma eficiente de transmissão de imagens e dados do dispositivo móvel para o servidor que possui o sistema de auxílio ao diagnóstico;
- armazenar as imagens e vídeos coletados de maneira adequada para uso posterior pelo sistema de auxílio ao diagnóstico.

1.3 Organização do documento

O restante do presente documento está organizado da maneira descrita a seguir. No Capítulo 2, os conceitos fundamentais referentes ao TEA, às medidas antropométricas, à detecção facial e ao reconhecimento de emoções são apresentados. O Capítulo 3 apresenta a revisão bibliográfica conduzida sobre o processamento em tempo real de imagens faciais em dispositivos móveis. Os métodos utilizados e a implementação do método são apresentados nos Capítulos 4 e 5, respectivamente. Os resultados e discussões são apresentados no Capítulo 6 e o Capítulo 7 apresenta as considerações finais do documento.

2 Conceitos fundamentais

Neste Capítulo os principais conceitos envolvendo o projeto de pesquisa são apresentados. A Seção 2.1 apresenta os conceitos gerais relacionados ao TEA. A Seção 2.2 define e exemplifica o conceito de medidas antropométricas. As seções 2.3 e 2.4 apresentam os principais conceitos utilizados por este trabalho relacionados ao processamento de imagens faciais e ao reconhecimento de emoções.

2.1 Transtorno do Espectro Autista

O TEA é caracterizado por um conjunto de distúrbios do neurodesenvolvimento, que envolvem prejuízos funcionais em relação à interação social e à comunicação de um indivíduo, assim como a presença de problemas de comportamento repetitivo restrito (BLUMBERG *et al.*, 2015). Independentemente da cultura, raça ou grupo social, o TEA pode comprometer duas principais áreas da vida de um indivíduo: comunicação social e comportamento neuromotor (LORD *et al.*, 2018).

A manifestação do comprometimento dessas duas áreas pode variar bastante de indivíduo para indivíduo. Por isso, definir o espectro autista é importante para mapear a diversidade desses distúrbios (FRITH; HAPPé, 2005).

Atualmente o diagnóstico do TEA é clínico, mas existem escalas e instrumentos de avaliação utilizados para auxiliar o diagnóstico. Os únicos considerados como "padrão-ouro" são:

1. **Protocolo de Observação para Diagnóstico de TEA (ADOS):** fornece uma escala de observação da comunicação e dos comportamentos típicos de indivíduos portadores de TEA;
2. **Entrevista Diagnóstica para Autismo Revisada (ADI-R):** é composta por uma escala obtida a partir de entrevistas como meio de avaliar uma sequência de comportamentos que são condizentes com pessoas que apresentam o TEA.

Embora reconhecidos como eficientes, esses instrumentos têm acesso limitado, visto que sua aplicação requer profissionais treinados e pagamento de direitos autorais.

Uma das descobertas mais importantes nos últimos anos foi a associação do crescimento do volume e peso do cérebro durante a infância dos indivíduos diagnosticados

com TEA. É importante destacar que em outras patologias relacionadas ao atraso no desenvolvimento do indivíduo, geralmente observa-se a diminuição do tamanho do cérebro (FRITH; HAPPé, 2005; CHASTE *et al.*, 2013; SACCO; GABRIELE; PERSICO, 2015; TANG *et al.*, 2021).

Por fim, é essencial que o diagnóstico seja realizado o mais rápido possível na vida de um indivíduo, pois quanto mais cedo acontecer o tratamento intensivo, maior será a probabilidade de melhorar os sintomas a longo prazo do TEA (BLUMBERG *et al.*, 2015).

2.2 Medidas antropométricas

Antropometria é a ciência responsável por medir o tamanho e as proporções de partes do corpo humano. A partir do uso dessas medidas é possível caracterizar variações e dismorfismos fenotípicos (FARKAS, 1982). As medidas podem ser coletadas em clínicas utilizando ferramentas tradicionais como fitas métricas e paquímetros (JUNQUEIRA *et al.*, 2016). Com o avanço da tecnologia, outras formas de coleta foram desenvolvidas como: escaneamento a laser, ressonância magnética, ultrassom e a utilização de câmeras fotográficas. Esta última representa uma maneira menos invasiva de se obter as medidas desejadas e as imagens não precisam ser capturadas necessariamente em uma clínica (JUNQUEIRA *et al.*, 2016).

Para descrever o rosto humano as medidas de distância são amplamente utilizadas. Geralmente são extraídas a partir dos pontos de controle (*landmarks*), que também são conhecidos como pontos de referência antropométrica (VEZZETTI; MARCOLIN, 2012). Em Farkas (1982) é proposto um manual descrevendo tais pontos de controle e definindo normas para a captura dos dados antropométricos. Estes pontos faciais são importantes porque estão presentes em todos os rostos e possuem um valor biológico (VEZZETTI; MARCOLIN, 2012).

As aplicações que podem ser desenvolvidas a partir do uso dessas medidas faciais são inúmeras, como a detecção e o reconhecimento facial, o reconhecimento de emoções faciais e estudo das mudanças morfológicas faciais devido ao envelhecimento ou dismorfologias (VEZZETTI; MARCOLIN, 2012).

2.3 Detecção facial

A detecção facial é uma tarefa que nossos cérebros realizam diariamente de maneira automática, porém não é tão simples para uma máquina realizar esse processamento. Uma visão geral de como o problema pode ser definido computacionalmente é: a partir de uma imagem ou vídeo, detectar e localizar um número desconhecido de faces humanas independentemente de iluminação ambiente, orientação e distância da câmera (HJELMåS; LOW, 2001).

Esse é um dos tópicos mais estudados na área de processamento de imagens. Constitui parte fundamental nos processos de alinhamento, reconhecimento facial e identificação de expressões faciais humanas (YANG *et al.*, 2016). Existem três principais categorias de métodos utilizados em dispositivos móveis para realizar essa tarefa (RATTANI; DERAKHSHANI, 2018), conforme explicitado a seguir.

Na primeira categoria (algoritmos baseados na coloração da pele), tais algoritmos são extremamente rápidos para o processamento das características faciais, porém são altamente impactados pela iluminação ambiente e reflexão especular. Além disso, imagens capturadas por câmeras fotográficas diferentes podem produzir diversos valores para a coloração de imagens semelhantes (RATTANI; DERAKHSHANI, 2018).

A segunda categoria (algoritmos de aprendizado de máquina) engloba algoritmos que são menos impactados pela iluminação ambiente e mais efetivos para a detecção facial, porém possuem dificuldades em reconhecer diferentes poses faciais que não estão no conjunto de treinamento e possuem um alto custo computacional (RATTANI; DERAKHSHANI, 2018).

A terceira categoria (híbrida) é uma combinação dos algoritmos baseados na coloração da pele com os algoritmos de aprendizado de máquina, tornando possível combinar as vantagens de ambas as técnicas citadas para aumentar a acurácia (RATTANI; DERAKHSHANI, 2018; KANTHARIA; PRAJAPATI, 2015).

Independentemente do método utilizado, a iluminação ambiente ainda pode ser um desafio para as tarefas que utilizam a detecção facial, sendo um dos tópicos mais citados nos artigos relacionados ao reconhecimento facial (PRADO *et al.*, 2016). Uma forma de corrigir a iluminação ambiente pode ser por meio do método apresentado em Tan e Triggs (2010), composto por uma sequência de técnicas de processamento de imagens.

O primeiro passo dessa sequência é a transformação não linear dos níveis de cinzas por meio da correção *Gamma*, técnica utilizada para compensar as características de exibição não lineares de um dispositivo (MCREYNOLDS; BLYTHE, 2005). A correção *Gamma* não remove a influência dos efeitos de sombreamento. Dessa forma, o próximo passo da sequência é a aplicação de um filtro baseado na diferença *Gaussiana* entre os *pixels* da imagem. Por fim, o contraste da imagem é equalizado por uma função definida pelos próprios autores.

Ainda em Tan e Triggs (2010) a etapa de correção da iluminação é comparada com outros métodos presentes no estado da arte: equalização de histogramas, variação total logarítmica (CHEN *et al.*, 2006), *Self-Quotient Image* (WANG; LI; WANG, 2004), *Multiscale Retinex* (MA *et al.*, 2017) e a suavização anisotrópica de *Gross & Brajovic* (GROSS RALPHAND BRAJOVIC, 2003). A análise do impacto do algoritmo de correção é realizada por meio da acurácia de um sistema de reconhecimento facial utilizando a base de dados *CMU-PIE* (SIM; BAKER; BSAT, 2003).

2.4 Reconhecimento de emoções

O reconhecimento de emoções depende dos gestos, poses, expressões faciais, pronúncia das palavras, comportamento do indivíduo e até mesmo sinais cerebrais (CHENG *et al.*, 2017). As expressões faciais representam uma parte essencial na comunicação humana, possibilitando uma rápida transmissão de informação de um indivíduo para o outro (BLAIR, 2003).

Por isso, pesquisadores da área de processamento de imagens faciais têm estudado cada vez mais o reconhecimento de expressões faciais (HAPPY; ROUTRAY, 2015). Para realizar esta tarefa, a maior parte das pesquisas se baseia no sistema "universal" das seis emoções básicas discretas: raiva, nojo, medo, felicidade, surpresa e tristeza (EKMAN *et al.*, 1987). Dessa forma pode-se utilizar o reconhecimento de emoções em aplicações relacionadas a interação humano-computador, assistência médica, vigilância, comportamento do motorista, dentre outras (HAPPY; ROUTRAY, 2015).

De acordo com Canedo e Neves (2019), um sistema de reconhecimento de expressões faciais é composto por quatro passos: captura da imagem, pré-processamento, extração de características, classificação ou regressão.

O pré-processamento é extremamente importante, pois realiza a detecção facial. Neste passo a maior parte do artigos selecionados na revisão bibliográfica apresentada por Canedo e Neves (2019) utilizaram o algoritmo proposto em Viola e Jones (2001). Esta etapa também garante uma qualidade aceitável dos dados utilizados pela próxima etapa de extração das características. Tal qualidade pode ser obtida por meio da normalização da iluminação, contraste e tamanho da imagem (CANEDO; NEVES, 2019).

Após a etapa de pré-processamento, a imagem passa por um processo de extração das características faciais. Durante essa etapa é possível encontrar uma série de desafios, pois a qualidade da extração dessas características apresenta impacto na precisão do sistema. É importante ressaltar que a maior parte das metodologias utilizadas dependem do reconhecimento dos pontos de controle (Seção 2.2) para atingir seus respectivos objetivos. Dessa forma, um dos desafios encontrados nesta etapa é o reconhecimento das características faciais em imagens faciais capturadas em ambientes com iluminação ou posição não controlados (HAPPY; ROUTRAY, 2015). Por exemplo, se a face estiver em direção ao lado direito de tal forma que não seja possível identificar o olho deste lado, então pode não ser possível capturar todas as características faciais necessárias.

Para realizar a extração das características faciais, pode-se utilizar o algoritmo *Facial Action Coding System* (FACS), que utiliza os pontos de controle para definir unidades de ação, que representam o movimento da musculatura do indivíduo. Tais unidades de ação consistem basicamente da contração ou relaxamento de determinados músculos do rosto, que em conjunto podem ser utilizadas para o reconhecimento de expressões faciais (EKMAN; ROSENBERG, 2005).

Por fim, a última etapa consiste na classificação, responsável por identificar a classe à qual a imagem pertence, considerando as seis emoções discretas anteriormente citadas. Nesta etapa, pode-se utilizar modelos de classificação ou de regressão para determinar as expressões do indivíduo (CANEDO; NEVES, 2019). Geralmente as técnicas utilizadas nesta etapa são adaptações dos algoritmos de máquinas de vetores de suporte (BOSER; GUYON; VAPNIK, 1992) ou de redes convolucionais (LIU *et al.*, 2017).

É importante ressaltar, que em dispositivos móveis pode-se aplicar outras etapas durante o processo de reconhecimento de expressões faciais, como será apresentado na Subseção 3.2.3. Independentemente das técnicas escolhidas nas etapas descritas anteriormente, os maiores desafios para o reconhecimento de expressões faciais continuam sendo a

posição facial, a iluminação e o reconhecimento de expressões em ambientes não controlados (CANEDO; NEVES, 2019; HAPPY; ROUTRAY, 2015).

3 Revisão bibliográfica sistemática

Com o objetivo de fornecer uma visão geral da área, este capítulo apresenta uma revisão sistemática visando a analisar as principais técnicas utilizadas para detecção facial, as áreas de aplicação, os algoritmos de processamento facial utilizados, as bibliotecas utilizadas e os aspectos relacionados à avaliação de abordagens apresentadas.

Além desta Seção introdutória, o presente Capítulo é constituído pelas seguintes seções: na Seção 3.1 será apresentado o método utilizado para desenvolver a revisão sistemática; na Seção 3.2 serão apresentados os resultados e discussões obtidos após o planejamento, condução e extração dos dados e na Seção 3.3 serão apresentadas as contribuições para esta pesquisa e as considerações finais.

3.1 Método

Esta revisão sistemática foi dividida em três etapas principais: protocolo, condução e extração de dados.

A elaboração do protocolo iniciou-se com a definição das principais questões para atingir os objetivos:

- quais algoritmos são usados para a detecção facial em dispositivos móveis?
- qual é o objetivo de cada técnica de processamento de imagem facial utilizada na literatura?
- quais são as bibliotecas e tecnologias utilizadas pelos artigos selecionados?
- quais métricas foram utilizadas pelos artigos para mensurar as técnicas aplicadas?
- como as técnicas de processamento de imagem facial foram aplicadas?

Após uma análise exploratória conduzida antes da primeira etapa da revisão, que utilizou termos relacionados à detecção facial em dispositivo móveis, foram selecionadas as seguintes bases de dados indexadas para identificar os artigos utilizados pela revisão bibliográfica: Association for Computing Machinery (ACM DL) [1], Institute of Electrical and Electronics Engineers (IEEE) Xplore [2] e Science Direct [3].

[1] https://dl.acm.org/
[2] https://ieeexplore.ieee.org/
[3] https://www.sciencedirect.com/

A Tabela 1 apresenta as *Strings* de busca aplicadas a cada base de dados. O resultado representa o número de trabalhos retornados, tendo sido excluídos os trabalhos duplicados. Estes artigos foram avaliados de acordo com os critérios de inclusão e exclusão, que foram definidos durante a etapa de protocolo e aplicados durante a etapa de condução. Tais critérios são apresentados na Tabela 2, que apresenta (na coluna Resultado) o número de artigos que possuem determinado critério e a porcentagem dos artigos que o critério foi atribuído. Dessa forma, ao mesmo artigo podem ser atribuídos vários critérios de exclusão ou inclusão.

Tabela 1 – *Strings* utilizadas em cada base indexada.

Base indexada	*String* de busca	Resultado
ACM	{ "query": { Abstract:(fac* OR emotion) AND Abstract:(detection OR recognition OR tracking OR "liveness detection") AND Abstract:(mobile) AND Abstract:("cloud computing"OR "real time") } "filter": { Article Type: Research Article, ACM Content: DL }}	350 (33%)
IEEE	(((*fac* OR emotion) AND (detection OR recognition OR tracking OR "liveness detection")) AND (pose OR illumination) AND ((mobile OR "mobile devices"OR "biometrics mobile") AND (communication OR computing OR "cloud computing"OR real-time)))	517 (48%)
Science Direct	(face OR facial OR emotion) AND (detection OR recognition OR tracking OR attributes) AND (mobile)	207 (19%)

Fonte: Henrique Silva Bortoletti, 2022.

Tabela 2 – Critérios de inclusão (I) e exclusão (E)

Critério	Descrição	Resultados (%)
I	Artigo sobre detecção facial, alinhamento ou reconhecimento de expressão facial em dispositivos móveis	99 (9%)
E	Artigo sem resumo	3 (0.27%)
E	Artigo não está escrito em inglês	3 (0.27%)
E	Artigo com menos de três páginas	20 (1.86%)
E	Artigo não é sobre processamento de imagem facial	921 (85.75%)
E	Artigo sobre processamento facial, mas não é aplicado em tempo real e não utiliza dispositivos móveis	28 (2.60%)
E	Artigo sem método ou resultado	23 (2.14%)

Fonte: Henrique Silva Bortoletti, 2022.

Durante a condução, primeiramente todo artigo que atendesse ao critério de inclusão e não atendesse a algum critério de exclusão foi selecionado para a avaliação de qualidade.

Em seguida, cada cada critério de qualidade (Tabela 3), recebeu um valor, sendo zero se o artigo não atendeu ao critério em questão; 0,5 se atendeu parcialmente ou 1 se atendeu plenamente ao critério. Desde modo, a soma total dos critérios de qualidade foi computada, representando a qualidade do artigo para a revisão sistemática.

Tabela 3 – Critérios de qualidade.

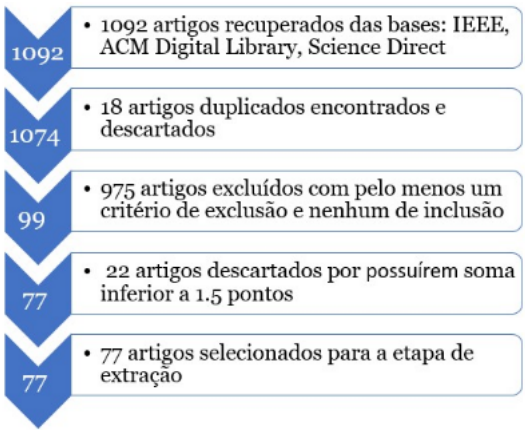
Critérios	Descrição
Q1	Testes de usuário são aplicados?
Q2	Foram realizados experimentos para medir o desempenho das técnicas utilizadas em tempo real?
Q3	Gráficos, figuras, pseudocódigos, diagramas ou qualquer outro elemento visual foi utilizado para descrever os métodos?
Q4	O artigo comparou os resultados com outros trabalho relacionados?
Q5	Métricas quantitativas foram aplicadas(acurácia, precisão e revocação)?
Q6	A base de dados utilizada está disponível para o público?

Fonte: Henrique Silva Bortoletti, 2022.

Após a aplicação dos critérios qualidade, aqueles que obtiveram uma soma igual ou maior a 1,5 foram selecionados para a etapa de extração de dados. Este limiar de pontuação foi definido baseando-se na média da soma dos critérios de qualidade que os artigos atingiram. Nesta etapa foram lidos 77 artigos e extraídos os dados relativos aos objetivos desta revisão sistemática.

A Figura 1 apresenta uma visão geral do processo conduzido. Os artigos incluídos na etapa de extração estão presentes na Tabela 7, que apresenta, para cada trabalho, a área de aplicação, a finalidade do artigo, as métricas e a classificação da base de dados utilizada (pública ou privada). A busca nas bases de dados indexadas retornou 1092 artigos publicados no período entre 2004 a 2020, dos quais 18 eram artigos duplicados. Após a aplicação dos critérios de inclusão e exclusão, 975 itens preencheram, no mínimo, um critério de exclusão. Em seguida, os 99 artigos restantes que atenderam ao critério de inclusão foram avaliados na etapa de avaliação da qualidade, na qual 22 artigos foram excluídos. Ao final, foram analisados 77 artigos na etapa de extração. No Apêndice A são descritos os principais dados de cada artigo analisado: a área de atuação, a finalidade, as métricas utilizadas e o tipo de base de dados utilizada no artigo (pública ou privada).

Figura 1 – Visão geral das etapas percorridas durante a revisão sistemática.



Fonte: Henrique Silva Bortoletti, 2022.

3.2 Resultados e discussões

3.2.1 Áreas de aplicação

Durante a etapa de extração dos dados da revisão sistemática, três diferentes áreas de aplicação foram identificadas: Segurança, Saúde e Educação. Adicionalmente, os artigos que contribuíam para o desenvolvimento da área de Computação, sem ter objetivos diretamente ligados a alguma área de aplicação, foram classificados na categoria de mesmo nome.

A maior parte do artigos (51%) foi classificada na área de computação, já que o objetivo dos autores foi contribuir tecnicamente, por meio da melhoria da experiência do usuário ou da melhoria das técnicas presentes no estado da arte. Por exemplo, em Liu *et al.* (2019) o objetivo do autor é desenvolver uma técnica eficiente de alinhamento facial, que possui uma velocidade maior que as outras técnicas presentes no estado da arte que podem ser aplicadas em dispositivos móveis. Em Cheng *et al.* (2012) os autores avaliaram como melhorar a detecção da posição facial do usuário, por meio de algoritmos de rastreamento facial. Neste caso, os autores não tinham o objetivo de aplicar esta técnica em alguma outra área específica do conhecimento, mas de melhorar a forma como os rostos humanos são identificados em ambientes não controlados.

Os estudos classificados na área da Segurança exploram as técnicas de reconhecimento facial ou de rastreamento facial, que podem ser aplicadas para identificar suspeitos ou para a autenticação do indivíduo. Por exemplo, em Wang *et al.* (2012) é descrito um sistema capaz de reconhecer imagens faciais baseado na localização do indivíduo. Os

autores afirmam que a técnica de busca facial desenvolvida é mais eficiente que as técnicas presentes no estado da arte, pois esta técnica busca primeiro pelos indivíduos na base de dados de moradores próximos à região do indivíduo que se deseja conhecer a identidade. Dessa forma, o objetivo do artigo é reconhecer imagens faciais e a sua finalidade é aplicar tais técnicas na área de Segurança para a identificação de suspeitos. Em Muslim e Islam (2017) foi desenvolvida uma aplicação em nuvem para o reconhecimento facial utilizando dispositivos móveis, que capturam imagens faciais e as enviam para o servidor responsável pelo processamento da imagem e reconhecimento. Os autores destacam que a técnica poderia ser utilizada por sistemas de vigilância e Segurança.

As áreas de Saúde e Educação apresentam poucos artigos. Na área da Saúde as técnicas podem ser aplicadas a algumas finalidades diferentes, por exemplo, identificação do cansaço de motoristas (GHAZAL *et al.*, 2018), medição do pulso (HUANG; YANG; CHENG, 2016) ou auxiliar crianças com autismo a reconhecerem expressões faciais (WASHINGTON *et al.*, 2017). Na área da Educação, os autores em Mittal, Agarwal e Nigam (2018) aplicaram uma técnica de reconhecimento facial para identificar os alunos presentes em aula.

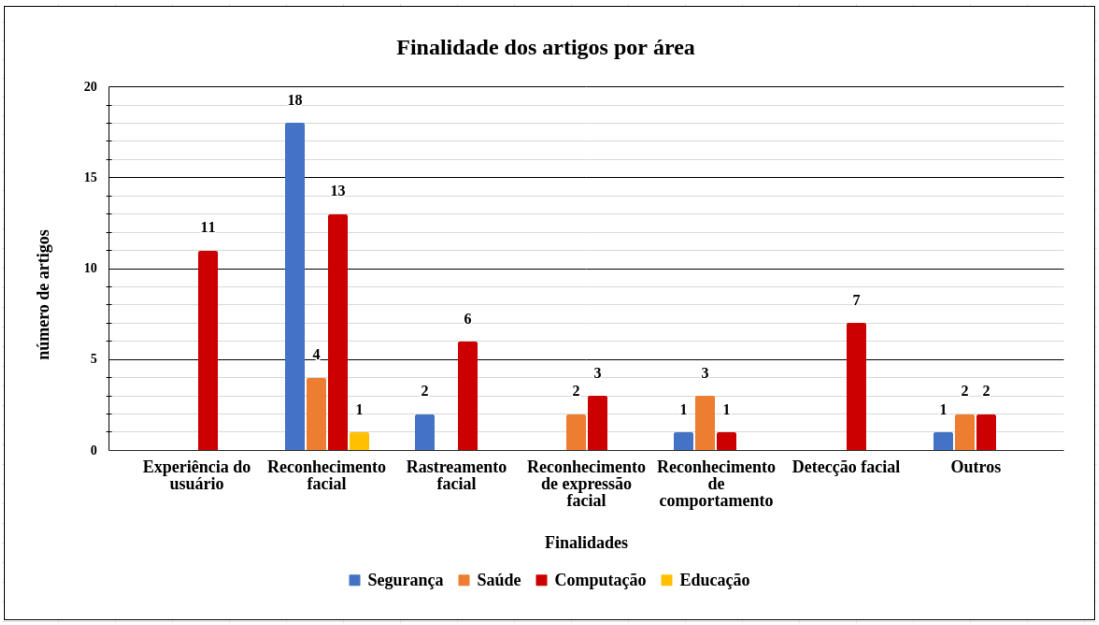
3.2.2 Objetivos dos artigos

A Figura 2 apresenta os artigos classificados de acordo com suas finalidades em sete categorias e sua relação com as áreas de aplicação citadas anteriormente. Os objetivos dos artigos são descritos a seguir.

Detecção facial: nesta categoria encontram-se os artigos cujo objetivo principal é aprimorar técnicas relacionadas à demarcação de rostos humanos em imagens capturadas por dispositivos móveis. Por exemplo, em Lee, Park e Park (2006) foi identificada uma técnica que melhorou o desempenho da detecção facial ao utilizar a diferença da intensidade das cores presentes entre as regiões vizinhas da imagem facial, tais como queixo, olhos e boca. Também nesta categoria, encontram-se os artigos com novas propostas de detecção facial. Em Mahbub, Sarkar e Chellappa (2019), um novo método baseado em redes convolucionais foi aplicado para a detecção facial de imagens que possuem algum tipo de oclusão durante a captura por dispositivos móveis.

Reconhecimento facial: os artigos neste grupo usam ou criam técnicas para o reconhecimento facial utilizando imagens faciais para resolver diferentes problemas.

Figura 2 – Categoria dos artigos incluídos de acordo com seus objetivos e áreas.



Fonte: Henrique Silva Bortoletti, 2022.

Em Mittal, Agarwal e Nigam (2018) um aplicativo em tempo real foi desenvolvido para reconhecer os alunos em uma sala de aula. De forma inovadora, os autores de Wong *et al.* (2010) implementaram uma técnica de reconhecimento facial com um sistema de áudio para auxiliar pessoas com deficiência visual a se autenticarem em seus dispositivos móveis.

Experiência do usuário: este grupo engloba artigos que visam a melhorar a experiência do usuário, explorando técnicas de processamento de imagens faciais. Em Ma *et al.* (2012) foi desenvolvida uma técnica de processamento de imagens faciais para ajustar o nível de brilho de um dispositivo móvel, partindo da premissa de que níveis muito altos de iluminação podem prejudicar a visão do usuário e níveis muito baixos podem dificultar a compreensão do conteúdo exibido pelo dispositivo móvel.

Reconhecimento de expressão facial: o objetivo principal dos artigos deste grupo é reconhecer as expressões faciais do usuário ou de quem está próximo ao usuário, visando principalmente a identificar emoções. Em Washington *et al.* (2017), o autor usou o reconhecimento de expressões faciais para auxiliar crianças autistas a identificar as emoções de seus familiares.

Reconhecimento de comportamento: neste grupo estão presentes os artigos que utilizam a detecção facial para reconhecer o comportamento dos indivíduos. A maior parte dos trabalhos (75%) está relacionada com o comportamento de motorista enquanto dirigem e a outra parte visa a detectar o comportamento da atividade facial enquanto um

usuário está tentando se autenticar em seu dispositivo móvel. Em Ghazal *et al.* (2018) imagens faciais de motoristas são utilizadas para identificar enquanto dirigem. Em Xiong *et al.* (2019) as técnicas de processamento de imagens faciais fazem parte de um sistema que identifica o nível de distração durante a jornada do motorista. Já em Vanitha, Vaidehi e Vasuhi (2018), uma nova abordagem é implementada para detectar atividades faciais baseando-se na diferença da intensidade das cores presentes nas imagens durante a análise do vídeo.

Rastreamento facial: os trabalhos deste grupo utilizam o rastreamento facial para auxiliar no reconhecimento facial e detecção facial. Em Elrefaei *et al.* (2017), um sistema de rastreamento facial foi desenvolvido para melhorar a técnica de reconhecimento facial utilizada durante a identificação de criminosos.

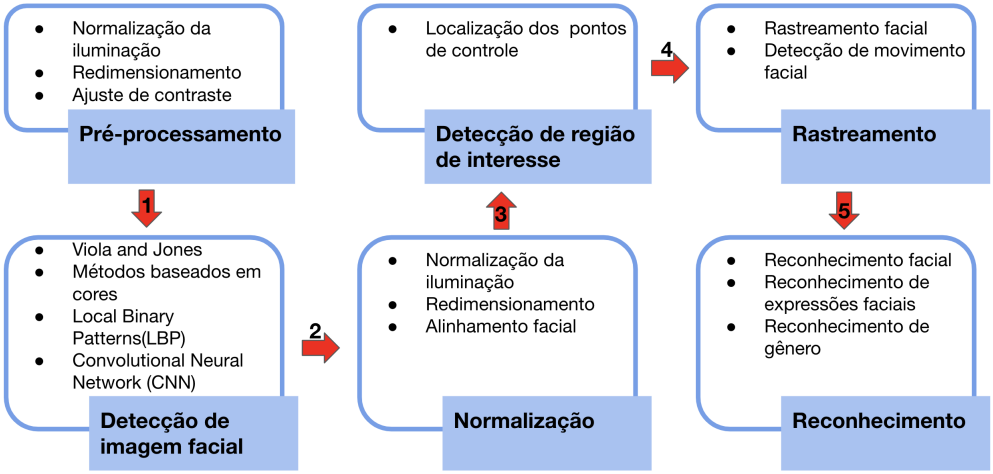
Outros: esta categoria representa 6,5% do total de artigos selecionados. Os objetivos presentes nesta categoria são variados, tais como reconhecimento de gênero (FOGGIA *et al.*, 2019), medição do pulso utilizando imagens faciais (HUANG; YANG; CHENG, 2016) e execução de técnicas com otimização de consumo de energia (MUKHERJEE *et al.*, 2019). Em Foggia *et al.* (2019), o artigo proposto em Viola e Jones (2001) foi utilizado para realizar a detecção facial em dispositivos móveis. Após esta etapa, duas técnicas de reconhecimento de gênero foram implementadas separadamente e suas acurácias foram comparadas. Em Huang, Yang e Cheng (2016) uma técnica foi desenvolvida para mensurar o pulso cardíaco do indivíduo. Por meio de técnicas de detecção facial e rastreamento de face, os autores utilizaram a intensidade das cores entre as imagens faciais capturadas para treinar um modelo de aprendizado de máquina, visando a identificar o pulso cardíaco do indivíduo de acordo com a variação da intensidade das cores nas regiões faciais relevantes. Em Mukherjee *et al.* (2019) uma aplicação cliente e servidor foi desenvolvida para realizar o reconhecimento facial. Segundo os autores, essa arquitetura possibilita avaliar o consumo de energia em um dispositivo móvel; em seguida, uma arquitetura cliente-servidor de reconhecimento facial com maior eficiência energética foi proposta. É importante destacar que a principal contribuição dos autores foi propor uma arquitetura energética eficiente para ser utilizada em dispositivos móveis, além disso, poderia ser utilizada para outras finalidades além do reconhecimento facial.

3.2.3 Técnicas utilizadas para o processamento facial em dispositivos móveis

A figura 3 apresenta o diagrama com as etapas envolvidas no processamento da imagem facial, construído a partir dos dados fornecidos pelos artigos incluídos. O processo está dividido em seis grupos de algoritmos que podem ser aplicados em sequência: **pré-processamento, detecção de imagem facial, normalização, detecção de região de interesse, rastreamento e reconhecimento**.

É importante ressaltar que nenhuma dessas etapas está presente em todos os artigos analisados, mesmo a detecção da imagem facial não é descrita por todos os artigos. Por exemplo, em Kalantarian *et al.* (2019) não é citado se um algoritmo de detecção de imagem facial foi aplicado, mas apenas uma técnica baseada em um conjunto de algoritmos para reconhecimento de emoção. Dessa forma, os autores poderiam ter utilizado uma técnica de detecção facial, porém isto não foi descrito no artigo.

Figura 3 – Diagrama sumarizado das técnicas de processamento facial utilizadas pelos artigos.



Fonte: Henrique Silva Bortoletti, 2022.

O conjunto de técnicas de **pré-processamento** engloba abordagens que preparam a imagem para as próximas etapas. Em alguns casos, uma sequência de técnicas é aplicada; por exemplo, em Zhang *et al.* (2018) os autores redimensionam as imagens para obter uma dimensão padrão e aplicam a equalização do histograma para normalização da iluminação e ajuste do contraste.

A **detecção da imagem facial** é a primeira etapa no processamento para 90% dos artigos. As técnicas mais citadas para esta tarefa são a técnica proposta em Viola e Jones (2001), métodos baseados na cor da imagem (VANITHA; VAIDEHI; VASUHI,

2018), padrões binários locais (FAUDZI; YAHYA, 2014) e métodos de aprendizado de máquina conforme descritos em Mahbub, Sarkar e Chellappa (2019).

A etapa de **normalização** consiste em algoritmos que podem ser aplicados após a detecção facial para ajustar a iluminação (CHEN; CHEN, 2012), alinhar e redimensionar as imagens faciais (LIU *et al.*, 2019). A diferença entre esta etapa e a etapa de pré-processamento, é a ordem em que as técnicas são aplicadas. Enquanto a primeira é aplicada antes que a imagem facial seja detectada, a normalização é aplicada após a etapa de detecção da imagem facial.

A etapa de **detecção de região de interesse** é responsável por identificar as regiões mais importantes da imagem facial a serem utilizados pelas técnicas nas próximas etapas. Em Zhang *et al.* (2018), por exemplo, os pontos de controle foram utilizados para calcular as distâncias euclidianas entre as regiões faciais, necessárias para o método de reconhecimento facial descrito pelos autores.

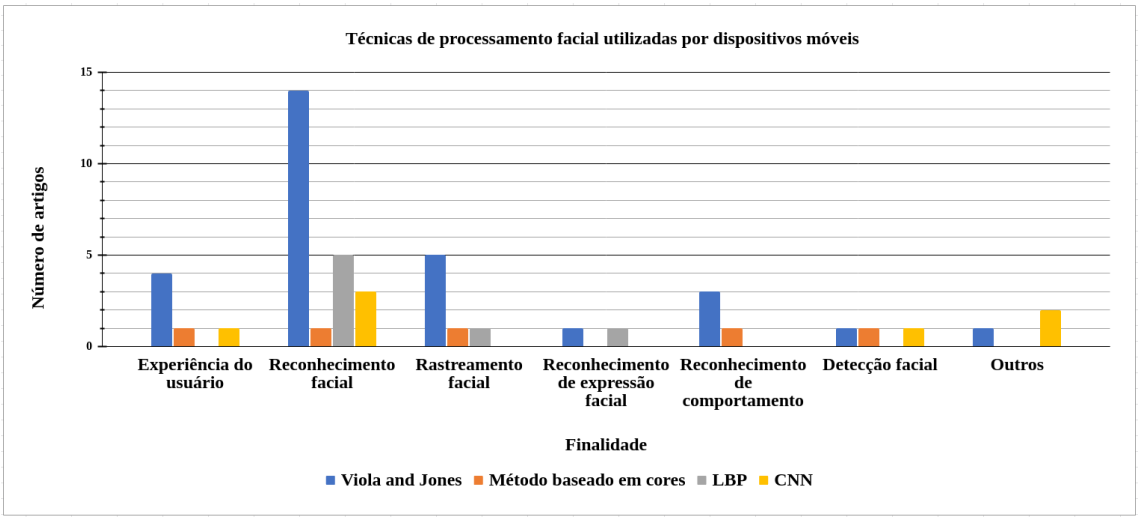
A etapa de **rastreamento** engloba técnicas para rastrear a imagem facial em vídeos (KODIROV *et al.*, 2013) ou para detectar se a imagem facial é de um usuário real. Em Maria *et al.* (2014) os autores citam o sistema de detecção de atividade facial do sistema operacional *Android*, mas desenvolvem uma nova técnica como parte de seu processo de reconhecimento facial para melhorar o custo computacional em relação à técnica oferecida por este sistema operacional *Android*.

Na última etapa do processamento da imagem facial, os artigos incluídos implementam a tarefa necessária para atingir seu objetivo, aqui denominada como etapa de **Reconhecimento**. As tarefas presentes neste bloco são reconhecimento de emoção facial (JEON *et al.*, 2016), reconhecimento facial (WANG *et al.*, 2017a) e reconhecimento de gênero (FOGGIA *et al.*, 2019).

A Figura 4 mostra os algoritmos mais utilizados para processamento de imagem facial em dispositivos móveis que são citados pelo menos por três categorias de objetivos diferentes. O algoritmo proposto em Viola e Jones (2001) foi a técnica mais citada para dispositivos móveis. De acordo com Wilhelm, Böhme e Gross (2004), este é um dos algoritmos mais eficientes encontrados na literatura para realizar a detecção facial. Seu desempenho e sua velocidade são decorrentes da eficiência em selecionar as regiões de interesse no início do processamento da imagem, o que diminui o esforço necessário para identificar a região facial (CHAUDHRY; CHANDRA, 2017). É um algoritmo apropriado computacionalmente para dispositivos móveis e possui alta precisão de detecção facial

(SMITH-CREASEY; ALBALOOSHI; RAJARAJAN, 2018). Em Smith, Albalooshi e Rajarajan (2018), os autores justificam sua escolha por esta técnica considerando sua robustez e o seu desempenho eficiente em relação ao tempo gasto, o que a torna adequada para aplicações em dispositivos móveis. No entanto, não é um bom algoritmo para realizar a detecção facial de imagens não frontais. Além disso, a iluminação ambiente pode prejudicar o desempenho da detecção, pois influencia as regiões de interesse que o algoritmo seleciona para determinar a posição facial (CHAUDHRY; CHANDRA, 2017; ZHANG; ZHANG, 2010).

Figura 4 – Algoritmos de processamento de imagem facial utilizados por dispositivos móveis.



Fonte: Henrique Silva Bortoletti, 2022.

É importante notar que o método apresentado em Tan e Triggs (2010) é uma possível solução para a iluminação ambiente não controlada. O método aplica uma sequência de técnicas de processamento de imagem para resolver o desafio da iluminação. O primeiro passo dessa sequência é a transformação não linear dos níveis de cinza por meio da correção *Gamma*, visando a reparar as características não lineares de exibição de um dispositivo (MCREYNOLDS; BLYTHE, 2005). A correção *Gamma* não remove a influência dos efeitos de sombreamento; assim, o próximo passo na sequência é a aplicação de um filtro baseado na diferença gaussiana dos *pixels* da imagem e, por fim, o contraste da imagem é equalizado por uma função definida pelos autores.

3.2.4 Tecnologias, plataformas e bibliotecas

Os dispositivos encontrados nos artigos selecionados foram categorizados da seguinte forma: *Smartphones* e *Tablets* (33%), *Arduino* (3%), outros dispositivos (30%) e dispositivo não citado (34%).

Ao menos 40% dos artigos selecionados utilizaram um dispositivo móvel, que possui o sistema operacional *Android* para a implementação dos métodos propostos, 11% utilizaram algum outro sistema operacional, como *IOS*, *Windows Mobile*, *Raspian*, *Symbian* ou *Blackberry OS*. Além disso, 49% dos artigos selecionados não citaram o sistema operacional utilizado.

A maior parte dos artigos não explica a motivação do sistema operacional escolhido. Em Chen e Chen (2012) os autores desenvolveram um sistema de rotação de tela automática baseado na posição do usuário. A biblioteca nativa de detecção facial oferecida pelo sistema operacional *IOS* foi utilizada, porém os autores destacam que seria possível implementar tais técnicas em um sistema operacional *Android*, pois ambos oferecem as mesmas funcionalidades necessárias para a implementação da técnica desenvolvida pelos autores.

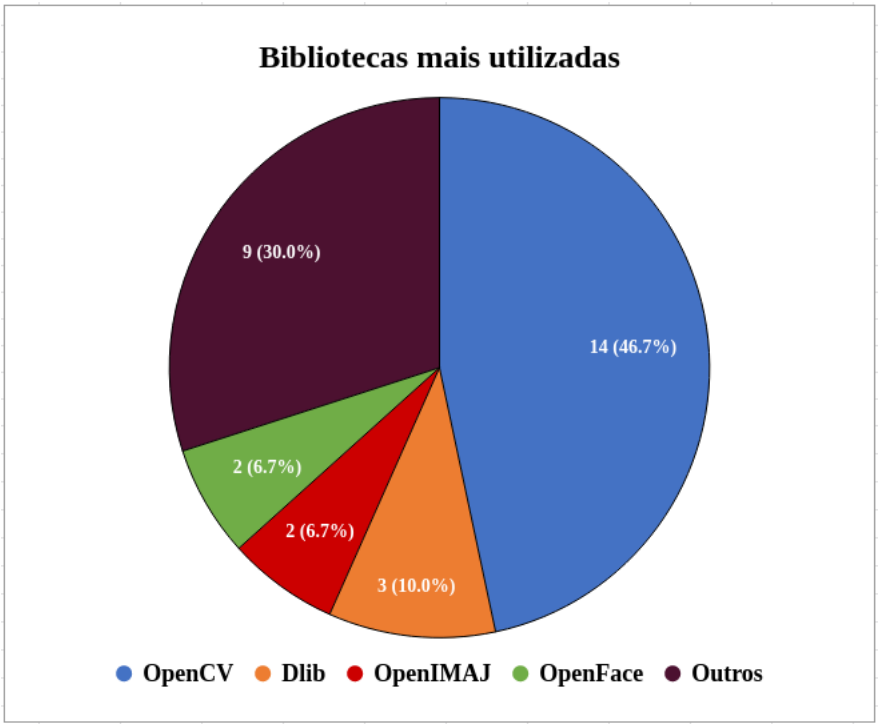
Na Figura 5 verifica-se que a biblioteca *OpenCV* foi a mais utilizada pelos artigos selecionados. Uma das possíveis causas pode ser sua flexibilidade em oferecer diversas opções de técnicas de processamento facial. Em Huang, Yang e Cheng (2016) a biblioteca é utilizada para a implementação do algoritmo desenvolvido em Viola e Jones (2001) para realizar a detecção facial. Já em Nambiappan e Datta (2018) a biblioteca é utilizada para realizar o reconhecimento facial por meio da técnica *Local Binary Patterns Histogram* (LBPH).

A biblioteca *Dlib* foi a segunda mais utilizada. Seu principal objetivo é oferecer a implementação de técnicas relacionadas ao aprendizado de máquina. Desenvolvida por meio da linguagem de programação *C++*, esta pode receber qualquer tipo de dado, por exemplo, vetores, imagens ou qualquer forma de estrutura de dados (KING, 2009). Já uma possibilidade de implementação para a linguagem de programação *Java* do algoritmo proposto em Viola e Jones (2001) pode ser por meio da biblioteca de análise multimídia *OpenIMAJ*, que foi utilizada em Smith, Albalooshi e Rajarajan (2018). Para realizar o reconhecimento facial pode-se utilizar a biblioteca *OpenFace*, que baseia-se no aprendizado

profundo para a extração das características utilizadas por máquinas de vetores de suporte, responsáveis pela classificação da imagem facial (WANG *et al.*, 2017a).

A categoria "Outros" representa todas as bibliotecas citadas em apenas um artigo, por exemplo, as bibliotecas nativas oferecidas pelos sistemas operacionais.

Figura 5 – Bibliotecas mais utilizadas de processamento facial utilizadas pelos artigos.



Fonte: Henrique Silva Bortoletti, 2022.

3.2.5 Bases de dados e métricas

Os conjuntos de dados utilizados para treinar ou avaliar os métodos são geralmente compostos por imagens faciais humanas (45%), vídeos (49%) ou ambos (6%). No entanto, cerca de 14% dos artigos (11 trabalhos) não descreveram o conjunto de dados de imagens privado utilizado. Em 33% dos artigos (26 trabalhos), os autores descrevem como o próprio conjunto de dados privado foi desenvolvido, 55% (43 trabalhos) utilizam um ou mais conjuntos de dados públicos e 4% (3 artigos) utilizam conjuntos de dados públicos e privados.

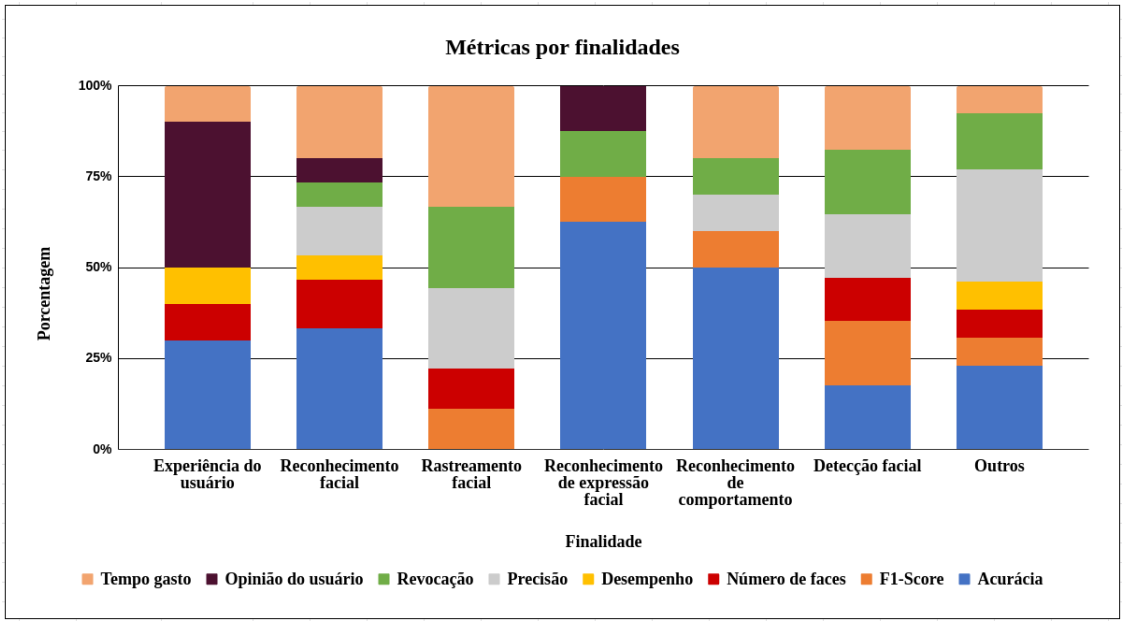
A base de dados pública mais utilizada foi a base *Labeled Faces in the Wild* (LFW) (HUANG *et al.*, 2007), citada por 18% dos artigos. Esta base representa uma referência para o reconhecimento facial contendo cerca de 13.000 imagens faciais coletadas da Web em ambientes não controlados. Cada imagem presente possui o nome correspondente

do indivíduo (HUANG *et al.*, 2007). Dez artigos que utilizaram este conjunto de dados também usaram algum outro banco de imagens, público ou privado, para avaliar ou treinar as técnicas implementadas (TIAN; ARBEL; CLARK, 2018; MAHBUB; SARKAR; CHELLAPPA, 2019).

Por exemplo, os autores em Washington *et al.* (2017) utilizaram um conjunto de dados público (CK+) para treinar o sistema de detecção facial, mas para medir a precisão do reconhecimento de emoção facial, os autores criaram um conjunto de dados privado com 3017 minutos de vídeos de 14 voluntários.

A Figura 6 apresenta as métricas utilizadas para avaliar os métodos apresentados pelos artigos que possuem os mesmos objetivos (Subseção 3.2.2). São incluídas todas as métricas que foram utilizadas por pelo menos três diferentes objetivos. A acurácia foi a métrica mais utilizada e apenas não está presente nos artigos cujos objetivos são o rastreamento facial ou o consumo de energia (incluídos na categoria "Outros" da Figura 6).

Figura 6 – Métricas mais utilizadas por objetivo dos artigos.



Fonte: Henrique Silva Bortoletti, 2022.

Geralmente a experiência do usuário não é medida apenas por métricas quantitativas, como acurácia, precisão ou revocação, mas também por técnicas qualitativas. Por exemplo, em Pham e Wang (2019) um questionário relacionado à atenção do usuário é aplicado para medir o envolvimento e identificar os sentimentos do usuário enquanto este assiste uma propaganda. O questionário também foi utilizado para medir o sucesso do método

de reconhecimento de emoção facial aplicado para identificar as emoções dos usuários, enquanto assistem ao comercial.

3.3 Considerações finais

Nesta revisão sistemática foram identificadas as técnicas mais utilizadas para cada objetivo específico, bem como as principais áreas da sociedade beneficiadas pelas técnicas de processamento de imagens faciais em dispositivos móveis. Adicionalmente, foi proposto um diagrama com a sequência de tarefas identificadas nos artigos incluídos com o objetivo de resumir as técnicas citadas pelos autores. Por fim, listamos e discutimos os *frameworks* mais utilizados para implementar os algoritmos identificados e as principais métricas e conjuntos de dados utilizados para avaliar os métodos descritos nos artigos. Os achados da revisão conduzida contribuíram para a proposta apresentada no Capítulo 4.

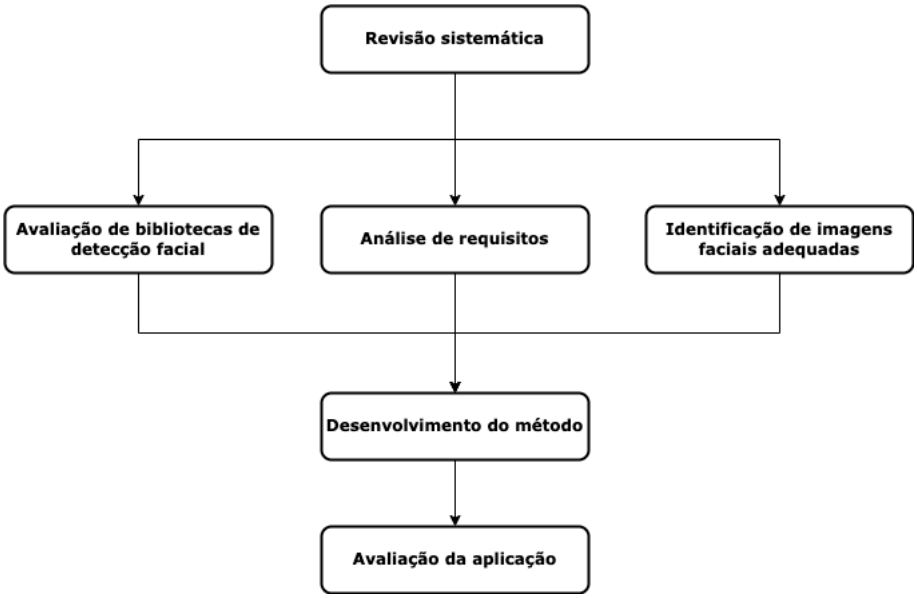
4 Materiais e métodos

Como anteriormente enunciado, este trabalho tem o objetivo de desenvolver e validar uma abordagem para adquirir e processar em tempo real imagens faciais em dispositivos móveis, garantindo a adequação para aplicação em sistemas de auxílio ao diagnóstico de TEA. Com isso, pretende-se contribuir para a democratização do acesso ao diagnóstico do TEA, por meio da aquisição de vídeos com as imagens faciais adequadas remotamente, os quais posteriormente serão submetidos a um sistema de auxílio ao diagnóstico. Neste capítulo são descritos os procedimentos e as etapas inerentes ao desenvolvimento desta pesquisa.

4.1 Etapas da pesquisa

Este projeto consiste em uma pesquisa experimental dividida em cinco etapas, conforme apresentado na Figura 7.

Figura 7 – Relação entre as diferentes etapas do projeto.



Fonte: Henrique Silva Bortoletti, 2022.

Na primeira fase foi realizada uma revisão sistemática que teve o objetivo de fornecer uma visão geral relacionada à área de aquisição e processamento de imagens faciais em dispositivos móveis. Foram analisadas quais as técnicas de processamento de imagem facial utilizadas em dispositivos móveis, as métricas utilizadas para avaliar os resultados obtidos

por meio destas técnicas, assim como as bibliotecas e tecnologias mais implementadas pelos autores. Os resultados da revisão foram apresentados no Capítulo 3.

A segunda fase é formada por dois experimentos, seguida de uma análise de requisitos. O primeiro teve o objetivo de identificar a influência dos algoritmos de detecção facial na identificação dos pontos de controle. O segundo experimento objetivou avaliar imagens faciais em um ambiente não controlado, visando definir regras que possibilitem classificar a adequação de uma imagem facial para a pesquisa. Os experimentos são apresentados, respectivamente, nas Seções 4.2 e 4.3. Ainda nesta fase se encontra a etapa de análise de requisitos descrita na Seção 4.4, responsável por definir os requisitos mínimos para o desenvolvimento da aplicação. Nesta etapa também são definidas as funções que cada usuário possui ao utilizar a aplicação.

A terceira fase é constituída pelo desenvolvimento do método proposto. Nesta etapa foram definidas quais as tecnologias a serem utilizadas durante o processo de desenvolvimento do método, assim como o processo de desenvolvimento, detalhado na Seção 4.5. A aplicação desenvolvida com suas respectivas funcionalidades é apresentada no Capítulo 5.

A quarta fase é responsável pela avaliação do método e pela análise dos resultados de tal avaliação. Para avaliar a aplicação desenvolvida foi realizado um experimento de usabilidade individual e em grupo. Para este experimento foi desenvolvido um questionário responsável por apresentar questões em relação à interação do usuário com a aplicação visando realizar as tarefas necessárias para a captura de imagens faciais adequadas. A avaliação da aplicação é apresentada na Seção 4.6. Os resultados relacionados à avaliação de usabilidade são apresentados e discutidos no Capítulo 6.

4.2 Avaliação de bibliotecas de detecção facial

Assumindo que as imagens faciais adequadas serão disponibilizadas para o treinamento ou teste de um sistema de auxílio ao diagnóstico, a próxima etapa é obter as distâncias entre os pontos de controle faciais, as quais serão utilizadas em um modelo de aprendizado de máquina. O modelo a ser usado no presente trabalho foi desenvolvido inicialmente em Pinheiros (2018). Visando aprimorar tal modelo, nesta Seção é apresentado

um experimento com o objetivo de identificar a influência da detecção facial na obtenção dos pontos de controle e, consequentemente, no desempenho do modelo.

4.2.1 Obtenção das distâncias entre os pontos de controle

O processo realizado para a obtenção das distâncias entre os pontos de controle possui quatro etapas: detecção facial, extração dos pontos de controle desejados, remoção de pontos de controle que dificultam a análise e cálculo das distâncias entre cada par de pontos de controle. Quatro algoritmos de detecção facial foram avaliados a fim de verificar sua influência no processo descrito:

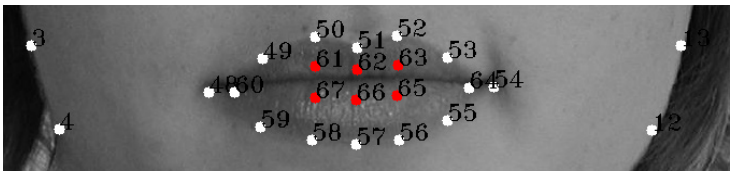
- *Convolutional Neural Network* (CNN) implementado pela biblioteca Dlib (KING, 2009);
- *Histogram of Oriented Gradients* (HOG) implementado pela biblioteca Dlib (KING, 2009);
- *Deep Neural Network* (DNN) implementado pela biblioteca OpenCV (BRADSKI, 2000);
- *Haar Cascade* implementado pela biblioteca OpenCV (BRADSKI, 2000).

Após a localização da região facial, os pontos de controle são identificados por um modelo matemático implementado pela biblioteca Dlib (KING, 2009). No experimento realizado foram identificados 68 pontos de controle. No caso de a região facial não ser localizada na etapa de detecção facial, a imagem não é utilizada na etapa de extração dos pontos de controle. Dos algoritmos citados anteriormente apenas o *Haar Cascade* não identificou a região facial em cinco imagens presentes na base utilizada, dessa forma tais imagens também não foram utilizadas durante a comparação dos algoritmos.

Durante o experimento foi identificado que em algumas imagens, alguns pontos de controle estavam posicionados exatamente nas mesmas coordenadas, todos estes sendo do contorno inferior do lábio superior e do contorno superior do lábio inferior (pontos “61”, “62”, “63”, “65”, “66”, “67” representados na Figura 8). Uma vez que as sobreposições foram observadas entre os pares de pontos (61,67), (62,66) e (63,65), optou-se por excluir das análises os pontos 61, 63 e 66, restando então 65 pontos de controle.

O conjunto de imagens utilizado neste experimento possui 43 imagens de crianças com TEA e 63 de crianças sem TEA. Cada imagem contém uma faixa de referência acima

Figura 8 – Pontos de controle da região da boca, que podem possuir as mesmas coordenadas.



Fonte: Henrique Silva Bortoletti, 2022.

da cabeça do indivíduo para normalizar as distâncias entre os pontos de controle. A faixa possui 50 centímetros de largura, cinco centímetros de altura e a distância do topo da cabeça do indivíduo para o centro da faixa não é relevante, pois apenas a largura e altura da faixa são utilizadas para a normalização das distâncias calculadas. A Figura 9 ilustra a posição e a dimensão da faixa em uma das imagens utilizadas. Dessa forma, a mesma técnica e a mesma base de dados descrita por Pinheiros (2018) foi aplicada para localizar a altura e a largura da faixa vermelha de referência.

Figura 9 – Faixa de de referência utilizada durante a coleta de imagens faciais pelo protocolo desenvolvido em (PINHEIROS, 2018).

Fonte: (PINHEIROS, 2018).

O cálculo da distância entre os pontos de controle é realizado por meio da equação 1, em que $d_n(p, q)$ é a distância euclidiana normalizada entre os pontos $p(x_1, y_1)$ e $q(x_2, y_2)$, w é a largura da faixa e h é a altura da faixa.

$$d_n(p,q) = \sqrt{[(x_1 - x_2)/w]^2 + [(y_1 - y_2)/h]^2} \tag{1}$$

Considerando os 65 pontos de controle, são calculadas as distâncias entre cada par de pontos, resultando em 2080 distâncias normalizadas pela faixa para cada item presente no conjunto de imagens.

4.2.2 Comparação dos algoritmos utilizados

Para avaliar o impacto dos algoritmos de detecção facial foram utilizados testes-T pareados (ROSS; WILLSON, 2017), para testar a hipótese de que a distância entre dois pontos de controle é influenciada pelos algoritmos de detecção facial.

A localização da face por meio de tais algoritmos pode ser exemplificada por meio da Figura 10, que possui a mesma imagem do indivíduo repetida quatro vezes. Cada uma destas possui um retângulo branco, que descreve o resultado do processo de detecção facial pelo algoritmo identificado no canto superior esquerdo da imagem. Os pontos brancos e vermelhos representam a localização dos pontos de controle após a identificação da face.

A diferença da localização da face entre os algoritmos pode não ser totalmente perceptível pela visão humana. Dessa forma, na Figura 10 determinados pontos foram evidenciados em vermelho, para facilitar a comparação entre as imagens.

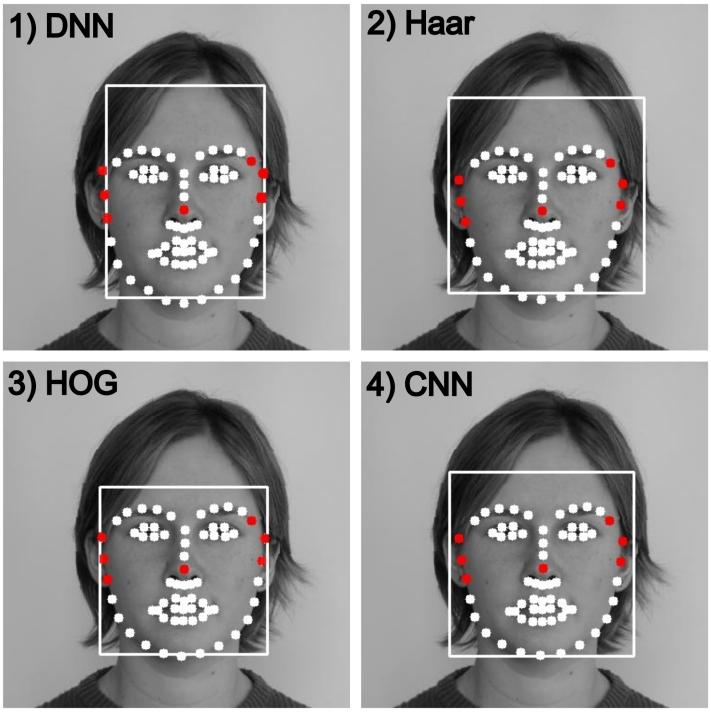
A avaliação da influência dos algoritmos de detecção facial no cálculo da distância entre os pontos de controle foi realizada por meio de testes-t pareados. Primeiramente, para cada algoritmo de detecção facial foram calculadas as 2080 distâncias conforme descrito na Subseção 4.2.1, para cada imagem presente no conjunto de imagens utilizado. Em seguida, para cada par de algoritmos de detecção facial, foram realizados 2080 testes-t pareados, um para cada distância. Cada teste visa avaliar se os dois algoritmos resultam em distâncias significativamente diferentes, pois avalia as seguintes hipóteses:

- H_0: $\Delta d = 0$ (não existem diferenças significativas para a distância calculada)
- H_1: $\Delta d \neq 0$ (existem diferenças significativas para a distância calculada)

As Hipóteses H_0 e H_1, respectivamente, representam se a diferença das distâncias calculadas entre os algoritmos é igual a zero ou diferente de zero.

Portanto, caso o p-valor do teste-t seja maior ou igual ao nível de significância estabelecido (neste trabalho foi utilizado o valor de 0.05), então a hipótese H_0 é verdadeira.

Figura 10 – Exemplo das diferentes localizações faciais pelas bibliotecas. O retângulo indica a área que o algoritmo identificou como sendo a face da pessoa e os pontos dentro do retângulo são os pontos de controle detectados pela biblioteca Dlib.



Fonte: Henrique Silva Bortoletti, 2022.

Porém, caso o p-valor do teste-t seja menor que o nível de significância estabelecido, então a hipótese H_1 é a verdadeira.

A Tabela 4 identifica a porcentagem das distâncias para cada par de algorítimos de detecção facial, que possuem diferenças significativas.

Tabela 4 – Número de distâncias entre pares de algoritmos que possuem diferenças significativas.

Algoritmo	CNN	HOG	DNN	Haar
CNN				
HOG	$98,5\%$			
DNN	$99,2\%$	$71,4\%$		
Haar	$99,8\%$	$21,4\%$	$72,7\%$	

Fonte: Henrique Silva Bortoletti, 2022.

Logo, pode-se observar na Tabela 4, uma diferença acima de 98% entre o algoritmo CNN e os demais algoritmos de detecção facial selecionados. Também é possível observar, que para os demais pares existe uma diferença acima de 71%, com exceção da comparação

entre os algoritmos Haar e HOG. Neste caso, existe uma diferença de 21% entre as distâncias calculadas.

Isso implica que para todos os casos existe uma diferença entre escolher um algoritmo de detecção facial ou escolher o outro, o que pode influenciar nos resultados do modelo de classificação utilizado.

4.3 Identificação de imagens faciais adequadas

Nesta seção são apresentados os resultados relacionados ao experimento que visa identificar as imagens adequadas em uma base de dados com diversas síndromes genéticas. Dessa forma o propósito deste experimento foi contribuir para a definição de regras que poderiam ser adotadas para utilizar em outras etapas deste projeto, visando a identificar uma imagem como adequada para uso no sistema de auxílio ao diagnóstico.

4.3.1 Base de dados

Neste experimento foi utilizada uma base de dados que possui 1573 imagens faciais com diferentes tipos de síndromes genéticas em diferentes ambientes (FERRY *et al.*, 2014). A Tabela 5 apresenta o número de imagens por síndrome presentes na base utilizada.

Tabela 5 – Número de imagens por síndrome.

Síndrome	Número de imagens
22q11	8
Turner	12
Marfan	18
Sotos	36
TreacherCollins	103
Progeria X	150
Fragile X	163
Down	197
Apert	200
Angelman	205
Williams	231
Cornelia De Lange (CDL)	250

Fonte: Henrique Silva Bortoletti, 2022.

Para classificar se uma imagem é adequada foi realizado um experimento, que consiste em responder quatro perguntas sobre as imagens presentes na base descrita:

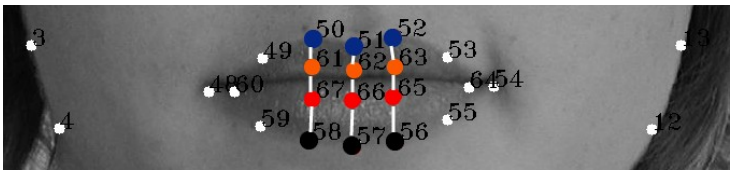
- a boca está aberta?
- a imagem facial é frontal?
- a expressão facial é neutra?
- existe algum tipo de oclusão na imagem?

Para responder tais perguntas, primeiramente foram adotadas algumas regras para identificar se a boca estava aberta ou fechada, assim como identificar se a imagem era frontal.

A Figura 11 apresenta os pontos utilizados para determinar se a boca estava aberta ou fechada. Foi calculada a média das distâncias entre cada ponto azul e cada ponto laranja conectados por meio de uma reta branca; essa média representa a largura do lábio superior. Também foi utilizada a largura do lábio inferior, que é igual à média das distâncias entre cada ponto vermelho e cada ponto preto conectados por meio de uma reta branca.

Após identificar a largura do lábio inferior e do lábio superior foi calculada a largura da região entre os lábios utilizando a média das distâncias entre cada ponto laranja conectado a cada ponto vermelho por meio das retas brancas. Se a largura da região entre os lábios for 50% maior que a largura do lábio superior ou inferior, então a boca estará aberta, caso contrário estará fechada.

Figura 11 – Pontos de controle da boca utilizados para determinar sua abertura.



Fonte: Pontos utilizados para identificar se a boca está fechada ou aberta(SUNG, 1996).

Para determinar se a imagem facial é frontal foi utilizada a proporção entre a distância do olho esquerdo ao nariz e a distância do olho direito ao nariz. Caso o indivíduo estivesse com a face levemente voltada para a direção do lado esquerdo ou direito, a proporção entre as distâncias seria menor que 0,8 ou maior que 1,3. Logo, caso a face esteja levemente voltada em direção para a esquerda ou para a direita, a imagem não será classificada como frontal.

Quando na imagem foi identificado que a boca estava aberta ou que a face não estava posicionada frontalmente, a imagem foi automaticamente classificada como não adequada.

Porém, se a boca estava fechada e a face estava posicionada frontalmente, então a imagem seria avaliada por dois pesquisadores para identificar algum tipo de oclusão presente na imagem ou se a expressão não era neutra.

Logo, a imagem foi classificada como adequada apenas quando a boca estivesse fechada, a imagem facial fosse frontal, a expressão facial fosse neutra e não houvesse algum tipo de oclusão na imagem.

4.3.2 Resultados do experimento

Ao final do experimento apenas 82 (5%) imagens foram identificadas como adequadas, 1417 (90%) foram descartadas por causa do alinhamento da face ou da abertura da boca e 74 (5%) imagens foram classificadas como não adequadas pelos pesquisadores, pois algum tipo de oclusão estava presente na imagem ou não possuíam a expressão neutra. Dessa forma, esse experimento foi importante para identificar possíveis heurísticas que podem ser utilizadas durante a proposta de pesquisa para identificar as características que tornam uma imagem facial adequada para ser submetida ao sistema de auxílio ao diagnóstico.

Também foi possível identificar que todos os indivíduos com a síndrome 22q11 possuíam a boca aberta e todos os indivíduos com a síndrome de Marfan também possuíam a boca aberta ou a imagem não era frontal. Assim, caso a proposta apresentada neste documento seja adaptada para outras síndromes, algumas características faciais para classificar uma imagem como adequada precisam ser adaptadas.

4.4 Análise de requisitos

Os requisitos identificados durante a construção de um sistema são normalmente classificados em funcionais e não funcionais. Enquanto os requisitos funcionais são responsáveis por descrever as funcionalidades do sistema, os requisitos não funcionais são responsáveis por descrever as propriedades e restrições do sistema (KURTANOVIć; MAALEJ, 2017).

Os distintos requisitos relacionados à construção da aplicação foram definidos durante diversas reuniões entre o PROTEA e o LApIS. Inicialmente os principais requisitos propostos pelo PROTEA foram relacionados a quem deveria ter acesso à aplicação, aos dados a serem armazenados dos pacientes e usuários. Já os requisitos propostos pelo LApIS estavam relacionados às principais funcionalidades presentes na aplicação.

Dessa forma, os seguintes requisitos funcionais foram estabelecidos:

- **Cadastro do usuário**: o sistema deve permitir que qualquer pessoa possa realizar o seu cadastro por meio da aplicação, com os seguintes dados: nome, cidade, estado, e-mail, telefone e senha do usuário;
- **Avaliação do cadastro do usuário**: o sistema deve enviar uma notificação para os responsáveis pelo projeto para avaliar se o usuário cadastrado possui permissão para utilizar a aplicação;
- **Troca de senha**: o sistema deve permitir ao usuário trocar a sua senha cadastrada;
- **Cadastro de novo paciente**: o sistema deve permitir ao usuário cadastrar um novo paciente, com os seguintes dados: nome, sexo e data de nascimento;
- **Consulta de pacientes cadastrados**: o sistema deve exibir apenas os pacientes cadastrados pelo usuário;
- **Captura de vídeo**: o sistema deve possibilitar registrar o peso e a altura do paciente durante a captura de vídeos faciais dos indivíduos;
- **Identificação de imagens faciais adequadas**: o sistema deve identificar as imagens faciais adequadas presentes no vídeo capturado;
- **Envio de dados**: o sistema deve ser capaz de enviar os dados da aplicação para o servidor;
- **Armazenamento de vídeos e imagens**: o sistema deve armazenar os vídeos capturados e as imagens faciais adequadas;
- **Configuração de parâmetros**: o sistema deve ser capaz de possibilitar ao usuário realizar a configuração dos parâmetros do sistema.

Os seguintes requisitos não funcionais foram estabelecidos:

- **Disponibilidade**: necessária a conexão do sistema à rede mundial de computadores para a troca de dados e a disponibilidade para o acesso de qualquer pessoa autorizada em território nacional ou estrangeiro;

- **Base de dados**: necessária a implementação de um banco de dados relacional para o armazenamento das informações relacionadas às capturas de vídeos, às imagens adequadas presentes nos vídeos e aos dados cadastrados pelos usuários;
- **Segurança dos dados do usuário**: necessária a criptografia da senha cadastrada pelo usuário na base de dados;
- **Segurança durante a transmissão dos dados**: necessária a criptografia durante a transmissão dos dados entre a aplicação e o servidor por meio do protocolo HTTPS;
- **Plataforma do aplicativo**: necessário utilizar a plataforma *Android* e acessível para qualquer pessoa;
- **Parâmetros de captura**: necessário estabelecer um tempo padrão para a duração do vídeo e um número mínimo de imagens adequadas encontradas durante a captura;
- **Protocolo de transmissão de dados**: necessário transmitir os dados do aplicativo para o servidor por meio de requisições REST;
- **Usuário temporário**: necessário o usuário receber uma senha temporária e estar na categoria de "usuário temporário"durante a troca de senha;
- **Tempo limite da sessão**: necessário estabelecer um tempo máximo de sessão, após o qual o usuário deve autenticar-se novamente. Neste projeto o tempo máximo foi estabelecido em cinco horas.

O resultado do desenvolvimento da aplicação utilizando os requisitos funcionais e não funcionais é apresentado no Capítulo 5.

4.5 Desenvolvimento do método

Nesta seção as tecnologias utilizadas para a implementação da proposta são apresentadas, a arquitetura do desenvolvimento da proposta e os módulos que compõem a estrutura para o desenvolvimento da proposta também são apresentados. A descrição das funcionalidades da aplicação resultante da abordagem definida nesta seção é apresentada no Capítulo 5.

4.5.1 Tecnologias

O Capítulo 3 apresenta os resultados e discussões da Revisão Bibliográfica Sistemática, que teve como um dos objetivos identificar no estado da arte as principais tecnologias utilizadas para a implementação de técnicas de processamento de imagens faciais em dispositivos móveis.

As principais tecnologias utilizadas durante a implementação da proposta de projeto, descrita no Capítulo 5 são:

- *Linguagem Java*: a linguagem de programação Java (BERTOLISSI; PREECE, 1998) foi utilizada para o desenvolvimento da aplicação presente no dispositivo móvel, assim como na implementação do servidor. Os motivos para a utilização desta linguagem são: vasta documentação encontrada, compatibilidade com todas as tecnologias utilizadas no projeto e gratuidade;
- *Sistema Operacional Android*: o sistema operacional, conforme mencionado anteriormente, foi selecionado para ser a plataforma de desenvolvimento da aplicação presente nos dispositivos móveis (ANDROID, 2022);
- *Biblioteca OpenCV*: foi a biblioteca mais citada durante a revisão bibliográfica sistemática, pois possui uma alta variedade de algoritmos de processamento de imagem facial (BRADSKI, 2000). Durante a implementação foi utilizada para realizar a extração das imagens adequadas presentes nos vídeos enviados ao servidor;
- *Biblioteca ML Kit*: arcabouço de algoritmos de aprendizado de máquina desenvolvido e disponibilizado pelo *Google*. Apresenta uma grande variedade de funcionalidades para aplicações presentes em dispositivos móveis (HETTIARACHCHI, 2021). Esta biblioteca foi utilizada para a detecção da face, a extração dos pontos de controle, identificação dos ângulos de *Euler* (PANCHAL; SHAH, 2013), cálculo da probabilidade de uma pessoa estar sorrindo e finalmente para a identificação dos olhos abertos ou fechados. Para a implementação deste projeto foi utilizada a versão *offline* da biblioteca, que possui apenas as funcionalidades relacionadas a detecção facial; desta maneira, quando o usuário instalar a aplicação em seu dispositivo móvel também estará instalando automaticamente todo o pacote relacionado a esta biblioteca;
- *Biblioteca AndroidX*: biblioteca de suporte que fornece compatibilidade com versões anteriores do *Android*, assim como oferece de maneira mais simplificada alguns

recursos do próprio sistema operacional, por exemplo, as funcionalidades da câmera frontal e traseira (ANDROIDX, 2022).

- *Retrofit*: uma API de código livre seguindo o padrão REST, provendo um padrão para transmissão de dados entre aplicação e servidor (RETROFIT, 2022). Utilizada para realizar as transmissões de dados do dispositivo móvel para o servidor;
- *Biblioteca Spring Framework*: biblioteca de código livre que providencia uma infraestrutura para a criação de aplicações *Java* (SPRING, 2022). Foi utilizada para o desenvolvimento do servidor; dessa forma as principais funcionalidades utilizadas foram: a integração com o banco de dados relacional, o gerenciador de envio de *e-mail*, o servidor embutido da biblioteca, os recursos de autenticação e autorização fornecidos para a segurança das APIs desenvolvidas;
- *Banco de dados MySql*: Sistema Gerenciador de Banco de Dados relacional (FILIP; ČEGAN, 2020) utilizado para o armazenamento de dados conforme descrito na Subseção 4.5.9;

É importante ressaltar que a escolha de cada tecnologia levou em consideração dois principais fatores: a tecnologia ainda é constantemente atualizada e existe documentação suficiente para a implementação das funcionalidades desejadas. Além disso, foi avaliado se as bibliotecas poderiam ser utilizadas por meio da linguagem de programação *Java*.

4.5.2 Arquitetura da proposta

Como enunciado no Capítulo 1, esta proposta faz parte de um projeto mais abrangente que visa construir um sistema de auxílio ao diagnóstico do TEA em que, uma das vertentes, visa classificar indivíduos com base no processamento de medidas antropométricas extraídas de imagens faciais. Na abordagem proposta neste trabalho, uma aplicação para dispositivos móveis é responsável por capturar um vídeo e verificar a adequação dos *frames* para uso diagnóstico. Uma vez identificados *frames* adequados, a aplicação deve enviá-los para o servidor que possui um modelo com base em Inteligência Artificial encarregado de realizar a previsão diagnóstica (classificação). Este modelo foi desenvolvido inicialmente por Pinheiros (2018) e está sendo aprimorado.

A Figura 12 apresenta a arquitetura da proposta de captura das imagens. Os símbolos dentro dos retângulos com descrições abaixo de cada um, representam os módulos

desenvolvidos que serão descritos ao longo desta seção. Já os retângulos representam os diferentes componentes que interagem entre si durante a captura do vídeo com as imagens faciais adequadas. Vale ressaltar que os módulos em cor verde foram implementados inicialmente em Pinheiros (2018). O módulo de transmissão de vídeo e dados não está dentro dos retângulos porque está presente tanto no componente do dispositivo móvel quanto no componente do servidor.

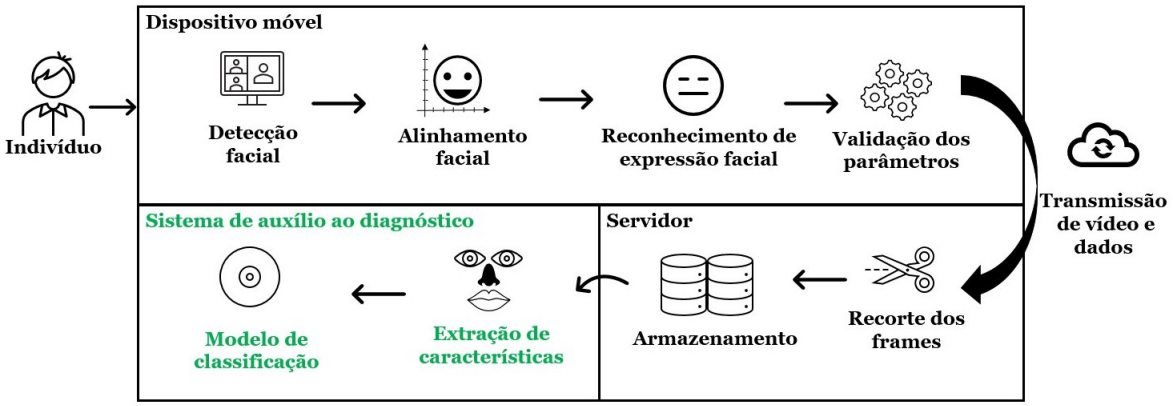
A aplicação no dispositivo móvel com sistema operacional *Android* consiste de um conjunto de algoritmos de processamento de imagem, responsáveis por determinar as melhores imagens faciais capturadas. Deste modo, o indivíduo deve se posicionar em frente à câmera do aparelho, enquanto um vídeo é feito pelo usuário por meio da aplicação para capturar as imagens adequadas.

Após identificar o número pré-estabelecido das primeiras melhores imagens faciais adequadas no vídeo, a posição dos *frames* com o vídeo e mais algumas informações relacionadas à captura são enviadas para o servidor. Este é responsável por extrair e armazenar as imagens adequadas do vídeo por meio da posição dos *frames*, assim como armazenar o vídeo e todas as informações enviadas pelo dispositivo móvel em um banco de dados relacional. É importante destacar, que o modelo de classificação e a extração de características não fazem parte da presente proposta, pois foram desenvolvidos inicialmente em Pinheiros (2018).

Nas subseções a seguir são descritos todos os módulos presentes na Figura 12.

Figura 12 – Arquitetura do processo de captura de imagens adequadas pela aplicação.O módulo destacado na cor verde é responsável pelo

auxílio ao diagnóstico, inicialmente desenvolvidos em Pinheiros (2018) e aperfeiçoado no presente trabalho.



Fonte: Henrique Silva Bortoletti, 2022.

4.5.3 Detecção facial

Essa é a primeira etapa da implementação da arquitetura proposta, iniciada no momento em que o vídeo do indivíduo começa a ser gravado. Este módulo é responsável por realizar a detecção das imagens faciais frontais do indivíduo.

Para isso, primeiramente o indivíduo é posicionado frontalmente ao dispositivo móvel. Em seguida, o usuário responsável pela captura das imagens do indivíduo pressiona um botão na interface gráfica para realizar a captura do vídeo. Dessa forma, a cada intervalo de tempo em milissegundos é realizada a detecção facial do indivíduo; após testes empíricos este intervalo foi estabelecido em 500 milissegundos, ou seja, a cada segundo de vídeo duas imagens serão avaliadas. Caso uma imagem facial seja detectada durante a captura do vídeo, esta será enviada para o módulo de alinhamento facial descrito na Subseção 4.5.4.

É importante ressaltar que inicialmente esperava-se avaliar todos os *frames* durante a captura em tempo real, porém durante a implementação deste módulo foi identificado que as técnicas implementadas podem sobrecarregar os dispositivos móveis causando lentidão e, em alguns casos, a aplicação pode ser encerrada inesperadamente. Por isso, estabeleceu-se avaliações periódicas, conforme relatado.

Outro motivo que inviabiliza a detecção facial de todos os *frames* em tempo real é a soma do tempo gasto pelas técnicas aplicadas, pois são necessários em média 68 milissegundos para realizar apenas a detecção facial de uma única imagem (HETTIARACHCHI, 2021). Logo, como são capturados 30 *frames* por segundo, então seria necessário que a técnica fosse capaz de realizar a detecção facial em menos de 34 milissegundos.

Uma possível solução para este desafio é a aplicação de tais técnicas para realizar a detecção facial após a captura do vídeo. Porém, o tempo de espera do usuário para realizar uma captura aumentaria consideravelmente. Como o objetivo deste projeto é a avaliação em tempo real das imagens faciais, então essa solução foi descartada.

Por fim, é importante destacar que a detecção de imagens faciais continua até que a captura do vídeo seja finalizada ao alcançar o tempo pré-estabelecido de duração. Se ao fim da gravação o número mínimo pré-estabelecido de imagens faciais adequadas for encontrado, então tais informações serão enviadas ao servidor. Porém, se o número mínimo de imagens faciais adequadas não for atingido, então as informações da captura

são descartadas e uma nova sessão de seleção das imagens será automaticamente iniciada. O tempo de duração do vídeo e o número mínimo de imagens faciais adequadas foram pré-estabelecidos durante a análise de requisitos conforme descrito na Seção 4.4.

É importante destacar que a detecção facial é realizada pela biblioteca *ML Kit* (GOOGLE, 2022), que após identificar a face do indivíduo retorna uma série de informações relacionadas à imagem detectada. Por exemplo, são retornados os pontos de controle (FARKAS, 1982), a probabilidade dos olhos estarem abertos ou fechados, a probabilidade da pessoa estar sorrindo, o contorno da face e os ângulos de *Euler* (PANCHAL; SHAH, 2013).

Dessa forma, durante o desenvolvimento da proposta notou-se que não seria necessário lidar com o desafio da iluminação, pois caso a biblioteca identifique a face do indivíduo e retorne todas as informações relacionadas ao rosto, então isso implica que a iluminação ambiente é minimamente adequada para a detecção facial. Logo, o tratamento da iluminação não foi incluído à proposta, pois a biblioteca é responsável por definir qual iluminação é ideal para a detecção facial.

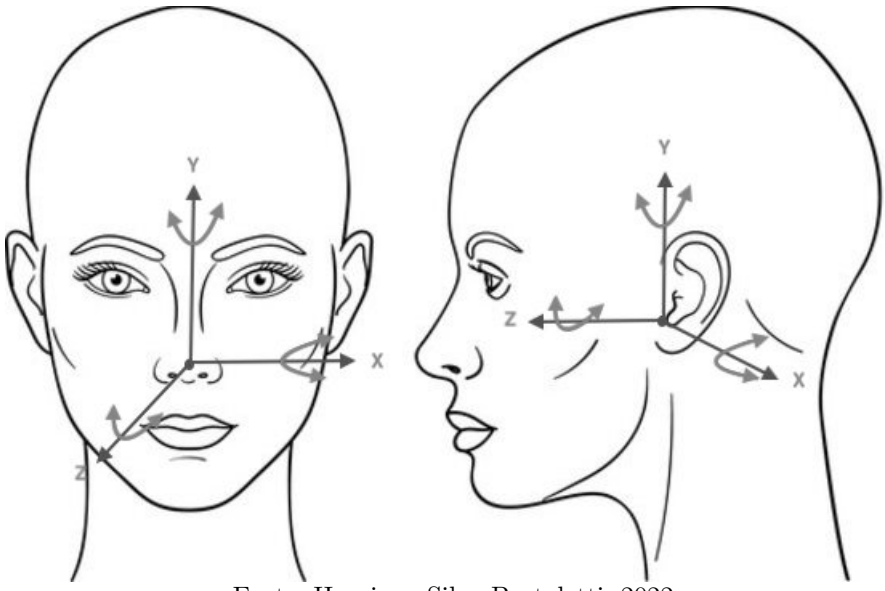
4.5.4 Alinhamento facial

Este módulo é responsável por avaliar o alinhamento facial do indivíduo identificando os valores dos ângulos de *Euler*, em conjunto com uma heurística baseada na diferença da altura entre os olhos e orelhas. Caso o alinhamento seja adequado, então a imagem será enviada para o módulo de **reconhecimento de expressão facial** descrito na Seção 4.5.5.

Ambas as avaliações são responsáveis por definir se a face do indivíduo é frontal e segue o padrão pré-estabelecido. Dessa forma, a detecção dos pontos de controle em diferentes indivíduos não será impactada pelo alinhamento da face, pois espera-se que todas as imagens faciais adequadas sejam frontais e possuam um alinhamento semelhante.

A avaliação do alinhamento da face foi realizada verificando-se os valores dos ângulos de *Euler*, que possibilitam identificar a rotação da cabeça acima do pescoço utilizando os valores de acordo com o posicionamento do objeto sob os eixos X, YeZ (CHEN; YU; ODOBEZ, 2015), conforme exemplificados por meio de duas diferentes perspectivas na Figura 13. Dessa forma, para avaliar o alinhamento facial são definidos intervalos de aceitação para cada ângulo.

Figura 13 – Perspectiva frontal e lateral dos ângulos de Euler sob os eixos $X, Y e Z$.



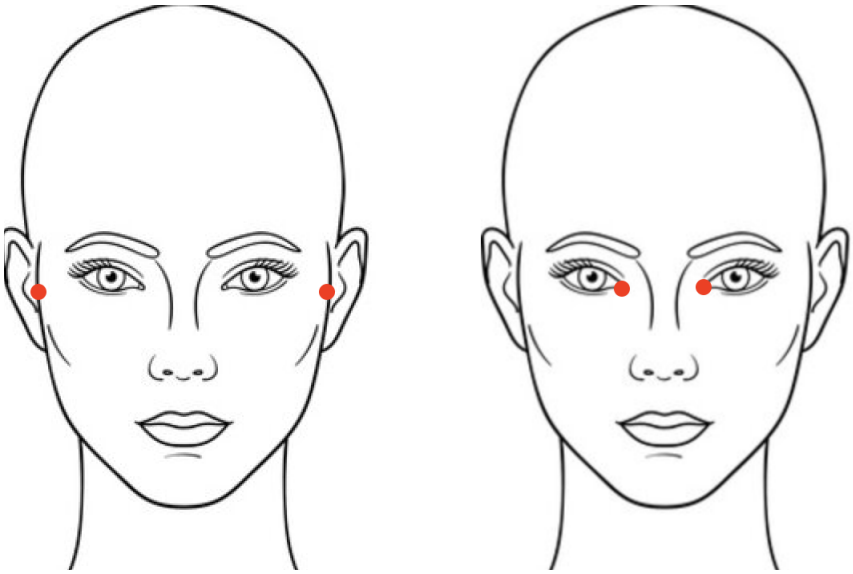
Fonte: Henrique Silva Bortoletti, 2022.

A Figura 13 exemplifica o alinhamento facial caso os ângulos de *Euler* sob os eixos X, Y e Z seja de aproximadamente 0°. Tais valores são calculados pela biblioteca *ML Kit* (GOOGLE, 2022) por meio dos pontos de controle detectados. Durante o desenvolvimento do método foi definida uma tolerância de 7° de aceitação para todos os ângulos; logo, os valores aceitáveis estão no intervalo [-7°, 7°] simultaneamente para os três ângulos.

Após esta primeira etapa é necessário avaliar a altura entre as orelhas e olhos do indivíduo. Para isso são utilizadas as distâncias entre os pontos em vermelho da Figura 14. Portanto, a distância vertical dos pontos de controle *landmarks* entre a orelha esquerda e a orelha direita, assim como a distância vertical dos pontos de controle entre o olho esquerdo e o olho direito são calculadas. Após obter essas distâncias, é necessário verificar se tais valores em *pixel* é menor que o valor pré-estabelecido de 100 *pixels* para tais medidas. Tal valor foi definido após testes realizados nas mesmas imagens utilizadas por (PINHEIROS, 2018). Caso o valor seja menor, então a posição da face é adequada, caso contrário o alinhamento da imagem não é considerado adequado.

Consequentemente, para uma imagem facial ter o alinhamento facial adequado é necessário avaliar os valores dos ângulos de *Euler* e avaliar os valores entre as alturas dos olhos e orelhas, conforme descrito. Se uma das avaliações não for adequada, então a imagem facial será descartada.

Figura 14 – Pontos de Landmark identificando pontos médios da orelha e cantos internos dos olhos.



Fonte: Henrique Silva Bortoletti, 2022.

4.5.5 Reconhecimento de expressão facial

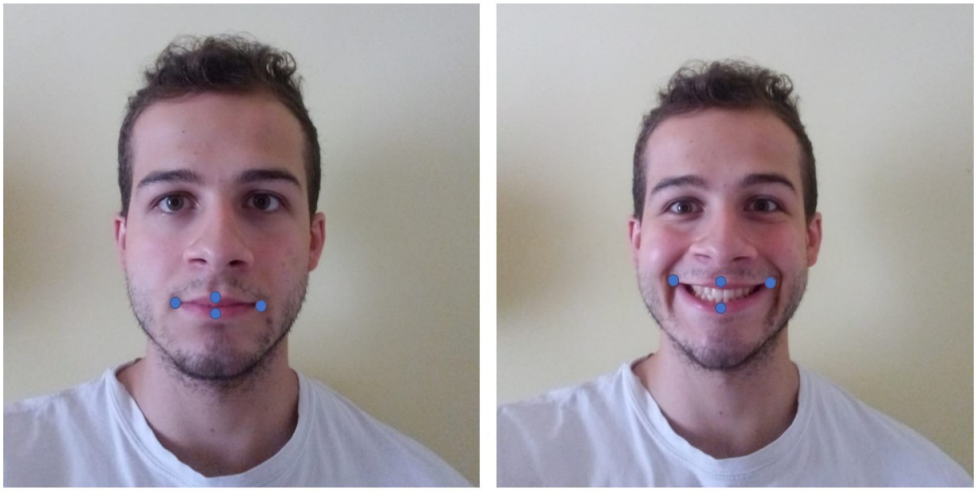
Este módulo é responsável por verificar se a expressão representada na imagem facial que está sendo adquirida é neutra. Caso a expressão facial seja adequada, então a imagem será enviada para o módulo de **validação dos parâmetros** (Seção 4.5.6); se não for adequada então a imagem será descartada.

A seleção da imagem facial com expressão neutra é necessária para que as regiões de interesse, por exemplo, boca, olhos, nariz e sobrancelhas estejam posicionadas da maneira mais semelhante possível entre diferentes indivíduos, pois a localização dos pontos de controle pode influenciar no cálculo das distâncias entre eles.

Por exemplo, na Figura 15 os pontos de controle azuis estão em localizações diferentes, pois o indivíduo à direita está sorrindo, enquanto o mesmo indivíduo à esquerda está com a boca fechada (expressão neutra). Esse posicionamento influencia no modelo de classificação, que espera receber imagens de indivíduos que possuem as regiões de interesse localizadas em posições referentes à mesma expressão facial.

Para realizar o reconhecimento da expressão facial neutra são realizadas duas validações: a primeira detecta a probabilidade do indivíduo estar sorrindo em uma imagem facial e a segunda identifica a probabilidade de ambos os olhos estarem abertos. Ambas as probabilidades são apresentadas pela biblioteca ML Kit (GOOGLE, 2022) que, após realizar a detecção facial, retorna um valor entre 0 e 1 indicando a probabilidade de que o

Figura 15 – Pontos de controle em azul identificam a parte central dos lábios superior e inferior do indivíduo.



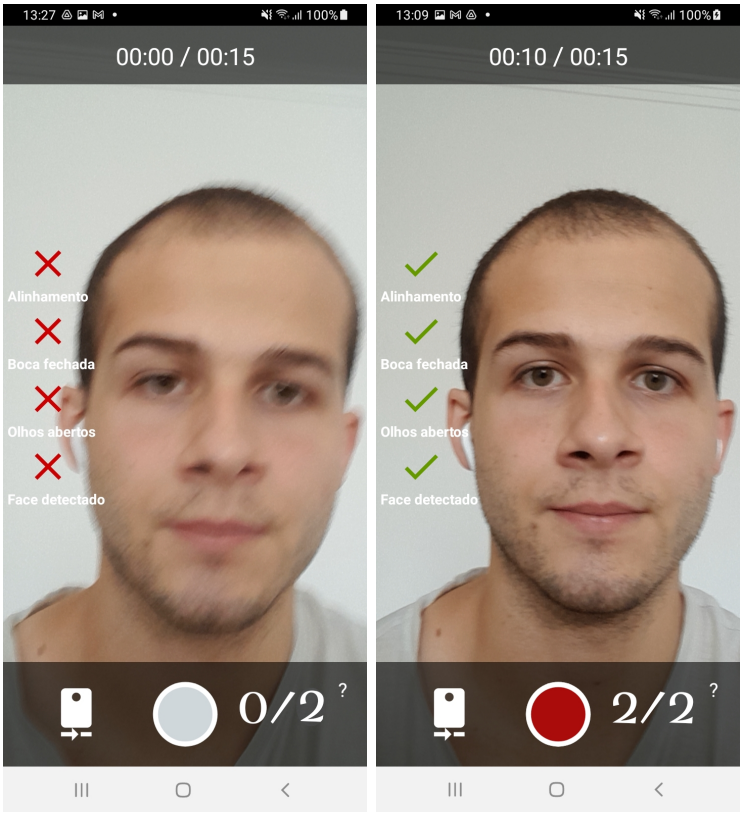
Fonte: Henrique Silva Bortoletti, 2022.

indivíduo esteja sorrindo. Quanto mais próximo o valor está do limite superior, maior é a probabilidade de se encontrar um sorriso. A mesma lógica se aplica para o reconhecimento de os olhos estarem abertos ou fechados. Dessa forma, um valor entre 0 e 1 é retornado para cada olho indicando a probabilidade de o olho estar aberto; quanto mais próximo o valor está do limite superior, maior é a probabilidade do olho estar aberto.

Caso a probabilidade do indivíduo estar sorrindo seja maior que a porcentagem limite ou se pelo menos um dos olhos apresentar uma probabilidade abaixo do esperado para ser considerado um olho aberto, então a imagem será descartada. De forma análoga, se a porcentagem da pessoa estiver sorrindo for inferior ao limite e a probabilidade de ambos os olhos estarem abertos for igual ou maior que o esperado, a imagem e os dados das outras validações provenientes dos módulos exemplificados anteriormente serão enviados para o módulo de validação dos parâmetros (Seção 4.5.6).

4.5.6 Validação dos parâmetros

Este módulo é responsável por validar quais os módulos que estão sendo utilizados durante a captura das imagens faciais, de informar ao usuário o estado da captura, como também de enviar para o módulo de transmissão (Seção 4.5.7) as informações que serão enviadas para o servidor.



(a) (b)

Figura 16 – Possíveis situações de falha e sucesso durante a captura das imagens faciais (a) Falha em todas as validações durante a captura. (b) Sucesso em todas as validações durante a captura.

Fonte: Henrique Silva Bortoletti, 2022.

Tal módulo pode ser descrito como um orquestrador da aplicação no dispositivo móvel, pois define quais módulos serão utilizados e quais os parâmetros presentes durante a captura.

É parte da proposta oferecer aos pesquisadores a opção de configurar diversos parâmetros utilizados durante a captura e de não utilizar todos os módulos descritos na Figura 12, com exceção do módulo de detecção facial. Por exemplo, para alguns casos pode ser difícil manter o indivíduo com o alinhamento facial adequado conforme descrito na Subseção 4.5.4, dessa forma é possível configurar a aplicação para ignorar essa validação ou alterar os valores necessários para a validação dos ângulos de *Euler*. Também é possível configurar número de imagens ideais necessárias, conforme descrito na Seção 5.6.

Após identificar quais validações devem ser utilizadas, as informações necessárias serão enviadas para este módulo, que informa ao usuário se houve falha em alguma etapa anterior. A Figura 16**A** representa um caso em que nenhuma validação ocorreu com sucesso, já a Figura 16**B** representa o caso em que todas as validações ocorreram com sucesso.

Se as validações forem bem sucedidas, então as seguintes informações serão enviadas para o módulo de transmissão (Seção 4.5.7): o vídeo, a posição das imagens faciais adequadas, a identificação do indivíduo, o seu peso e a sua altura.

4.5.7 Transmissão de imagem e dados

Este módulo é responsável por estabelecer a comunicação entre o dispositivo móvel e o servidor. A aplicação cliente presente no dispositivo móvel é responsável pela transmissão do vídeo e dos dados referentes às imagens faciais adequadas, que são transmitidas por meio da rede mundial de computadores. Já o servidor é responsável por disponibilizar inúmeras *Application Programming Interfaces* (APIs) que, por meio da arquitetura *Representational State Transfer* (REST) utilizando o protocolo *Hyper Text Transfer Protocol Secure*(HTTPS), estabelecem padrões de comunicação que serão utilizados por ambas as partes.

A escolha da arquitetura REST é baseada na premissa que apenas a aplicação cliente é responsável por enviar os dados e não existe um fluxo intenso de troca de informações entre um dispositivo móvel e o servidor. A transmissão dos dados não é intensa, pois todos os dados necessários para o armazenamento das imagens adequadas são enviados por meio de uma única requisição REST utilizando o protocolo HTTPS, que garante a privacidade e segurança dos dados durante a comunicação (CALLEGATI; CERRONI; RAMILLI, 2009).

Vale destacar que inicialmente uma requisição REST seria responsável pela transmissão do vídeo e de algumas informações da captura, enquanto que assincronamente outras requisições REST iriam realizar a transmissão das imagens faciais, pois dessa forma esperava-se agilizar o processo de envio dos dados e diminuir o tempo que o cliente precisa esperar para o envio das informações necessárias. Porém, durante a implementação do método foi identificado que enviar as imagens adequadas de maneira assíncrona causou uma maior lentidão dos envios de todos os dados, pois em alguns testes as imagens eram enviadas apenas após o envio do vídeo. Dessa forma, a maneira mais eficiente encontrada para enviar os dados foi utilizando uma única requisição REST transmitindo a posição das imagens adequadas no vídeo, o vídeo e as informações do paciente como peso e altura.

A posição das imagens adequadas será utilizada pelo servidor para extrair e armazenar as imagens adequadas conforme descrito na Seção 4.5.8.

4.5.8 Recorte dos *frames*

Essa é a primeira etapa do servidor após o envio bem sucedido do vídeo e dos dados da captura. Dessa forma, uma lista com as posições das imagens faciais adequadas em milissegundos no vídeo é utilizada para a extração de cada *frame* que será armazenado por meio das técnicas descritas no módulo de armazenamento (Subseção 4.5.9).

Ao percorrer a lista com as posições das imagens faciais adequadas, a aplicação armazena cada *frame* encontrado em um diretório específico para aquela captura. Após armazenar todas as imagens faciais, o servidor envia para o módulo de armazenamento (Seção 4.5.9) o caminho de acesso para as imagens.

4.5.9 Armazenamento

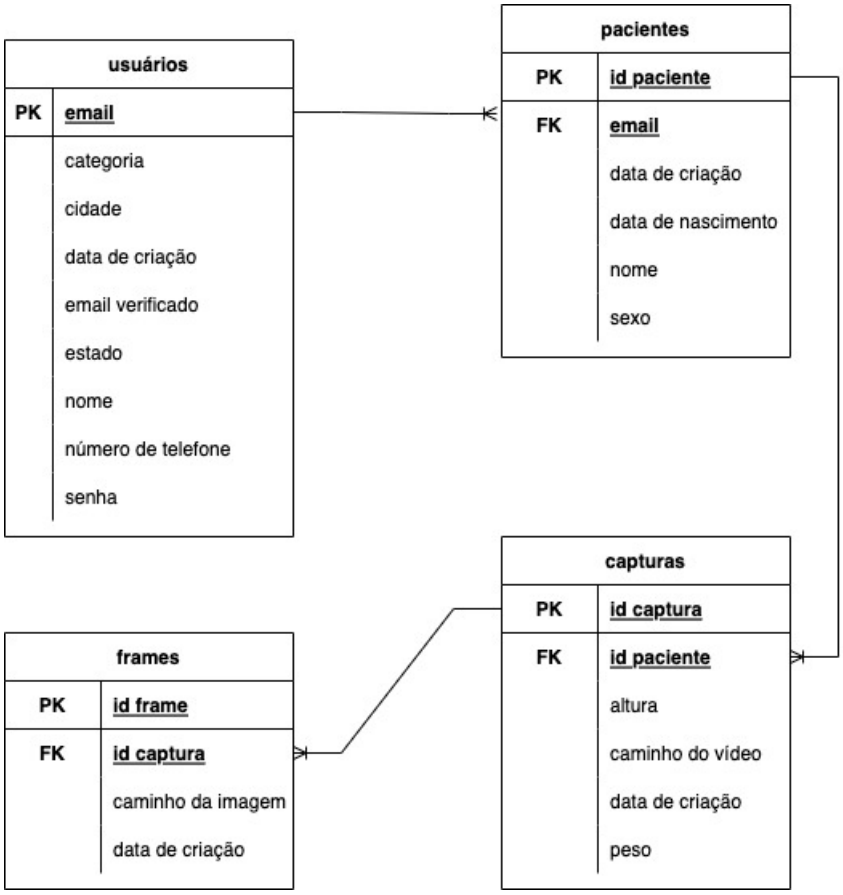
Este módulo é responsável pelo armazenamento dos dados da captura, que poderão ser utilizados para o treinamento de modelos de classificação responsáveis pelo auxílio ao diagnóstico do TEA. Por meio de um banco de dados relacional são armazenados tais dados conforme o modelo relacional apresentado na Figura 17.

Neste modelo, existem 4 entidades: usuários, pacientes, capturas e *frames*. Cada entidade possui uma chave primaria identificada pelas letras "PK"ao lado esquerdo do atributo e algumas entidades possuem uma chave estrangeira identificada pelas letras "FK"ao lado esquerdo do atributo conforme apresentado na Figura 17.

As informações relacionadas aos usuários, que são responsáveis por realizar a captura das imagens faciais adequadas do indivíduo, são armazenadas pela entidade usuários. Inicialmente esse usuário pode ser algum pesquisador responsável pela coleta de imagens faciais para a pesquisa.

Portanto, o usuário pode realizar a captura de imagens faciais de vários indivíduos, que são identificados como pacientes dentro do modelo relacional apresentado. Assim, para cada paciente existem diversas capturas realizadas e para cada captura existem diversos *frames*, que representam as imagens faciais adequadas armazenadas. A Tabela 6 apresenta detalhadamente cada atributo presente nas entidades apresentadas na Figura 17.

Figura 17 – Modelo relacional para o armazenamento das imagens adequadas.



Fonte: Henrique Silva Bortoletti, 2022.

Os dados do usuário são registrados durante o cadastro do usuário (Seção 5.1), os dados do paciente são registrados durante o registro de paciente (Seção 5.4), assim como, as demais entidades são registradas durante a captura.

4.5.10 Extração de características e modelo de classificação

Esta Subseção apresenta o módulo de extração de características e o módulo responsável pelo modelo de classificação. É importante ressaltar que ambos os módulos foram desenvolvidos inicialmente em Pinheiros (2018) e ao longo do desenvolvimento deste trabalho alguma mudanças foram realizadas.

Portanto, o primeiro módulo é responsável por identificar as características necessárias para o modelo de classificação. Na versão inicial deste módulo, as características são as distâncias entre os pontos de controle extraídos por meio do algoritmo *Constrained Local Neural Fields* (CLNF). Para isto, tal módulo utilizou a ferramenta *OpenFace* (BALTRUSAITIS; ROBINSON; MORENCY, 2016), que identifica 68 pontos de controle

Tabela 6 – Descrição dos atributos presentes nas entidades do modelo relacional.

Entidade	Atributo	Descrição
Usuários	email	E-mail de registro do usuário
	categoria	Usuário com senha fixa ou com senha temporária
	cidade	Cidade em que vive
	data de criação	Data em que o usuário foi criado
	e-mail verificado	Assume o valor verdadeiro se o e-mail for verificado
	estado	Estado em que vive
	nome	Nome do usuário
	número de telefone	Número de telefone para contato
	senha	Senha para autenticação
Pacientes	id paciente	Identificador do paciente
	email	E-mail de registro do usuário vinculado ao paciente
	data de criação	Data em que o paciente foi inserido
	data de nascimento	Data de nascimento do paciente
	nome	Nome do paciente
	sexo	Sexo do paciente
Capturas	id captura	Identificador da captura
	id paciente	Identificador do paciente vinculado à captura
	altura	Altura do paciente na ocasião da captura
	caminho do vídeo	Caminho do vídeo para acesso no servidor
	data de criação	Data e hora em que a captura foi realizada
	peso	Peso do paciente na ocasião da captura
Frames	id frame	Identificador do frame
	id captura	Identificador da captura vinculada ao frame
	caminho da imagem	Caminho da imagem para acesso no servidor
	data de criação	Data e hora em que o frame foi capturado

Fonte: Henrique Silva Bortoletti, 2022.

facial. Após a identificação destes pontos é gerado um vetor contendo a localização de cada um destes. Então, as distâncias entre cada par de pontos de controle são calculadas e utilizadas pelo modelo de classificação.

Logo, as imagens a serem coletadas pela aplicação proposta neste trabalho serão utilizadas pelo modelo de classificação, para treiná-lo e aprimorá-lo utilizando uma nova amostra composta pelas novas imagens capturadas. É importante destacar que o resultado do modelo de classificação não será enviado para os usuários que utilizarem a aplicação.

4.6 Avaliação da aplicação

Após o desenvolvimento da proposta apresentada neste capítulo, avaliações experimentais foram realizadas para avaliar a interação do usuário com o sistema. Dessa

forma, um mesmo experimento foi aplicado para dois grupos diferentes de usuários. O primeiro grupo realizou o experimento no mesmo dia e horário e é composto por dois profissionais de saúde responsáveis pela atual coleta de imagens faciais conforme proposto em Pinheiros (2018), por dois convidados que não conheciam o contexto sobre a coleta de imagens faciais e por um pesquisador da área da computação que já conhecia o contexto sobre o projeto. O segundo grupo realizou o experimento de maneira individual em dias e horários diferentes, sendo composto por sete profissionais da área da computação, que inicialmente não conheciam o contexto prévio sobre o projeto.

Em ambos os grupos foi utilizada a mesma forma de avaliação da implementação da proposta. Os voluntários receberam um questionário com todas as informações necessárias para avaliar o sistema desenvolvido. Ambos os grupos possuíam os recursos necessários para instalar a aplicação em seus respectivos dispositivos móveis compatíveis com as necessidades do sistema.

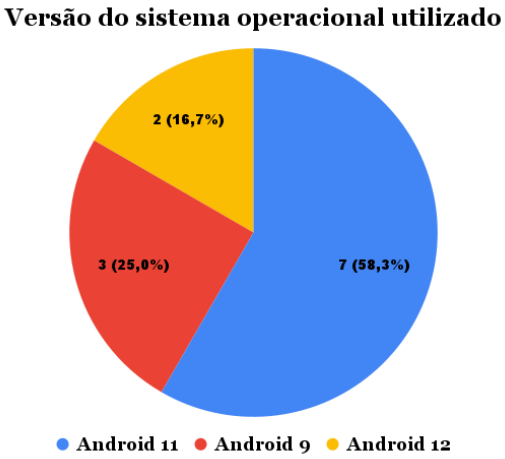
É importante destacar que no primeiro grupo os participantes tiveram uma ajuda externa, já que os pesquisadores responsáveis pelo presente projeto coordenaram todas as etapas da avaliação da aplicação. Dessa forma, em alguns momentos os responsáveis pelo projeto ajudaram os participantes a entender determinadas perguntas do questionário e funcionalidades da aplicação. Isso ocorreu apenas no início do experimento, enquanto os participantes instalavam a aplicação em seus respectivos dispositivos móveis. Já o segundo grupo não obteve tal ajuda externa enquanto os participantes realizavam o experimento.

Na Figura 18 são apresentadas as versões do sistema operacional *Android* utilizadas pelos participantes durante o experimento. A maior parte dos voluntários realizou o experimento por meio de um dispositivo com a versão 11 do sistema operacional *Android*, enquanto que alguns voluntários realizaram o experimento utilizando um dispositivo móvel com a versão 12, sendo esta a mais recente versão no momento em que foi realizada a avaliação da aplicação.

4.6.1 Questionário

Desde o lançamento dos primeiros dispositivos móveis a usabilidade das aplicações presentes nestes dispositivos tem sido uma preocupação constante dos criadores de *softwares*

Figura 18 – Tempo necessário para cada voluntário da área da computação realizar as tarefas.



Fonte: Henrique Silva Bortoletti, 2022.

para tais aparelhos. Portanto, entender a usabilidade de uma aplicação é fundamental para modelar sistemas com qualidade para o público alvo (WEICHBROTH, 2020).

Para isso, um questionário baseado na ISO 9241-11 (GENEVA, 2022) foi desenvolvido para avaliar três principais vertentes da implementação da proposta: eficácia, eficiência e satisfação. Portanto, o objetivo do método proposto é de identificar os pontos fortes e fracos de usabilidade da aplicação desenvolvida, pois espera-se que tal ferramenta seja utilizada por diversos tipos de usuários. Por meio do questionário também almejou-se avaliar o esforço cognitivo e a curva de aprendizado necessária para utilizar a aplicação.

O esforço cognitivo aplicado durante o uso de uma aplicação pode ser dividido em três tipos: externo, interno e sobressalente. Primeiramente, o esforço cognitivo externo se refere à interação do usuário com elementos que não estão presentes na aplicação, mas impactam diretamente como o usuário se relaciona com o sistema, tais como manuais de instrução, elementos visuais relacionados ao uso de determinada funcionalidade ou tutoriais de como utilizar determinada funcionalidade. O esforço cognitivo interno se refere à complexidade das funcionalidades presentes na aplicação, tal complexidade pode ser avaliada por meio da composição dos elementos presentes na aplicação e como estes diferentes elementos se relacionam. Já o esforço sobressalente refere-se ao grau de desgaste mental sofrido pelo usuário para entender a integração das diversas atividades da aplicação e integrar ativamente novas informações com conhecimento prévio durante o uso do aplicativo. Independentemente do tipo de esforço cognitivo que o pesquisador pretende

avaliar, este pode utilizar diversos tipos de avaliações subjetivas como uma escala de classificação para avaliar o uso da aplicação pelos usuários (WEICHBROTH, 2020).

Assim, a curva de aprendizado pode ser dividida em duas partes. A primeira parte se refere ao grau de dificuldade que um usuário possui ao interagir pela primeira vez com um sistema sem um guia ou um manual. Dessa forma, pode-se medir o sucesso dessa primeira parte pela quantidade de vezes necessárias para resolver uma determinada tarefa, o número de assistências necessárias e o número de erros cometidos pelo usuário. Já a segunda parte da curva de aprendizado se refere ao desempenho do usuário ao longo do tempo e pode ser observado por meio de tarefas executadas de maneira consecutiva pelo usuário (WEICHBROTH, 2020).

Para atingir o objetivo do questionário, este é dividido em tarefas que descrevem os passos que devem ser seguidos para avaliar determinadas funcionalidades da aplicação. A Figura 19 ilustra a ordem que as tarefas são apresentadas no questionário e o objetivo de cada uma destas.

Figura 19 – Fluxograma de tarefas presentes no questionário.

Fonte: Henrique Silva Bortoletti, 2022.

A tarefa 0 é responsável por avaliar a dificuldade da instalação do aplicativo e do manual. A tarefa 1 descreve ao usuário como realizar o cadastro para utilizar a aplicação. Ao fim desta tarefa é necessário que o pesquisador responsável aprove o cadastro de cada usuário, pois durante a construção do sistema este foi um dos requisitos. Assim, a cada novo usuário cadastrado, os seus dados são validados por um profissional da equipe de saúde. Com a aprovação do cadastro, o usuário pode realizar a tarefa 2, que avalia o processo de autenticação da aplicação.

Posteriormente à autenticação do sistema, o usuário é direcionado para a tela de registro de paciente; a tarefa 3 avalia a usabilidade do registro de um novo paciente pelo

usuário. Finalmente o usuário inicia a tarefa 4, que é responsável pela avaliação da captura do vídeo do paciente registrado anteriormente. Nesta tarefa é necessário que os voluntários do experimento encontrem uma segunda pessoa para realizar a captura das imagens faciais do indivíduo. Prosseguindo para a próxima etapa, a tarefa 6 avalia o processo de troca de senha pelo usuário.

As tarefas 5 e 7 repetem os passos das tarefas descritas anteriormente, com exceção das etapas relacionadas à criação de um usuário e troca de senha, pois dessa forma será possível avaliar a curva de aprendizado durante a utilização da aplicação. Com isso espera-se avaliar a eficácia e a eficiência da aplicação.

Para que a realização de cada tarefa possa ser avaliada posteriormente foram definidas as seguintes questões:

- **Q1**: Que horas a tarefa foi realizada?
- **Q2**: Que horas a tarefa foi finalizada?
- **Q3**: O(a) Senhor(a) teve alguma dificuldade ao realizar a tarefa?
- **Q4**: O(a) Senhor(a) se sentiu mentalmente cansado após realizar a tarefa?
- **Q5**: O(a) Senhor(a) encontrou algum erro (alertas em vermelho avisando que algo não funcionou, aplicativo fechou inesperadamente etc) durante a execução da tarefa?
- **Q6**: Caso tenha encontrado algum erro durante a execução da tarefa, por favor, descreva aqui.
- **Q7**: O(a) Senhor(a) precisou utilizar o manual?
- **Q8**: O(a) Senhor(a) entendeu o que significa todos os elementos (textos, botões, caixas de entrada e etc.) presentes na interface?
- **Q9**: O(a) Senhor(a) sentiu algum incômodo (botões não funcionando como esperado, interface travando, tempo de resposta do aplicativo alto, etc) ao realizar a tarefa?
- **Q10**: O(a) Senhor(a) encontrou dificuldades de leitura da interface (tamanho, cor, posição etc dos elementos em tela)?
- **Q11**: Caso tenha encontrado alguma dificuldade, por favor, descreva aqui.

Tais perguntas foram aplicadas para todas as tarefas, com exceção da tarefa 0 que não possui as perguntas **Q7, Q8, Q9, Q10 e Q11**, pois esta tarefa apenas avalia a instalação da aplicação e do manual, porém tais questões são aplicáveis ao contexto das tarefas responsáveis por avaliar a usabilidade da aplicação.

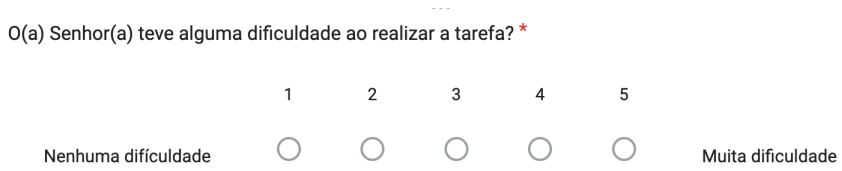
Para responder as perguntas **Q1 e Q2**, é solicitado aos participantes que escrevam respectivamente os horários de início e fim de realização da tarefa. Nas perguntas **Q6 e Q11**, caso seja encontrado algum erro ou dificuldade, então é disponibilizado um espaço para o participante descrever em um parágrafo o comportamento observado.

Para responder a pergunta **Q7** os participantes escolheram uma das opções abaixo:

- Não
- Sim, o manual me ajudou a resolver minhas dúvidas
- Sim, mas o manual não me ajudou a resolver minhas dúvidas

Para responder as perguntas **Q3, Q4, Q5, Q8, Q9 e Q10**, foi disponibilizado no questionário uma escala *Likert* para cada questão com valores de 1 a 5. Na Figura 20, pode-se observar a escala *Likert* aplicada à pergunta **Q3**. Para as demais perguntas, os valores nas extremidades foram descritos conforme a avaliação a ser feita pelo participante.

Figura 20 – Escala *Likert* para a pergunta Q3(O(a) Senhor(a) teve alguma dificuldade ao realizar a tarefa?)



Fonte: Henrique Silva Bortoletti, 2022.

Dessa forma, nas perguntas **Q3 e Q4** os valores próximos a 1 representam que o participante encontrou alguma dificuldade ao realizar a tarefa e se sentiu mentalmente cansado após realizar a tarefa, já os valores próximos a cinco representam que o usuário não teve dificuldades e não se sentiu mentalmente cansado após realizar a tarefa. Na pergunta **Q5**, os valores próximos a um representam que muitos erros foram encontrados durante a execução da tarefa, porém os valores próximos a cinco demonstram que nenhuma falha foi detectada.

Nas perguntas **Q8, Q9, e Q10** quanto mais próximo de um, então mais positiva é a avaliação e quanto mais próximo de cinco mais negativa é a avaliação da aplicação. Por exemplo, se o participante entende o significado de todos os elementos presentes na interface e não encontra dificuldades de leitura então o valor que representa isso na escala *Linkert* será um para as questões **Q8 e Q9**. Caso contrário o valor será mais próximo de cinco.

Logo, as tarefas constituem o questionário de forma que seja possível avaliar as funcionalidades presentes na aplicação, assim como avaliar a usabilidade da aplicação pelos usuários de forma que seja possível identificar os possíveis pontos de melhoria e avaliar se é possível atingir o objetivo da proposta.

4.6.2 Condução dos experimentos

Ao concordarem em participar do experimento, os voluntários receberam um questionário com as instruções necessárias para a avaliação da aplicação por meio de um *link* contendo a aplicação desenvolvida e o manual de como realizar cada funcionalidade presente no sistema. A primeira página do questionário está em anexo no Apêndice B e o manual está em anexo no Apêndice C. O manual também foi avaliado pelos participantes, pois dessa forma pode-se observar o impacto deste durante o uso da aplicação e a sua necessidade para a realização das funcionalidades.

Inicialmente foi realizado o experimento em grupo por via de uma videoconferência *online*. Um dos motivos para isto foi a necessidade do pesquisador responsável estar presente para aprovar o cadastro realizado pelos participantes para a utilização das demais funcionalidades da aplicação. Outro motivo para o experimento ter sido conduzido desta forma foi a possibilidade de observar de maneira mais próxima as dificuldades encontradas pelos usuários.

Assim, a execução do experimento seguiu o seguinte procedimento: cada participante recebeu o mesmo questionário descrito anteriormente e o pesquisador responsável além de aprovar o cadastro realizado pelos participantes, também coordenou a execução de cada tarefa para que todos os participantes as executassem ao mesmo tempo.

É importante ressaltar que o pesquisador responsável apenas auxiliou alguns participantes durante a instalação da aplicação, após isto este não interferiu na execução das demais tarefas, pois isto poderia interferir na avaliação do manual.

Esperava-se que os participantes executassem todas as tarefas em trinta minutos. Devido à dificuldade que alguns participantes encontraram ao realizar determinadas tarefas, o tempo total para a realização das atividades foi de aproximadamente uma hora. No Capítulo 6 é possível observar o tempo gasto por cada participante para a execução de cada tarefa.

Em seguida, foi realizado o experimento de maneira individual sem a presença dos pesquisadores responsáveis. Voluntários foram convidados para participarem do experimento de maneira que cada indivíduo tivesse a liberdade de escolher o melhor horário para a realização do experimento, mas foi necessário informar o momento que o experimento fosse realizado para que o pesquisador responsável validasse os dados de cadastro do usuário. Neste experimento o pesquisador responsável não auxiliou em nenhuma tarefa, assim como nenhum participante pediu sua ajuda.

Ao fim de ambos os experimentos, os participantes realizaram observações e sugestões de melhorias, tais informações e os resultados provenientes das avaliações dos participantes estão presentes no Capítulo 6.

4.7 Considerações finais

Este capítulo apresentou as etapas realizadas durante o desenvolvimento desta pesquisa, como a realização de experimentos quantitativos, o desenvolvimento da aplicação para a democratização do acesso ao auxílio diagnóstico de indivíduos com TEA, a avaliação da aplicação desenvolvida e a análise dos resultados.

No Capítulo 5, a aplicação desenvolvida é apresentada, detalhando-se suas funcionalidades. Os resultados dos experimentos e da avaliação da aplicação são apresentados no Capítulo 6.

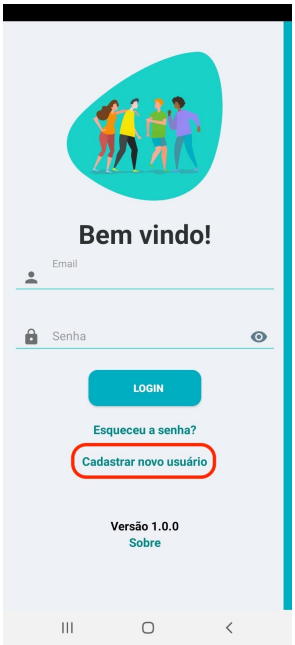
5 Aplicação desenvolvida

A aplicação apresentada neste capítulo é resultado da implementação dos métodos apresentados no Capítulo 4. Este sistema auxilia os pesquisadores durante a coleta de vídeos faciais contendo as imagens faciais adequadas. Cada seção deste capítulo apresenta o comportamento de uma funcionalidade específica presente na aplicação, com exceção apenas da Seção 5.7, que apresenta as considerações finais do Capítulo.

5.1 Cadastro de usuário

Com a aplicação devidamente instalada em seu dispositivo móvel, o primeiro passo para o usuário poder utilizar todas as funcionalidades presentes na aplicação é a realização do seu cadastro por meio da opção "cadastrar novo usuário"presente na tela inicial da aplicação, conforme apresentado na Figura 21.

Figura 21 – Interface inicial da aplicação com a opção para cadastro de novo usuário destacada por um contorno vermelho.



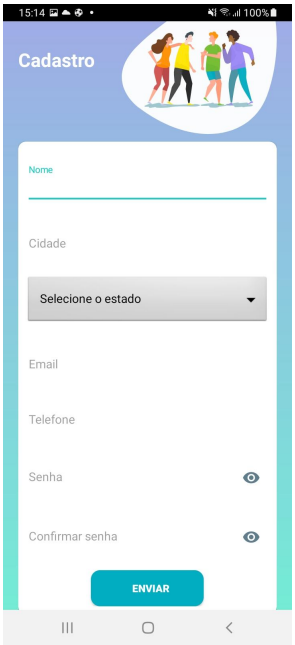
Fonte: Henrique Silva Bortoletti, 2022.

Ao iniciar o processo de cadastro o usuário é redirecionado para uma nova interface conforme a Figura 22, desse modo é possível cadastrar os dados relacionados ao usuário como: nome, cidade, estado *e-mail*, telefone e senha. Tais dados são necessários para que

os pesquisadores possam entrar em contato com o usuário, assim como para a validação do cadastro.

A senha cadastrada pelo usuário é criptografada no banco de dados por meio do algoritmo de criptografia simétrico *Bruce Schneier's Blowfish cipher*(FINCH, 1995). Tal técnica utiliza caracteres aleatórios para a criptografia da senha, portanto não é possível identificar os valores cadastrados pelos usuários na base de dados.

Figura 22 – Interface da aplicação para cadastro de usuário.



Fonte: Henrique Silva Bortoletti, 2022.

Após o cadastro, os usuários ainda não possuem o direito de usar a aplicação, pois um grupo de pesquisadores é responsável por avaliar se cada usuário possui permissão para utilizar a aplicação e realizar a captura dos vídeo com as imagens faciais adequadas.

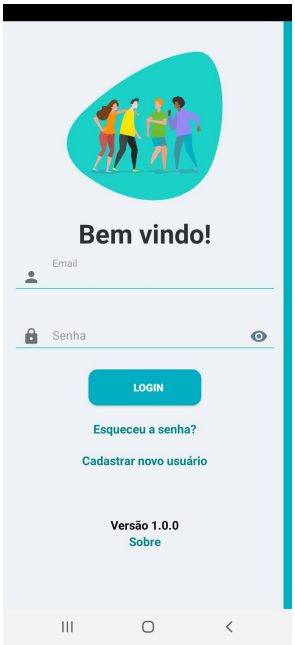
Esse processo ocorre inicialmente por meio de uma requisição REST enviada para o servidor, que armazena os dados cadastrados e envia um *e-mail* para cada pesquisador presente no grupo de responsáveis por determinar se cada usuário possui permissão para utilizar a aplicação. Após a aprovação do cadastro, um *e-mail* é enviado para o endereço cadastrado informando que o indivíduo pode utilizar a aplicação.

5.2 Autenticação

A autenticação do usuário é necessária apenas uma única vez durante o uso da aplicação e ocorre no momento em que o programa é iniciado. Dessa maneira, após o

usuário acessar a aplicação será exibida uma interface que possui um campo de entrada para o *e-mail* e um campo de entrada para a senha cadastrada pelo usuário, conforme apresentado na Figura 23.

Figura 23 – Interface inicial da aplicação desenvolvida.



Fonte: Henrique Silva Bortoletti, 2022.

Após o usuário inserir os seus dados e selecionar a opção "Login", estes são enviados para o servidor por meio de uma requisição REST, o qual irá autenticar o usuário por meio de uma verificação de *e-mail* e senha. Caso estes estejam armazenados no banco de dados conforme enviado pelo usuário, a autenticação é realizada com sucesso. Caso contrário, uma mensagem é enviada à aplicação informando ao usuário que uma falha ocorreu durante a autenticação pelo servidor.

Em seguida à autenticação, o servidor irá gerar um *token* de autenticação da aplicação que é armazenado no banco de dados com validade fixa de cinco horas. Caso este tempo expire, o *token* será removido do banco de dados. Dessa forma, cada requisição REST deve ser realizada ao servidor com um *token* válido e caso este não seja encontrado pelo servidor no banco de dados, a requisição REST irá automaticamente retornar um erro de autenticação.

Dessa forma, após o usuário realizar a autenticação, a aplicação recebe o *token* de autenticação e para cada requisição futura ao servidor tal *token* será enviado junto com os dados específicos da requisição para o servidor.

É importante ressaltar que o envio dos dados para o servidor é feito por meio do protocolo HTTPS; portanto, durante a comunicação entre a aplicação e o servidor a senha do usuário e os demais dados são criptografados, minimizando o risco de terceiros terem acesso à senha utilizada do usuário.

5.3 Troca de senha

Para realizar a troca de senha, o usuário necessita acessar a aplicação e na interface inicial selecionar a opção "Esqueceu a senha?"para a troca de senha, que está destacada com contorno em vermelho na Figura 24.

Figura 24 – Interface da aplicação para iniciar o processo de troca de senha.

Fonte: Henrique Silva Bortoletti, 2022.

Uma nova interface será exibida possibilitando ao usuário informar o seu *e-mail*. Por meio de uma requisição REST será possível informar ao servidor, que aquele usuário necessita alterar a senha.

Em seguida, o servidor altera a senha do usuário para um conjunto de caracteres aleatórios, envia um *e-mail* para o usuário contendo a nova senha e altera o atributo "categoria"da entidade usuários para usuário temporário no banco de dados. Portanto, quando o usuário realiza uma nova autenticação, o servidor avalia se é necessária a troca de senha e, caso necessária, informa a aplicação.

Por meio da aplicação o usuário realiza a troca de senha informando uma nova senha, que será enviada em uma requisição REST ao servidor responsável por alterar a senha conforme o valor enviado. O atributo "categoria"também é atualizado para apenas usuário, dessa forma, não será necessário realizar a troca de senha novamente. Pois, apenas é necessário realizar a troca de senha quando o valor do atributo "categoria"é usuário temporário.

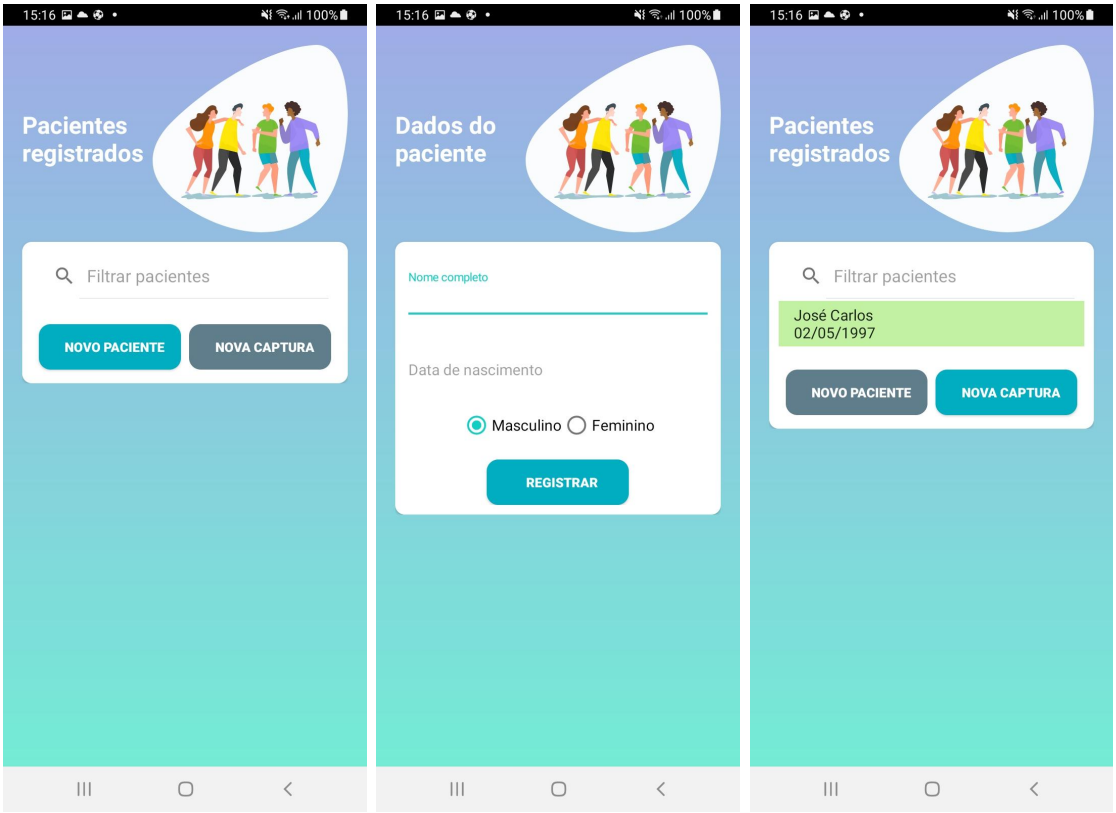
5.4 Novo paciente

O usuário cadastrado e autenticado deve registrar todos os pacientes que são de sua responsabilidade; por consequência, cada usuário pode ter acesso apenas aos pacientes sob seus cuidados.

Após a autenticação, o usuário é direcionado para uma interface com todos os pacientes já cadastrados por ele. Nesta interface também é possível realizar o cadastro de um novo paciente ou realizar a captura do vídeo facial de um paciente selecionado, conforme exemplificado na Seção 5.5. Para cadastrar um novo paciente, o usuário deve selecionar a opção "novo paciente", conforme a Figura 25**A**. Então, este é direcionado para a interface descrita na Figura 25**B**, possibilitando ao usuário preencher o nome completo, a data de nascimento e o sexo do paciente. Com os dados preenchidos o usuário seleciona a opção "registrar"conforme a Figura 25**B**, esta ação envia os dados registrados do paciente por meio de uma única requisição REST para serem armazenados pelo servidor. Por fim, o usuário é direcionado para a interface exemplificada pela Figura 25**C**, contendo o novo paciente adicionado à lista de pacientes registrados.

5.5 Captura de vídeo do paciente

O processo de identificação de imagens faciais adequadas é feito por meio da captura de um vídeo do paciente, que possui uma quantidade máxima de segundos e no qual espera-se encontrar uma quantidade mínima de imagens adequadas para que os dados sejam enviados ao servidor a fim de serem utilizados no auxílio ao diagnóstico. O tempo máximo do vídeo e o número mínimo de imagens adequadas podem ser alterados a qualquer momento conforme exemplificado na Seção 5.6.



(a) (b) (c)

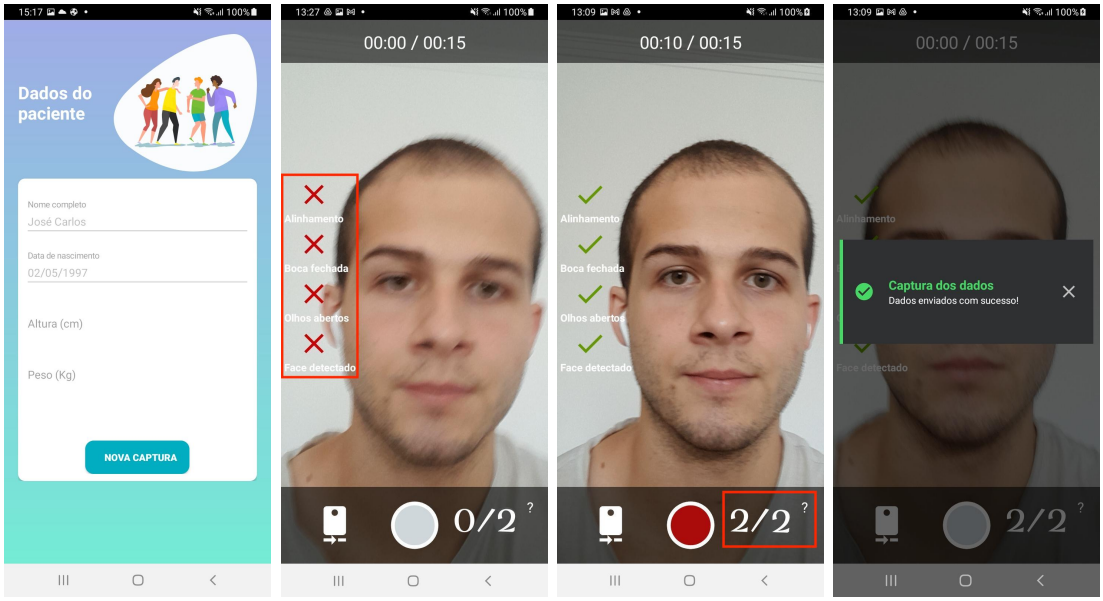
Figura 25 – Interface da aplicação para registrar novo paciente ou iniciar uma nova captura (a) Interface dos pacientes registrados sem nenhum paciente. (b) Interface dos dados do paciente a ser registrado. (c) interface dos pacientes registrados com um paciente selecionado

Fonte: Henrique Silva Bortoletti, 2022.

O primeiro passo para a identificação das imagens adequadas é a seleção do paciente pelo usuário responsável pela captura. Portanto, após realizar o registro do paciente, conforme descrito na Seção 5.4, o usuário seleciona o paciente cadastrado conforme a Figura 25**C**.

Após o usuário selecionar a opção "nova captura", será necessário preencher um breve formulário conforme descrito na Figura 26**A**, que contém os dados relacionados ao paciente no momento da captura como: altura em centímetros e peso em quilos. Estes dados são importantes para um acompanhamento do paciente pelos profissionais da saúde, assim como, podem ser utilizados na construção de modelos matemáticos em conjunto com as imagens adequadas.

Após este passo, o usuário será direcionado para uma nova interface, conforme descrita pela Figura 26**B**, que requisita preferencialmente a utilização da câmera traseira



(a) (b) (c) (d)

Figura 26 – Processo de captura de vídeo facial do paciente (a) dados relacionados ao paciente necessários para a captura do vídeo facial (b) Captura em andamento. (c) Imagens adequadas encontradas (d) Mensagem de captura realizada com sucesso.

Fonte: Henrique Silva Bortoletti, 2022.

do dispositivo móvel. Nesta interface, o usuário necessita selecionar o círculo na área inferior da interface para iniciar uma nova captura.

A Figura 26**B** representa o instante inicial da captura, pois o cronômetro na parte superior da tela do dispositivo móvel exibe o tempo de zero segundos dos 15 segundos totais de gravação. Também é possível observar na Figura 26**B**, que a imagem atual do vídeo não é a ideal, pois os parâmetros necessários destacados no retângulo vermelho à esquerda da tela: alinhamento, boca fechada, olhos abertos e face detectada, estão com um "X"em vermelho. Isso ocorre porque a imagem facial adequada não foi detectada; dessa maneira, todos os outros atributos também não serão considerados adequados, já que necessitam da localização da face.

Já na Figura 26**C**, a captura está em 10 dos 15 segundos totais da gravação e no canto destacado pelo retângulo em vermelho é possível identificar que o envio do vídeo pode ser realizado, pois as duas imagens faciais adequadas e necessárias foram encontradas. Também é possível observar que os parâmetros de captura estão em verde.

Quando o tempo de duração do vídeo for de 15 segundos e caso todas as imagens faciais adequadas tenham sido encontradas, então os seguintes itens serão enviados para o servidor: os dados do paciente, o vídeo, a localização das imagens adequadas no vídeo e o

e-mail do usuário responsável pela captura. Caso as imagens adequadas não tenham sido encontradas em 15 segundos, um novo processo de captura será automaticamente iniciado. Este processo se repete até que o número mínimo de imagens tenha sido encontrado ou até que o usuário finalize o aplicativo. Após o envio de todos os dados para o servidor, por meio de uma requisição REST, será exibida uma mensagem para o usuário conforme apresentada na Figura 26**D**.

No servidor, uma aplicação é responsável por receber os dados enviados pela requisição REST, armazena-los e extrair do vídeo as imagens adequadas, por meio dos tempos de cada uma destas, como sendo as localizações de cada imagem adequada no vídeo. Após a extração das imagens faciais, estas são armazenadas em um diretório do servidor.

5.6 Configuração dos parâmetros

Após a autenticação bem sucedida do usuário, uma requisição REST assíncrona é feita para obter os parâmetros que serão utilizados pela aplicação. Tais parâmetros são responsáveis pela forma como a captura é realizada e podem ser alterados por meio de um *script javascript*, que envia ao servidor por meio de uma requisição REST os novos valores a serem utilizados por todas as novas capturas, independente do usuário responsável.

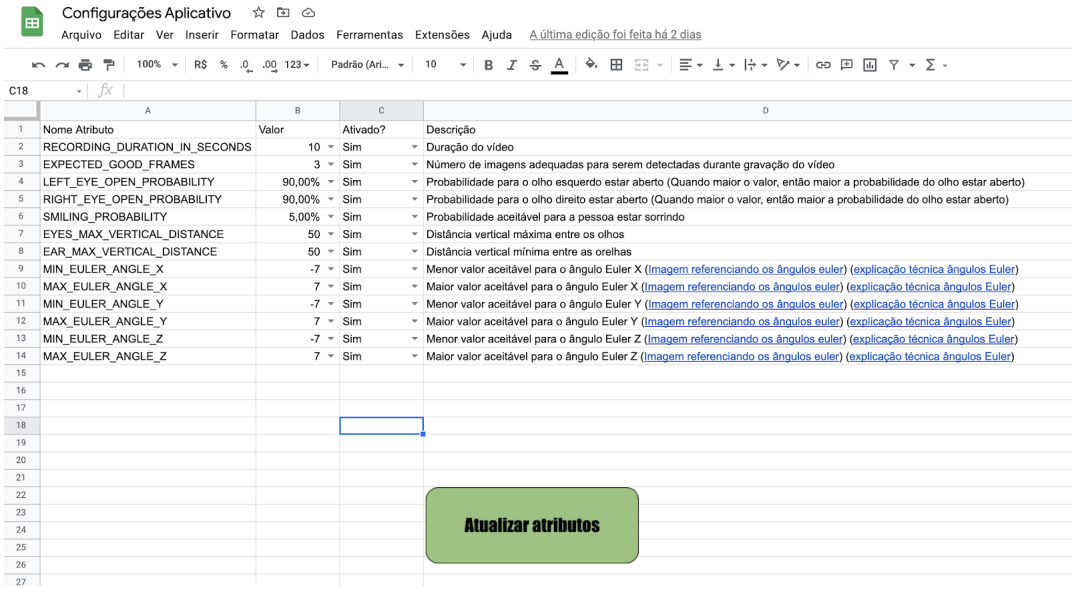
A Figura 27 descreve a planilha de texto conectada à rede mundial de computadores, que é responsável por exibir os valores atuais dos parâmetros configuráveis da aplicação. Caso seja necessária a mudança de algum parâmetro, então o usuário pode alterar o valor na planilha conforme descrito no manual (Apêndice C) e selecionar a opção "atualizar atributos"dentro do retângulo verde da Figura 27.

É importante ressaltar que no futuro espera-se que os usuários sejam divididos em categorias e uma interface seja implementada na aplicação presente no dispositivo móvel possibilitando que apenas usuários da categoria "administrador"possam realizar a mudança dos parâmetros.

5.7 Considerações finais

O atual Capítulo exemplificou as funcionalidades presentes na aplicação desenvolvida. No Apêndice C pode-se encontrar o manual, responsável por auxiliar os usuários na

Figura 27 – Planilha de texto responsável por gerenciar os parâmetros da aplicação.



	A	B	C	D
1	Nome Atributo	Valor	Ativado?	Descrição
2	RECORDING_DURATION_IN_SECONDS	10	Sim	Duração do vídeo
3	EXPECTED_GOOD_FRAMES	3	Sim	Número de imagens adequadas para serem detectadas durante gravação do vídeo
4	LEFT_EYE_OPEN_PROBABILITY	90,00%	Sim	Probabilidade para o olho esquerdo estar aberto (Quando maior o valor, então maior a probabilidade do olho estar aberto)
5	RIGHT_EYE_OPEN_PROBABILITY	90,00%	Sim	Probabilidade para o olho direito estar aberto (Quando maior o valor, então maior a probabilidade do olho estar aberto)
6	SMILING_PROBABILITY	5,00%	Sim	Probabilidade aceitável para a pessoa estar sorrindo
7	EYES_MAX_VERTICAL_DISTANCE	50	Sim	Distância vertical máxima entre os olhos
8	EAR_MAX_VERTICAL_DISTANCE	50	Sim	Distância vertical mínima entre as orelhas
9	MIN_EULER_ANGLE_X	-7	Sim	Menor valor aceitável para o ângulo Euler X (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
10	MAX_EULER_ANGLE_X	7	Sim	Maior valor aceitável para o ângulo Euler X (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
11	MIN_EULER_ANGLE_Y	-7	Sim	Menor valor aceitável para o ângulo Euler Y (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
12	MAX_EULER_ANGLE_Y	7	Sim	Maior valor aceitável para o ângulo Euler Y (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
13	MIN_EULER_ANGLE_Z	-7	Sim	Menor valor aceitável para o ângulo Euler Z (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
14	MAX_EULER_ANGLE_Z	7	Sim	Maior valor aceitável para o ângulo Euler Z (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)

Fonte: Henrique Silva Bortoletti, 2022.

utilização dessas funcionalidades. O Capítulo 6 apresenta o resultado do uso desta aplicação e o seu impacto durante a coleta de vídeos para o auxílio ao diagnóstico TEA.

6 Resultados e discussões

Conforme apresentado na Seção 4.6, dois grupos realizaram o mesmo experimento em dois momentos diferentes para avaliar a usabilidade da aplicação desenvolvida. Dessa forma, esta seção apresenta os resultados da avaliação dos voluntários durante o experimento realizado.

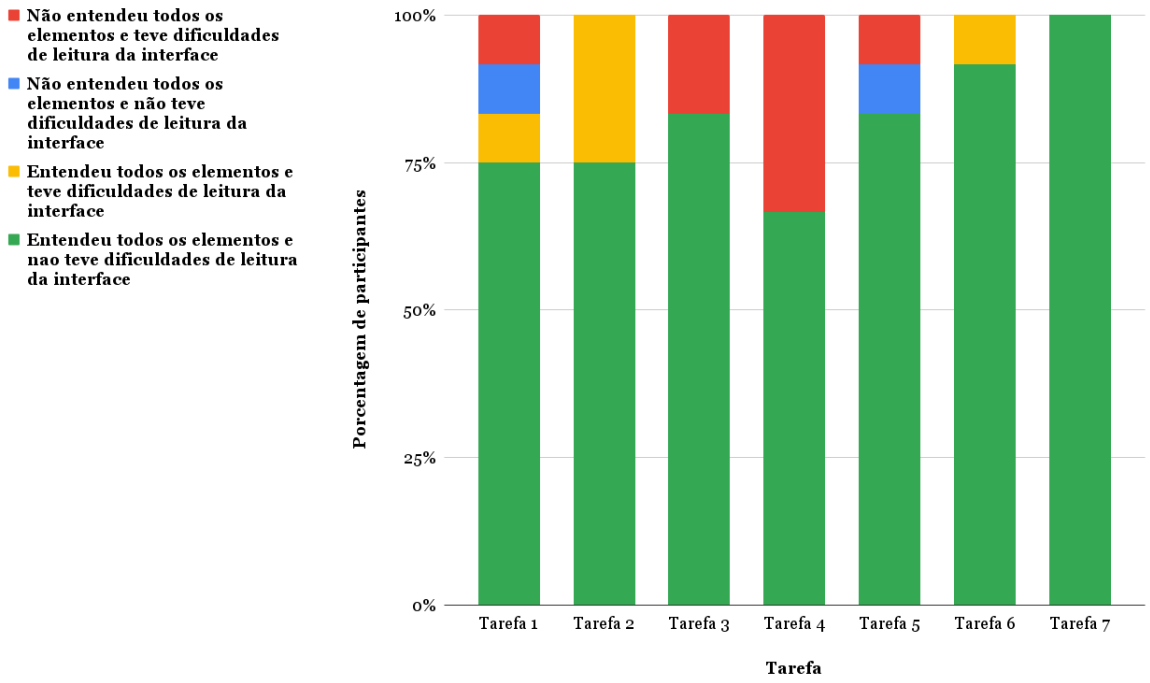
6.1 Esforço cognitivo

Para avaliar o esforço cognitivo necessário para a utilização da aplicação pelos voluntários foram realizadas as perguntas **Q4, Q8 e Q10** conforme descrito na Subseção 4.6.1. Pode-se observar na Figura 28 que pelo menos 65% dos voluntários conseguiram compreender todos os elementos da interface gráfica da aplicação e não apresentaram dificuldades de leitura da interface em todas as tarefas.

No entanto, é importante destacar as dificuldades encontradas pelos voluntários durante a realização da tarefa 4 (realizar a captura do paciente registrado na tarefa 3) na Figura 28, pois nesta tarefa a porcentagem de voluntários que não entenderam os elementos e encontraram dificuldades de leitura da interface foi maior que nas demais tarefas, com exceção da tarefa zero (instalação da aplicação no dispositivo móvel). Um dos motivos mencionados pelos voluntários foi falta de entendimento para a necessidade de selecionar um paciente antes de inciar o processo de captura. Alguns voluntários também destacaram a dificuldade em compreender se a captura havia se iniciado automaticamente após a seleção do paciente ou qual ação era necessária para começar o processo. Por fim, alguns voluntários também destacaram dificuldade em alterar da câmera frontal para a câmera traseira para realizar a detecção facial do indivíduo ao seu lado.

Na Figura 29 é possível observar que para compreender a interface da aplicação a maioria dos usuários não se sentiu cansada após realizar as tarefas necessárias. Pode-se inferir que a maior parte dos usuários não despendeu um grande esforço cognitivo para realizar as tarefas. Podemos observar em ambas as Figuras 28 e 29, que conforme as tarefas foram sendo executadas pelos usuários, estes tiveram uma melhor compreensão e leitura dos elementos presentes na interface gráfica. A quantidade média de usuários que se sentiram mentalmente desgastados após a realização de uma tarefa também se manteve constante entre nenhum e dois usuários, com exceção da tarefa 5 (realizar uma

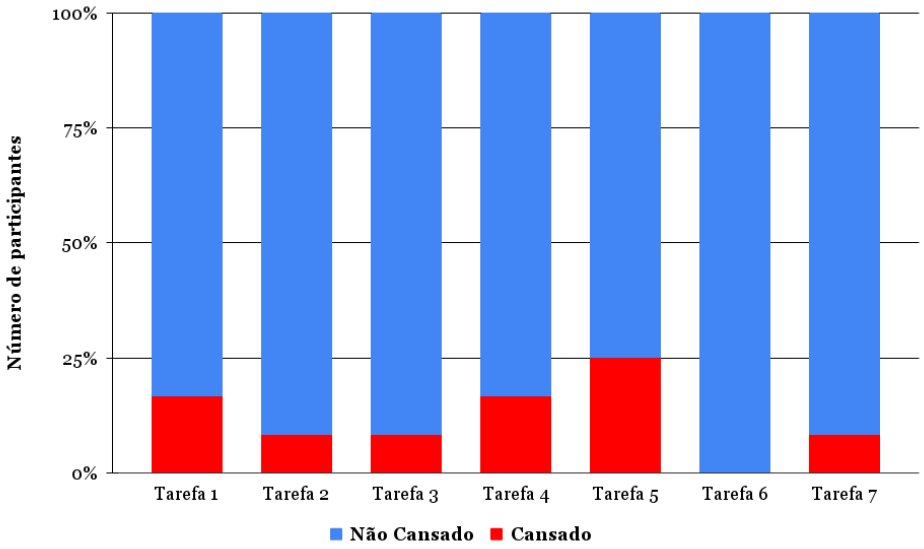
Figura 28 – Cognição da interface da aplicação pelos voluntários a cada tarefa realizada.



Fonte: Henrique Silva Bortoletti, 2022.

nova captura para um paciente já registrado) que teve três usuários que se declararam cansados.

Figura 29 – Porcentagem de voluntários cansados e de voluntários não cansados após a realização de cada tarefa por meio da aplicação desenvolvida.



Fonte: Henrique Silva Bortoletti, 2022.

Tais dificuldades podem ter sido observadas porque a tarefa 4 é responsável por demandar aos voluntários pela primeira vez a realização da captura do vídeo com as

imagens faciais adequadas. Pode-se notar que na tarefa 7 os usuários não tiveram nenhum tipo de dificuldade, sendo que esta tarefa possui mais passos que as demais tarefas, pois é responsável por avaliar o processo de autenticação do usuário, cadastro de novo paciente e realização de nova captura. Logo, pode-se concluir que conforme os voluntários realizaram novas capturas, a carga cognitiva decresce gradativamente, fazendo com que a compreensão e a leitura da interface gráfica também se torne mais simples, indicando que há um aprendizado acumulado.

6.2 Curva de aprendizado

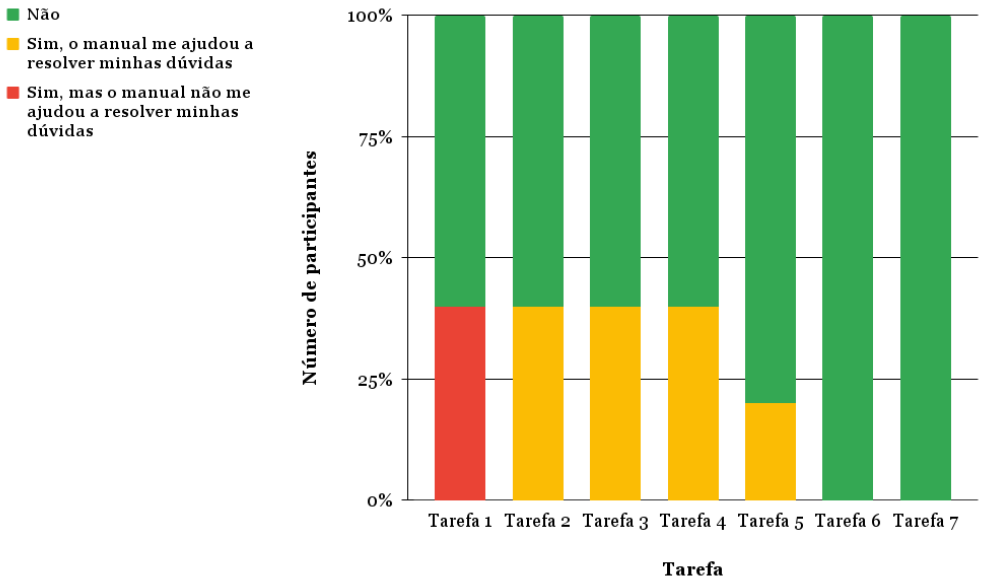
Durante a realização do experimento descrito na Seção 4.6 foi necessário que os voluntários respondessem um questionário, indicando se precisaram utilizar o manual para cada tarefa realizada e o grau de dificuldade para realizar a tarefa exigida. Consequentemente, foi verificado o grau de dificuldade e o desempenho dos usuários.

Os gráficos presentes nas Figuras 30 e 32 avaliam a dificuldade do grupo de voluntários majoritariamente composto por pesquisadores da área da saúde e possuíam algum conhecimento prévio do projeto. Os gráficos das Figuras 31 e 33 representam os mesmos gráficos anteriormente citados referentes aos dados dos voluntários da área da computação, que majoritariamente não possuíam conhecimento prévio do projeto de pesquisa.

Pode-se observar na Figura 30, que o manual foi necessário por menos da metade dos participantes até a realização da tarefa 5 (realizar uma nova captura para um paciente já registrado); contudo, é importante observar a ineficiência do manual durante a realização da primeira tarefa responsável por avaliar o processo de cadastro do usuário, segundo os voluntários.

Em contrapartida, o grupo dos pesquisadores da área da computação apresentou maior costume de utilizar o manual de acordo com a Figura 31, já que pelo menos um dos participantes necessitou do manual até a realização da tarefa 6 (realizar uma troca de senha). Além disso, durante a tarefa 4 (realizar a captura do paciente registrado na tarefa 3) a maior parte dos participantes precisou utilizar o manual. Provavelmente um dos motivos para isto ocorrer é que as ações necessárias para realizar uma captura de imagens faciais adequadas não tenha ficado claras para os novos usuários sem conhecimento prévio

Figura 30 – Porcentagem de pesquisadores da área da saúde por tarefa que precisaram utilizar o manual de instruções durante o experimento.



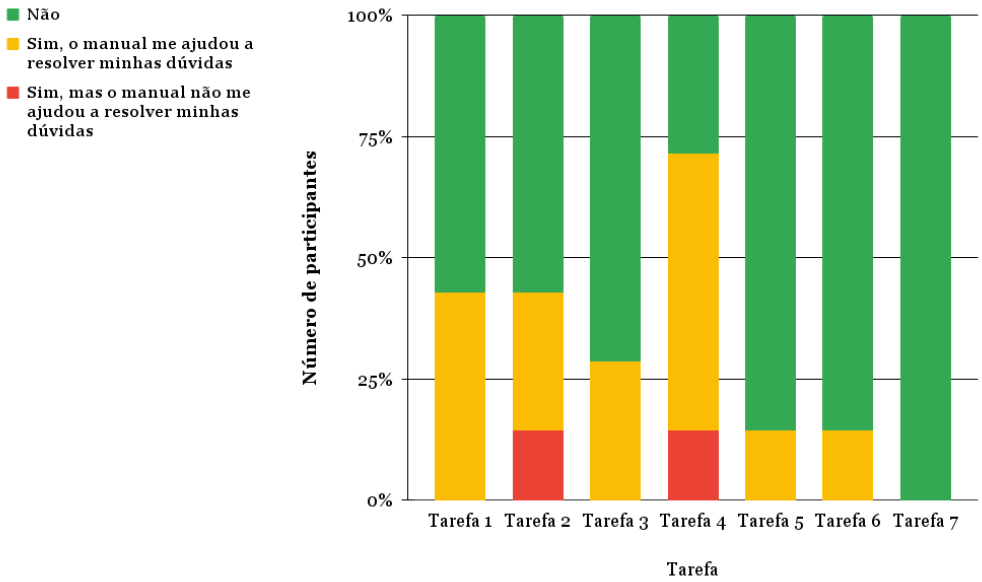
Fonte: Henrique Silva Bortoletti, 2022.

do projeto de pesquisa. Entretanto, o questionário aplicado possui uma breve descrição do projeto de pesquisa e da aplicação. Além disso, antes de iniciar o experimento, o grupo de pesquisadores da área da computação recebeu uma mensagem detalhada sobre a pesquisa realizada. Vale ressaltar que uma sugestão proposta por um dos participantes foi a utilização de vídeo para explicar o uso de cada funcionalidade da aplicação junto com o manual.

Durante o experimento os participantes também avaliaram a dificuldade encontrada para realizar cada tarefa. Os níveis de dificuldade de cada participante são apresentados nas Figuras 32 e 33. Pode-se observar na Figura 32 que os participantes tiveram maior dificuldade ao realizar as tarefas comparados com os participantes na Figura 33. Também é possível observar que conforme as tarefas são desenvolvidas, o nível de dificuldade decresce, indicando que os indivíduos aprendem com a repetição das mesmas ações em diversas tarefas.

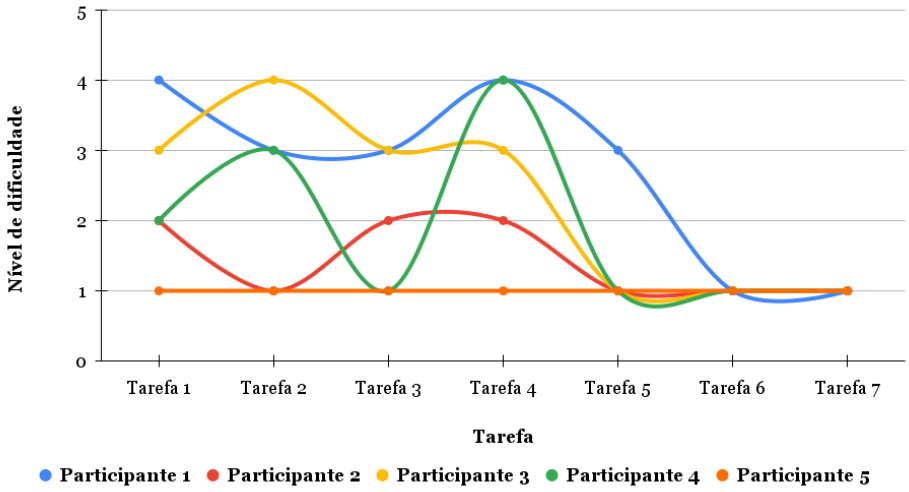
Pode-se notar um contraste entre os gráficos das Figuras 30 e 31 em relação aos gráficos as Figuras 32 e 33. O grupo de pesquisadores da área da saúde não utilizou o manual de instruções na mesma intensidade que os pesquisadores da área da computação. Porém, o grupo da área da saúde relatou maior dificuldade ao realizar as tarefas, enquanto estes mantiveram-se constantes atribuindo um nível de dificuldade igual ou inferior a dois

Figura 31 – Porcentagem de pesquisadores da área da computação por tarefa que precisaram utilizar o manual de instruções durante o experimento.



Fonte: Henrique Silva Bortoletti, 2022.

Figura 32 – Dificuldade dos pesquisadores da área da saúde ao realizar determinada tarefa durante o experimento.

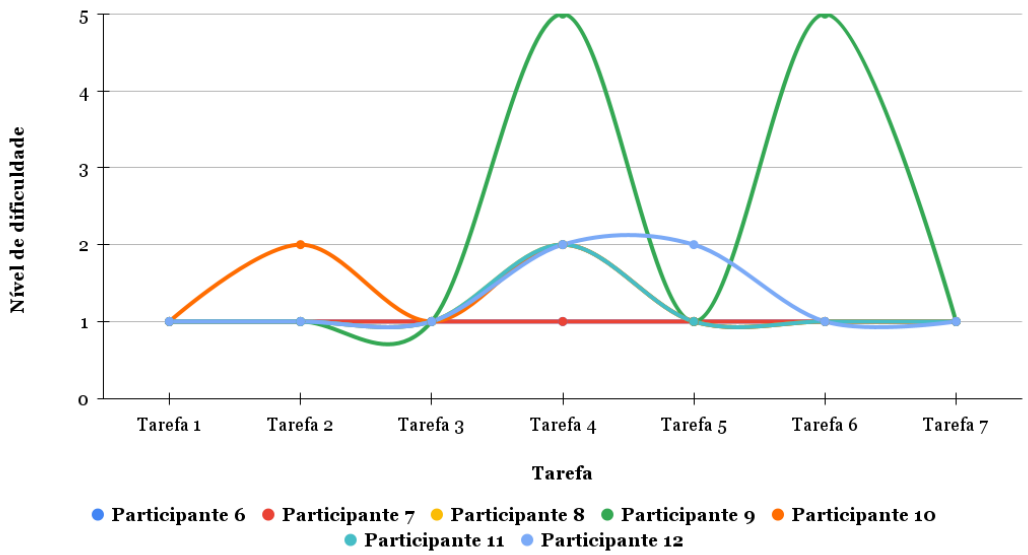


Fonte: Henrique Silva Bortoletti, 2022.

para todas as tarefas, com exceção do participante nove que teve problemas técnicos ao realizar a tarefa 4 e inicialmente não entendeu a tarefa 6, que é responsável por avaliar a troca de senha pela aplicação.

Dessa forma, observa-se a necessidade do manual para os novos usuários compreenderem a aplicação sem muitas dificuldades, pois verifica-se que houve dificuldade mesmo para os usuários que possuíam um conhecimento prévio sobre o projeto de pesquisa.

Figura 33 – Dificuldade dos pesquisadores da área da computação ao realizar determinada tarefa durante o experimento.



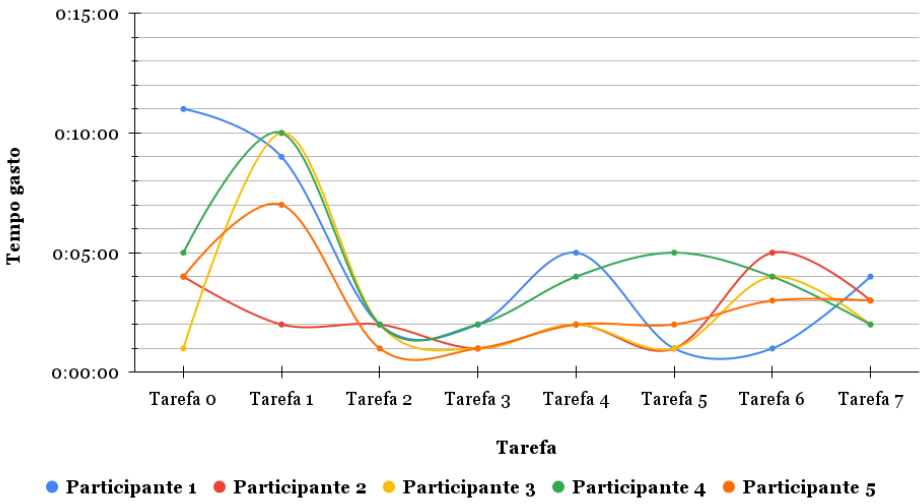
Fonte: Henrique Silva Bortoletti, 2022.

Para avaliar o desempenho de cada voluntário foi mensurado o tempo gasto por cada participante ao realizar cada tarefa presente no experimento (Figuras 34 e 35). A partir da tarefa 2, ambos os grupos mantiveram uma média próxima de tempo para realizar as tarefas. Dessa forma, considera-se o desempenho dos participantes de ambos os grupos semelhante.

Portanto, ao observar as Figuras 34 e 35, pode-se notar que o tempo gasto por cada grupo decresce de maneira constante, mesmo com a complexidade de cada tarefa aumentando gradativamente.

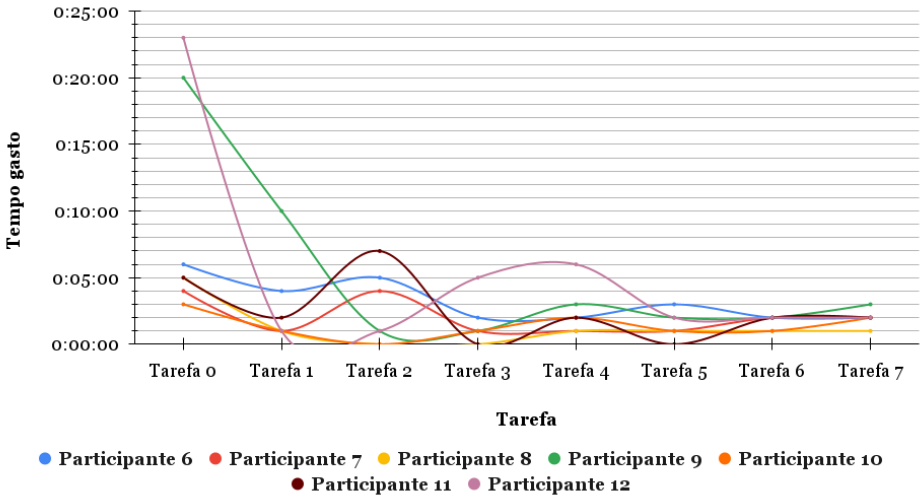
Ao se avaliar os itens anteriores em conjunto, há indícios de que os voluntários apresentaram uma maior dificuldade ao realizar as primeiras tarefas, pois enquanto o grupo de pesquisadores da área da computação utiliza de maneira mais constante o manual para realizar as tarefas iniciais, o grupo de pesquisadores da área da saúde relata maior dificuldade ao concluir tais atividades. Adicionalmente, verifica-se que a dificuldade foi diminuindo ao longo do experimento, visto que ambos os grupos necessitaram menos da utilização do manual e relataram uma menor dificuldade ao realizar as novas tarefas. Este achado é ratificado durante a avaliação do desempenho do usuário por meio do tempo dispendido em cada tarefa. Em ambos os grupos todos os indivíduos necessitaram de um tempo menor para executar as últimas tarefas quando comparado ao tempo das tarefas iniciais.

Figura 34 – Tempo necessário para cada voluntário da área da saúde realizar as tarefas.



Fonte: Henrique Silva Bortoletti, 2022.

Figura 35 – Tempo necessário para cada voluntário da área da computação realizar as tarefas.



Fonte: Henrique Silva Bortoletti, 2022.

6.3 Falhas identificadas

Durante a realização do experimento descrito na Seção 4.6, os participantes, por meio de duas questões opcionais, descreveram os erros encontrados no sistema e relataram quais foram as principais dificuldades de interpretação ao realizar cada tarefa.

A primeira pergunta está relacionada às falhas que podem surgir enquanto a aplicação está sendo executada, por exemplo, alguma funcionalidade não é executada como esperado, longo tempo de espera ao realizar uma ação, a aplicação não funciona

no dispositivo em determinado cenário ou falha de comunicação entre o dispositivo e o servidor.

Já a segunda pergunta está relacionada às dificuldades que os voluntários descreveram ao realizar determinada tarefa, por exemplo, o voluntário não entender a descrição da tarefa, as interfaces do sistema não serem intuitivas ou erro de interpretação dos elementos gráficos da interface da aplicação pelo voluntário.

6.3.1 Falhas de disponibilidade

Ao observar as respostas relacionadas aos erros encontrados do sistema, foi possível identificar um padrão na descrição feita pelos voluntários, pois apenas erros relacionados à disponibilidade da aplicação foram encontrados.

Tais erros estão relacionados ao aplicativo encerrar inesperadamente enquanto o usuário tenta realizar uma determinada tarefa ou uma mensagem de erro é exibida ao usuário informando que o servidor está indisponível e será necessário realizar uma nova tentativa de envio dos dados.

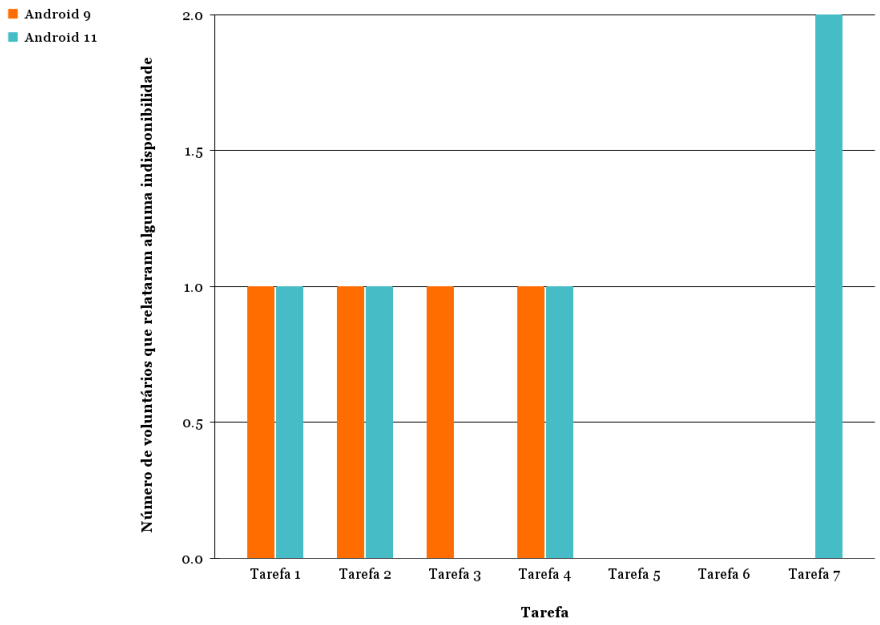
Na Figura 36, o número total de falhas de disponibilidade, por versão do sistema operacional, é apresentado para cada tarefa realizada pelos voluntários. Pode-se observar que o número de falhas encontradas por sistema operacional é muito próximo e falhas de disponibilidade não foram identificadas pelos voluntários que utilizaram o *Android 12*.

Também não foram encontradas falhas de disponibilidade ao realizar as Tarefas 5 ou 6, que respectivamente são: a realização de uma nova captura para um paciente já registrado e a troca de senha pelo usuário. Isto pode significar que talvez tais falhas de disponibilidade sejam esporádicas, o que não necessariamente descarta uma possível falha de compatibilidade entre o dispositivo móvel e o sistema desenvolvido.

Como exemplo, um dos participantes utilizando o *Android 9* reportou falhas de indisponibilidade do servidor em três diferentes tarefas. Essas falhas podem ter sido causadas porque a conexão deste voluntário com a rede mundial de computadores não apresentou-se estável durante o experimento. Dessa forma, em versões futuras a aplicação deve possuir mecanismos para não perder os dados registrados pelo usuário enquanto a conexão deste com a rede mundial de computadores não for estável. Uma possível solução para este problema pode ser aumentar o número de tentativas de conexão entre o dispositivo

e o servidor para cada requisição REST realizada sem sucesso. Outra possibilidade pode ser o desenvolvimento de um mecanismo de armazenamento dos dados no dispositivo do usuário, que apenas são enviados quando a conexão do dispositivo móvel com a rede mundial de computadores se tornar estável.

Figura 36 – Indisponibilidade do sistema desenvolvido durante o experimento.



Fonte: Henrique Silva Bortoletti, 2022.

O número de voluntários que relataram alguma falha de disponibilidade na Tarefa 0 não é apresentado na Figura 36, pois o propósito desta Figura é de apresentar as falhas relacionadas ao sistema desenvolvido e não durante a instalação do aplicativo no dispositivo móvel.

6.3.2 Falhas de interpretação da interface ou tarefas

Também é possível observar um padrão nas respostas dos voluntários relacionadas às dificuldades encontradas em relação às interfaces. Foram identificados dois tipos de dificuldade: os voluntários não entenderam algum elemento visual da interface do sistema enquanto realizavam alguma tarefa ou apresentaram dificuldades para compreender a descrição da tarefa a ser realizada.

Como exemplo, um dos participantes relatou uma falha de interpretação dos elementos visuais da interface, pois ao realizar a captura não compreendeu a necessidade de

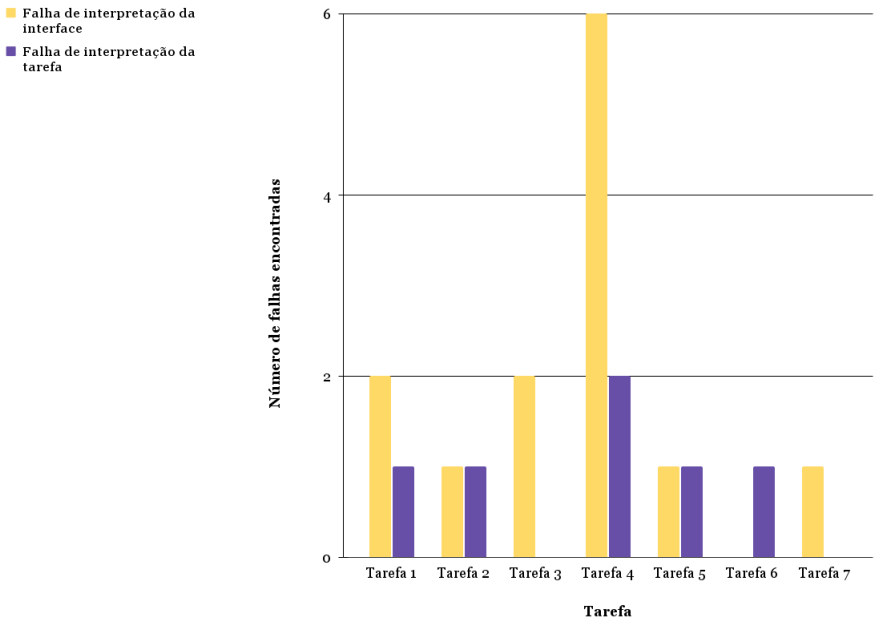
selecionar o círculo inferior da interface para iniciar uma nova captura, conforme descrito na Subseção 5.5. Outro participante relatou uma falha ao interpretar a descrição da Tarefa 6 (realizar uma troca de senha), pois não entendeu a motivação para realizar a troca de senha.

Na Figura 37 os números de falhas devidas a questões de interface e o número de falhas decorrentes de interpretação da tarefa são apresentados para cada tarefa realizada. É possível observar que a maior parte das falhas encontradas para ambos os casos está presente na Tarefa 4, que solicita ao voluntário para realizar a primeira captura de vídeo facial de outro indivíduo. Isso mostra que talvez sejam necessárias melhorias na interface desenvolvida, para que novos usuários, ao realizarem a captura facial pela primeira vez, não encontrem obstáculos de interpretação dos elementos gráficos presentes na interface.

Como exemplo, alguns voluntários relataram na Tarefa 4 que não entenderam como realizar a troca da câmera frontal para a câmera traseira, outros voluntários relataram que não entenderam como selecionar o paciente para a captura e por fim diferentes usuários relataram que não compreenderam os ajustes necessários para encontrar as imagens adequadas. Apesar dos relatos apresentados, pode-se observar um decréscimo nos números de falhas reportadas após a Tarefa 4. Uma possível explicação para este achado é que a repetição de tarefas contribuiu para o aprendizado. Como exemplo, um dos voluntários reportou na Tarefa 5 (realizar uma nova captura para um paciente já registrado), que apenas compreendeu como alterar entre as câmeras traseira e frontal do dispositivo móvel após finalizar a Tarefa 4.

O número de voluntários que relataram alguma falha de interpretação não é apresentado para a Tarefa 0 na Figura 37 porque a questão relacionada às dificuldades encontradas durante o uso da aplicação não foi aplicada para esta tarefa, em razão de descrever o processo de instalação da aplicação.

Figura 37 – Falhas identificadas de interpretação pelos participantes ao utilizar a aplicação ou ao realizar determinada tarefa durante o experimento.



Fonte: Henrique Silva Bortoletti, 2022.

7 Conclusões e trabalhos futuros

O objetivo do presente trabalho foi conceber e desenvolver um sistema de processamento de vídeos faciais capaz de identificar em tempo real imagens adequadas para serem utilizadas em um sistema de auxílio ao diagnóstico do TEA.

Tal objetivo pôde ser alcançado por meio de uma aplicação desenvolvida, a qual poderá ser empregada por especialistas para coletar imagens para o auxílio ao diagnóstico, assim como remotamente por indivíduos que não possuem o acesso a profissionais especializados. Esta também poderá ser adaptada para o auxílio ao diagnóstico de outros distúrbios com base nas características das imagens faciais coletadas pelo dispositivo móvel. Contudo, é importante destacar, que a aplicação não indica ao usuário o resultado do modelo de classificação.

Tal aplicação foi avaliada por diversos profissionais da área da saúde e da computação, que demonstraram conseguir realizar a captura de vídeos contendo as imagens faciais adequadas. Tais indivíduos reportaram inicialmente possuir algumas dificuldades ao utilizarem a aplicação pela primeira vez, porém ao longo das capturas realizadas tal dificuldade foi decaindo gradativamente.

A principal contribuição foi o desenvolvimento do método para a avaliação e seleção em tempo real de imagens faciais adequadas em dispositivos móveis. A avaliação com voluntários mostrou que, em geral, não houve dificuldade uso do aplicativo após o aprendizado das tarefas ser absorvido.

Além dos impactos científicos e tecnológicos que procurou-se evidenciar neste documento, este possui um potencial de impacto social ao abrir possibilidades para disponibilizar o auxílio ao diagnóstico do TEA para regiões sem acesso a uma infraestrutura de saúde especializada. Dessa forma, com um amparo inicial espera-se que os familiares ou pessoas próximas dos indivíduos possam procurar recursos para o diagnóstico definitivo, possibilitando um tratamento precoce do TEA.

Além do impacto social, também é possível identificar um impacto econômico, pois é possível contribuir com o diagnóstico automatizado minimizando o uso de testes com alto custo que exigem profissionais especializados.

A principal conclusão extraída da avaliação conduzida é a curva de aprendizado, que apresentou evidências de que a repetição da tarefa contribui para diminuição da dificuldade em realizá-la.

Dessa forma, espera-se que por meio do método desenvolvido seja possível democratizar o acesso ao diagnóstico e que futuramente seja possível utilizar o método proposto para a detecção de outras síndromes.

7.1 Trabalhos futuros

Considerando as contribuições desta pesquisa e os resultados decorrentes das avaliações da aplicação, nota-se que alguns aspectos podem ser desenvolvidos em trabalhos futuros. Um destes enfoques está relacionado à análise da aplicação para a captura de vídeos faciais de indivíduos com outro tipo de transtorno. Para isso, espera-se que os usuários responsáveis pelas capturas dos vídeos faciais sejam categorizados por meio de uma interface responsável por determinar quais funcionalidades os usuários podem ter acesso.

Dessa forma, determinados usuários podem realizar a captura de imagens faciais, enquanto outros usuários possuem o acesso para administrar os parâmetros de captura de vídeos de imagens faciais. Tais parâmetros podem ser configurados para algum outro tipo de transtorno ou para algum outro tipo de paciente, com isso espera-se alterar o modelo relacional para armazenar a configuração dos parâmetros utilizados por captura realizada.

Outro enfoque a ser desenvolvido é a avaliação do impacto da aplicação pelas pessoas que irão realizar as capturas dos vídeos faciais com uma maior frequência e em regiões geográficas distantes umas das outras. Portanto, espera-se realizar um novo experimento responsável por aplicar a avaliação da aplicação para um número maior de indivíduos e pacientes, que inicialmente nunca tiveram um contato com a aplicação.

Por fim, espera-se que seja possível replicar tal projeto desenvolvido para outros tipos de sistemas operacionais presentes em dispositivos móveis, dessa forma aumentando ainda mais o alcance para a captura de vídeos faciais ao redor do Brasil e do Mundo.

Referências

ABULKHAIR, M.; ALSAHLI, A. H.; TALEB, K. M.; BAHRAN, A. M.; ALZAHRANI, F. M.; ALZAHRANI, H. A.; IBRAHIM, L. F. Mobile platform detect and alerts system for driver fatigue. *Procedia Computer Science*, v. 62, p. 555 – 564, 2015. ISSN 1877-0509. Proceedings of the 2015 International Conference on Soft Computing and Software Engineering (SCSE'15). Citado na página 109.

ALDRIDGE, K.; GEORGE, I.; COLE, K.; AUSTIN, J.; TAKAHASHI, T.; DUAN, Y.; MILES, J. Facial phenotypes in subgroups of prepubertal boys with autism spectrum disorders are correlated with clinical phenotypes. *Molecular autism*, v. 2, p. 15, 10 2011. Citado na página 16.

AN, S.; MA, X.; SONG, R.; LI, Y. Face detection and recognition with surf for human-robot interaction. In: *2009 IEEE International Conference on Automation and Logistics*. [S.l.: s.n.], 2009. p. 1946–1951. ISSN 2161-816X. Citado na página 105.

ANDROID. *What is Android*. 2022. Disponível em: ⟨https://www.android.com/what-is-android/⟩. Citado na página 50.

ANDROIDX. *AndroidX Overview*. 2022. Disponível em: ⟨https://developer.android.com/jetpack/androidx⟩. Citado na página 51.

BALTRUSAITIS, T.; ROBINSON, P.; MORENCY, L.-P. Openface: An open source facial behavior analysis toolkit. *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*, p. 1–10, 2016. Citado na página 61.

BARRA, S.; MARSICO, M. D.; GALDI, C.; RICCIO, D.; WECHSLER, H. FAME: Face authentication for mobile encounter. In: *2013 IEEE Workshop on Biometric Measurements and Systems for Security and Medical Applications*. [S.l.: s.n.], 2013. p. 1–7. Citado na página 108.

BERTOLISSI, E.; PREECE, C. Java in real-time applications. *IEEE Transactions on Nuclear Science*, v. 45, n. 4, p. 1965–1972, 1998. Citado na página 50.

BERTóK, K.; FAZEKAS, A. Face recognition on mobile platforms. In: *2016 7th IEEE International Conference on Cognitive Infocommunications (CogInfoCom)*. [S.l.: s.n.], 2016. p. 000037–000042. Citado na página 105.

BLAIR, R. Facial expressions, their communicatory functions and neuro-cognitive substrates. *Philos Trans R Soc Lond B Biol Sci*, v. 358, n. 1431, p. 561–572, 2003. Citado na página 22.

BLUMBERG, S.; ZABLOTSKY, B.; AVILA, R.; COLPE, L.; PRINGLE, B.; KOGAN, M. Diagnosis lost: Differences between children who had and who currently have an autism spectrum disorder diagnosis. *Autism : the international journal of research and practice*, v. 20, 10 2015. Citado 2 vezes nas páginas 19 e 20.

BOSER, B. E.; GUYON, I. M.; VAPNIK, V. N. A training algorithm for optimal margin classifiers. In: *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*. New York, NY, USA: Association for Computing Machinery, 1992. (COLT '92), p. 144–152. ISBN 089791497X. Citado na página 23.

BRADSKI, G. The OpenCV Library. *Dr. Dobb's Journal of Software Tools*, 2000. Citado 2 vezes nas páginas 41 e 50.

CALLEGATI, F.; CERRONI, W.; RAMILLI, M. Man-in-the-middle attack to the HTTPS protocol. *IEEE Security Privacy*, v. 7, n. 1, p. 78–81, 2009. Citado na página 59.

CANEDO, D.; NEVES, A. Facial expression recognition using computer vision: A systematic review. *Applied Sciences*, v. 9, n. 21, 2019. ISSN 2076-3417. Citado 3 vezes nas páginas 22, 23 e 24.

CHASTE, P.; KLEI, L.; SANDERS, S. J.; MURTHA, M. T.; HUS, V.; LOWE, J. K.; WILLSEY, A. J.; MORENO-DE-LUCA, D.; YU, T. W.; FOMBONNE, E.; GESCHWIND, D.; GRICE, D. E.; LEDBETTER, D. H.; LORD, C.; MANE, S. M.; MARTIN, C. L.; MARTIN, D. M.; MORROW, E. M.; WALSH, C. A.; SUTCLIFFE, J. S.; STATE, M. W.; DEVLIN, B.; COOK, E. H.; KIM, S.-J. Adjusting head circumference for covariates in autism: Clinical correlates of a highly heritable continuous trait. *Biological Psychiatry*, v. 74, n. 8, p. 576–584, 2013. ISSN 0006-3223. Attention-Deficit/Hyperactivity Disorder: Brain Development and Circuit Function. Citado na página 20.

CHAUDHRY, S.; CHANDRA, R. Face detection and recognition in an unconstrained environment for mobile visual assistive system. *Applied Soft Computing*, v. 53, p. 168 – 180, 2017. ISSN 1568-4946. Citado 3 vezes nas páginas 33, 34 e 108.

CHEN, C.-P.; CHEN, C.-S. Intrinsic illumination subspace for lighting insensitive face recognition. *IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics)*, v. 42, n. 2, p. 422–433, April 2012. ISSN 1941-0492. Citado 3 vezes nas páginas 33, 35 e 105.

CHEN, S.; LU, X.; SUN, L.; GE, S. Target domain adaptation for face detection in a smart camera network with peer-to-peer communications. In: *2015 IEEE Global Communications Conference (GLOBECOM)*. [S.l.: s.n.], 2015. p. 1–6. Citado na página 106.

CHEN, T.; YIN, W.; ZHOU, X.; COMANICIU, D. Total variation models for variable lighting face recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 28, n. 9, p. 1519–1524, 2006. Citado na página 22.

CHEN, Y.; YU, Y.; ODOBEZ, J.-M. Head nod detection from a full 3d model. In: . [S.l.: s.n.], 2015. p. 528–536. Citado na página 54.

CHENG, B.; WANG, Z.; ZHANG, Z.; LI, Z.; LIU, D.; YANG, J.; HUANG, S. Robust emotion recognition from low quality and low bit rate video: A deep learning approach. 09 2017. Citado na página 22.

CHENG, L.-P.; HSIAO, F.-I.; LIU, Y.-T.; CHEN, M. Y. IRotate: Automatic screen rotation based on face orientation. In: *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems*. New York, NY, USA: Association for Computing Machinery, 2012. (CHI '12), p. 2203–2210. ISBN 9781450310154. Citado 2 vezes nas páginas 28 e 105.

CHIA, S.-C.; MANDAL, B.; XU, Q.; LI, L.; LIM, J.-H. Enhancing social interaction with seamless face recognition on google glass: Leveraging opportunistic multi-tasking on smart phones. In: *Proceedings of the 17th International Conference on Human-Computer*

Interaction with Mobile Devices and Services Adjunct. New York, NY, USA: Association for Computing Machinery, 2015. (MobileHCI '15), p. 750–757. ISBN 9781450336536. Citado na página 105.

CHOI, K.; TOH, K.-A.; BYUN, H. Realtime training on mobile devices for face recognition applications. *Pattern Recognition*, v. 44, n. 2, p. 386 – 400, 2011. ISSN 0031-3203. Citado na página 105.

CRUZ, C.; SUCAR, L. E.; MORALES, E. F. Real-time face recognition for human-robot interaction. In: *2008 8th IEEE International Conference on Automatic Face Gesture Recognition*. [S.l.: s.n.], 2008. p. 1–6. Citado na página 105.

EKMAN, P.; FRIESEN, W.; O'SULLIVAN, M.; CHAN, A.; DIACOYANNI-TARLATZIS, I.; HEIDER, K.; KRAUSE, R.; LECOMPTE, W.; PITCAIRN, T.; BITTI, P. R. Universals and cultural differences in the judgments of facial expressions of emotion. *Journal of personality and social psychology*, v. 53, p. 712–7, 11 1987. Citado na página 22.

EKMAN, P.; ROSENBERG, E. L. *What the face reveals: Basic and applied studies of spontaneous expression using the facial action coding system (FACS), 2nd ed.* [S.l.]: New York, NY, US: Oxford University Press, 2005. Citado na página 23.

ELREFAEI, L. A.; ALHARTHI, A.; ALAMOUDI, H.; ALMUTAIRI, S.; AL-RAMMAH, F. Real-time face detection and tracking on mobile phones for criminal detection. In: *2017 2nd International Conference on Anti-Cyber Crimes (ICACC)*. [S.l.: s.n.], 2017. p. 75–80. Citado 2 vezes nas páginas 31 e 108.

FARKAS, L. G. *Anthropometry of the head and face*. [S.l.]: New York, Elsevier, 1982. Citado 2 vezes nas páginas 20 e 54.

FATHY, M. E.; PATEL, V. M.; CHELLAPPA, R. Face-based active authentication on mobile devices. In: *2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.: s.n.], 2015. p. 1687–1691. ISSN 2379-190X. Citado na página 107.

FAUDZI, S. A. A. M.; YAHYA, N. Evaluation of lbp-based face recognition techniques. In: *2014 5th International Conference on Intelligent and Advanced Systems (ICIAS)*. [S.l.: s.n.], 2014. p. 1–6. Citado na página 33.

FERRY, Q.; STEINBERG, J.; WEBBER, C.; FITZPATRICK, D.; PONTING, C.; ZISSERMAN, A.; NELLAKER, C. Diagnostically relevant facial gestalt information from ordinary photos. *eLife*, v. 3, 2014. Citado na página 45.

FILIP, P.; ČEGAN, L. Comparison of mysql and mongodb with focus on performance. In: *2020 International Conference on Informatics, Multimedia, Cyber and Information System (ICIMCIS)*. [S.l.: s.n.], 2020. p. 184–187. Citado na página 51.

FINCH, P. J. M. *A Study of the Blowfish Encryption Algorithm*. Tese (Doutorado) — University of New York, USA, 1995. UMI Order No. GAX95-21269. Citado na página 71.

FOGGIA, P.; GRECO, A.; PERCANNELLA, G.; VENTO, M.; VIGILANTE, V. A system for gender recognition on mobile robots. In: *Proceedings of the 2nd International Conference on Applications of Intelligent Systems*. New York, NY, USA: Association for

Computing Machinery, 2019. (APPIS '19), p. 6. ISBN 9781450360852. Citado 3 vezes nas páginas 31, 33 e 109.

FRITH, U.; HAPPé, F. Autism spectrum disorder. *Current Biology*, Elsevier, v. 15, n. 19, p. R786–R790, Oct 2005. ISSN 0960-9822. Citado 2 vezes nas páginas 19 e 20.

GENEVA, S. *Ergonomic Requirements for Office Work With Visual Display Terminals (VDTs)—Part 11: Guidance on Usability*. 2022. Disponível em: ⟨https://www.iso.org/obp/ui/#iso:std:iso:9241:-11:ed-1:v1:en⟩. Citado na página 64.

GHAZAL, M.; HAEYEH, Y. A.; ABED, A.; GHAZAL, S. Embedded fatigue detection using convolutional neural networks with mobile integration. In: *2018 6th International Conference on Future Internet of Things and Cloud Workshops (FiCloudW)*. [S.l.: s.n.], 2018. p. 129–133. Citado 3 vezes nas páginas 29, 31 e 108.

GOOGLE. *Detect faces with ML Kit on Android*. 2022. Disponível em: ⟨h5ps://developers.google.com/ml-kit/vision/face-detection/android⟩. Citado 3 vezes nas páginas 54, 55 e 56.

GROSS RALPHAND BRAJOVIC, V. An image preprocessing algorithm for illumination invariant face recognition. In: KITTLER JOSEFAND NIXON, M. S. (Ed.). *Audio- and Video-Based Biometric Person Authentication*. Berlin, Heidelberg: Springer Berlin Heidelberg, 2003. p. 10–18. Citado na página 22.

HAPPY, S. L.; ROUTRAY, A. Automatic facial expression recognition using features of salient facial patches. *IEEE Transactions on Affective Computing*, v. 6, 05 2015. Citado 3 vezes nas páginas 22, 23 e 24.

HELLMUND, T.; SEITZ, A.; HALADJIAN, J.; BRUEGGE, B. IPRA: Real-time face recognition on smart glasses with fog computing. In: *Proceedings of the 2018 ACM International Joint Conference and 2018 International Symposium on Pervasive and Ubiquitous Computing and Wearable Computers*. New York, NY, USA: Association for Computing Machinery, 2018. p. 988–993. ISBN 9781450359665. Citado na página 106.

HETTIARACHCHI, S. *Analysis of different face detection andrecognition models for Android*. 73 p. Dissertação (Mestrado) — Mid Sweden University, Department of Information Systems and Technology, 2021. Citado 2 vezes nas páginas 50 e 53.

HJELMåS, E.; LOW, B. Face detection: A survey. *Computer Vision and Image Understanding*, v. 83, p. 236–274, 09 2001. Citado na página 21.

HU, J.-Y.; SUENG, C.-C.; LIAO, W.-H.; HO, C. C. Android-based mobile payment service protected by 3-factor authentication and virtual private ad hoc networking. In: *2012 Computing, Communications and Applications Conference*. [S.l.: s.n.], 2012. p. 111–116. Citado na página 107.

HUANG, C.; YANG, X.; CHENG, K.-T. T. Accurate and efficient pulse measurement from facial videos on smartphones. In: *2016 IEEE Winter Conference on Applications of Computer Vision (WACV)*. [S.l.: s.n.], 2016. p. 1–8. Citado 4 vezes nas páginas 29, 31, 35 e 109.

HUANG, G. B.; RAMESH, M.; BERG, T.; LEARNED-MILLER, E. *Labeled Faces in the Wild: A Database for Studying Face Recognition in Unconstrained Environments*. [S.l.], 2007. Citado 2 vezes nas páginas 36 e 37.

HULSEBOSCH, R.; EBBEN, P. Enhancing face recognition with location information. In: *2008 Third International Conference on Availability, Reliability and Security*. [S.l.: s.n.], 2008. p. 397–403. Citado na página 108.

JEON, J.; PARK, J.-C.; JO, Y.; NAM, C.; BAE, K.-H.; HWANG, Y.; KIM, D.-S. A real-time facial expression recognizer using deep neural network. In: *Proceedings of the 10th International Conference on Ubiquitous Information Management and Communication*. New York, NY, USA: Association for Computing Machinery, 2016. (IMCOM '16), p. 4. ISBN 9781450341424. Citado 2 vezes nas páginas 33 e 109.

JIN, Y.; KIM, J.; KIM, B.; MALLIPEDDI, R.; LEE, M. Smart cane: Face recognition system for blind. In: *Proceedings of the 3rd International Conference on Human-Agent Interaction*. New York, NY, USA: Association for Computing Machinery, 2015. p. 145–148. ISBN 9781450335270. Citado na página 108.

JOHANSSEN, J. O.; BERNIUS, J. P.; BRUEGGE, B. Toward usability problem identification based on user emotions derived from facial expressions. In: *2019 IEEE/ACM 4th International Workshop on Emotion Awareness in Software Engineering (SEmotion)*. [S.l.: s.n.], 2019. p. 1–7. Citado na página 106.

JUNQUEIRA, A. A. J.; MAGRI, L. V.; MELCHIOR, M. d. O.; SILVA, A. M. B. R. D.; SILVA, M. A. M. R. D. Facial anthropometric analysis of a healthy group of young brazilian adults by means of stereophotogrammetry technique. *Revista de Odontologia da UNESP*, scielo, v. 45, p. 139 – 145, 06 2016. ISSN 1807-2577. Citado na página 20.

KALANTARIAN, H.; JEDOUI, K.; WASHINGTON, P.; TARIQ, Q.; DUNLAP, K.; SCHWARTZ, J.; WALL, D. P. Labeling images with facial emotion and the potential for pediatric healthcare. *Artificial Intelligence in Medicine*, v. 98, p. 77 – 86, 2019. ISSN 0933-3657. Citado 2 vezes nas páginas 32 e 109.

KANTHARIA, K. J.; PRAJAPATI, G. I. Facial behavior recognition using soft computing techniques: A survey. In: *2015 Fifth International Conference on Advanced Computing Communication Technologies*. [S.l.: s.n.], 2015. p. 30–34. Citado na página 21.

KING, D. E. Dlib-ml: A machine learning toolkit. *Journal of Machine Learning Research*, v. 10, p. 1755–1758, 2009. Citado 2 vezes nas páginas 35 e 41.

KODIROV, E.; PN, A. F.; LEE, G.-S.; CHOI, D.-J.; NA, I. S. Robust real time face tracking in mobile devices. In: *Proceedings of the 7th International Conference on Ubiquitous Information Management and Communication*. New York, NY, USA: Association for Computing Machinery, 2013. (ICUIMC '13), p. 8. ISBN 9781450319584. Citado 2 vezes nas páginas 33 e 108.

KURTANOVIć, Z.; MAALEJ, W. Automatically classifying functional and non-functional requirements using supervised machine learning. In: *2017 IEEE 25th International Requirements Engineering Conference (RE)*. [S.l.: s.n.], 2017. p. 490–495. Citado na página 47.

LEE, T.; PARK, S.-K.; PARK, M. An effective method for detecting facial features and face in human–robot interaction. *Information Sciences*, v. 176, n. 21, p. 3166 – 3189, 2006. ISSN 0020-0255. Citado 2 vezes nas páginas 29 e 106.

LI, A.; MORARIU, V. I.; DAVIS, L. S. Selective encoding for recognizing unreliably localized faces. In: *2015 IEEE International Conference on Computer Vision (ICCV)*. [S.l.: s.n.], 2015. p. 3613–3621. ISSN 2380-7504. Citado na página 105.

LIU, W.; WANG, Z.; LIU, X.; ZENG, N.; LIU, Y.; ALSAADI, F. E. A survey of deep neural network architectures and their applications. *Neurocomputing*, v. 234, p. 11–26, 2017. ISSN 0925-2312. Citado na página 23.

LIU, Z.; ZHU, X.; TANG, M.; LEI, Z.; WANG, J. Efficient face alignment with fast normalization and contour fitting loss. *ACM Trans. Multimedia Comput. Commun. Appl.*, Association for Computing Machinery, New York, NY, USA, v. 15, n. 3s, p. 16, November 2019. ISSN 1551-6857. Citado 3 vezes nas páginas 28, 33 e 107.

LORD, C.; ELSABBAGH, M.; BAIRD, G.; VEENSTRA-VANDERWEELE, J. Autism spectrum disorder. *The Lancet*, v. 392, 08 2018. Citado na página 19.

LUNERTI, C.; GUEST, R. M.; BLANCO-GONZALO, R.; SANCHEZ-REILLO, R.; BAKER, J. Environmental effects on face recognition in smartphones. In: *2017 International Carnahan Conference on Security Technology (ICCST)*. [S.l.: s.n.], 2017. p. 1–6. Citado na página 106.

MA, J.; FAN, X.; NI, J.; ZHU, X.; XIONG, C. Multi-scale retinex with color restoration image enhancement based on gaussian filtering and guided filtering. *International Journal of Modern Physics B*, v. 31, p. 1744077, 05 2017. Citado na página 22.

MA, T.-Y.; LIN, C.-Y.; HSU, S.-W.; HU, C.-W.; HOU, T.-W. Automatic brightness control of the handheld device display with low illumination. In: *2012 IEEE International Conference on Computer Science and Automation Engineering (CSAE)*. [S.l.: s.n.], 2012. v. 2, p. 382–385. Citado 2 vezes nas páginas 30 e 106.

MAHBUB, U.; SARKAR, S.; CHELLAPPA, R. Partial face detection in the mobile domain. *Image and Vision Computing*, v. 82, p. 1 – 17, 2019. ISSN 0262-8856. Citado 4 vezes nas páginas 29, 33, 37 e 106.

MARIA, D. M.; GALDI, C.; NAPPI, M.; RICCIO, D. FIRME: Face and iris recognition for mobile engagement. *Image and Vision Computing*, v. 32, n. 12, p. 1161 – 1172, 2014. ISSN 0262-8856. Citado 2 vezes nas páginas 33 e 108.

MCREYNOLDS, T.; BLYTHE, D. CHAPTER 3 - color, shading, and lighting. In: MCREYNOLDS, T.; BLYTHE, D. (Ed.). *Advanced Graphics Programming Using OpenGL*. San Francisco: Morgan Kaufmann, 2005, (The Morgan Kaufmann Series in Computer Graphics). p. 35 – 56. ISBN 978-1-55860-659-3. Citado 2 vezes nas páginas 22 e 34.

MITTAL, S.; AGARWAL, S.; NIGAM, M. J. Real time multiple face recognition: A deep learning approach. In: *Proceedings of the 2018 International Conference on Digital Medicine and Image Processing*. New York, NY, USA: Association for Computing Machinery, 2018. (DMIP '18), p. 70–76. ISBN 9781450365789. Citado 3 vezes nas páginas 29, 30 e 109.

MUKHERJEE, D.; DAS, R.; MAJUMDAR, S.; GHOSH, S.; THEPADE, S.; BASU, A. Energy efficient face recognition in mobile-fog environment. *Procedia Computer Science*, v. 152, p. 274 – 281, 2019. ISSN 1877-0509. International Conference on Pervasive

Computing Advances and Applications- PerCAA 2019. Citado 2 vezes nas páginas 31 e 108.

MUSLIM, N.; ISLAM, S. Face recognition in the edge cloud. In: *Proceedings of the International Conference on Imaging, Signal Processing and Communication*. New York, NY, USA: Association for Computing Machinery, 2017. (ICISPC 2017), p. 5–9. ISBN 9781450352895. Citado 3 vezes nas páginas 16, 29 e 107.

MY, V. D.; ZELL, A. Real time face tracking and pose estimation using an adaptive correlation filter for human-robot interaction. In: *2013 European Conference on Mobile Robots*. [S.l.: s.n.], 2013. p. 119–124. Citado na página 106.

NAMBIAPPAN, H. R.; DATTA, S. PISS-IoT: Person identification and spotting system in an internet-of-things way. In: *Proceedings of the 11th PErvasive Technologies Related to Assistive Environments Conference*. New York, NY, USA: Association for Computing Machinery, 2018. (PETRA '18), p. 8. ISBN 9781450363907. Citado 2 vezes nas páginas 35 e 107.

NI, Z.; LI, Q. Fusion learning model for mobile face safe detection and facial gesture analysis. *IEEE Access*, v. 8, p. 61043–61050, 2020. ISSN 2169-3536. Citado na página 107.

PANCHAL, K.; SHAH, H. 3d face recognition based on pose correction using euler angle method. In: *2013 International Conference on Machine Intelligence and Research Advancement*. [S.l.: s.n.], 2013. p. 467–471. Citado 2 vezes nas páginas 50 e 54.

PASCHALAKIS, S.; BOBER, M. Real-time face detection and tracking for mobile videoconferencing. *Real-Time Imaging*, v. 10, n. 2, p. 81 – 94, 2004. ISSN 1077-2014. Citado na página 107.

PENG, P.; SHEN, Y. Efficient face verification in mobile environment using component-based PCA. In: *2013 6th International Congress on Image and Signal Processing (CISP)*. [S.l.: s.n.], 2013. v. 2, p. 753–757. Citado na página 108.

PHAM, P.; WANG, J. Attentivevideo: A multimodal approach to quantify emotional responses to mobile advertisements. *ACM Trans. Interact. Intell. Syst.*, Association for Computing Machinery, New York, NY, USA, v. 9, n. 2–3, p. 30, March 2019. ISSN 2160-6455. Citado 2 vezes nas páginas 37 e 105.

PINHEIROS, T. D. Classificação de imagens faciais para o auxílio ao diagnóstico do transtorno do espectro autista. 2018. Citado 10 vezes nas páginas 8, 16, 17, 40, 42, 51, 52, 55, 61 e 63.

POWELL, B. M.; GOSWAMI, G.; VATSA, M.; SINGH, R.; NOORE, A. fgCAPTCHA: Genetically optimized face image captcha 5. *IEEE Access*, v. 2, p. 473–484, 2014. Citado na página 106.

PRADO, K. S.; ROMAN, N. T.; SILVA, V. F.; BERNARDES, J. L.; DIGIAMPIETRI, L. A.; ORTEGA, E. M.; LIMA, C. A. M.; CURA, L. M. V.; ANTUNES, M. M. Automatic facial recognition: A systematic review on the problem of light variation. In: KUROSU, M. (Ed.). *Human-Computer Interaction. Interaction Platforms and Techniques*. Cham: Springer International Publishing, 2016. p. 211–221. Citado na página 21.

RAHMAN, M.; KEHTARNAVAZ, N.; REN, J. A hybrid face detection approach for real-time depolyment on mobile devices. In: *2009 16th IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2009. p. 3233–3236. ISSN 2381-8549. Citado na página 106.

RAO, A.; S., V.; C., A.; MURTHY, K. B.; SUBRAMANYAM; A., V. Feature extractionusing ORB-RANSAC for face recognition. *Procedia Computer Science*, v. 70, p. 174 – 184, 2015. ISSN 1877-0509. Citado na página 105.

RATTANI, A.; DERAKHSHANI, R. A survey of mobile face biometrics. *Computers & Electrical Engineering*, v. 72, p. 39 – 52, 2018. ISSN 0045-7906. Citado na página 21.

REN, J.; JIANG, X.; YUAN, J. A complete and fully automated face verification system on mobile devices. *Pattern Recognition*, v. 46, n. 1, p. 45–56, 2013. ISSN 0031-3203. Citado na página 105.

RETROFIT. *Retrofit*. 2022. Disponível em: ⟨https://square.github.io/retrofit/⟩. Citado na página 51.

RIESCH, P. J.; DU, X.; LING, H.; MAYHEW, M. J. Face recognition with environment tolerance on a mobile device. In: *2015 IEEE 2nd International Conference on Cyber Security and Cloud Computing*. [S.l.: s.n.], 2015. p. 342–348. Citado na página 108.

RODIER, P.; BRYSON, S.; WELCH, J. Minor malformations and physical measurements in autism: data from nova scotia. *Teratology*, v. 55, n. 5, p. 319—325, May 1997. ISSN 0040-3709. Citado na página 16.

ROSS, A.; WILLSON, V. L. Paired samples t-test. In: _____. *Basic and Advanced Statistical Tests: Writing Results Sections and Creating Tables and Figures*. Rotterdam: SensePublishers, 2017. p. 17–19. ISBN 978-94-6351-086-8. Disponível em: ⟨https://doi.org/10.1007/978-94-6351-086-8_4⟩. Citado na página 43.

SACCO, R.; GABRIELE, S.; PERSICO, A. M. Head circumference and brain size in autism spectrum disorder: A systematic review and meta-analysis. *Psychiatry Research: Neuroimaging*, v. 234, n. 2, p. 239–251, 2015. ISSN 0925-4927. Citado na página 20.

SAMANGOUEI, P.; PATEL, V. M.; CHELLAPPA, R. Facial attributes for active authentication on mobile devices. *Image and Vision Computing*, v. 58, p. 181 – 192, 2017. ISSN 0262-8856. Citado na página 108.

SARKAR, S.; PATEL, V. M.; CHELLAPPA, R. Deep feature-based face detection on mobile devices. In: *2016 IEEE International Conference on Identity, Security and Behavior Analysis (ISBA)*. [S.l.: s.n.], 2016. p. 1–8. Citado na página 106.

SCHNEIDER, C.; ESAU, N.; KLEINJOHANN, L.; KLEINJOHANN, B. Feature based face localization and recognition on mobile devices. In: *2006 9th International Conference on Control, Automation, Robotics and Vision*. [S.l.: s.n.], 2006. p. 1–6. Citado na página 107.

SHEN, Y.; YANG, M.; WEI, B.; CHOU, C. T.; HU, W. Learn to recognise: Exploring priors of sparse face recognition on smartphones. *IEEE Transactions on Mobile Computing*, v. 16, n. 6, p. 1705–1717, 2017. Citado na página 105.

SHI, F.; TSAI, S.-F.; WANG, Y.; LIANG, C.-K. Steadiface: Real-time face-centric stabilization on mobile phones. In: *2019 IEEE International Conference on Image Processing (ICIP)*. [S.l.: s.n.], 2019. p. 4599–4603. ISSN 2381-8549. Citado na página 106.

SIM, T.; BAKER, S.; BSAT, M. The CMU pose, illumination, and expression database. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, v. 25, n. 12, p. 1615 – 1618, December 2003. Citado na página 22.

SMITH-CREASEY, M.; ALBALOOSHI, F. A.; RAJARAJAN, M. Continuous face authentication scheme for mobile devices with tracking and liveness detection. *Microprocessors and Microsystems*, v. 63, p. 147–157, 2018. ISSN 0141-9331. Citado 2 vezes nas páginas 34 e 105.

SMITH, M.; ALBALOOSHI, F. A.; RAJARAJAN, M. Context awareness for improved continuous face authentication on mobile devices. In: *2018 IEEE 16th Intl Conf on Dependable, Autonomic and Secure Computing, 16th Intl Conf on Pervasive Intelligence and Computing, 4th Intl Conf on Big Data Intelligence and Computing and Cyber Science and Technology Congress(DASC/PiCom/DataCom/CyberSciTech)*. [S.l.: s.n.], 2018. p. 644–652. Citado 3 vezes nas páginas 34, 35 e 107.

SPRING. *Spring*. 2022. Disponível em: ⟨https://spring.io/⟩. Citado na página 51.

SUK, M.; PRABHAKARAN, B. Real-time mobile facial expression recognition system – a case study. In: *2014 IEEE Conference on Computer Vision and Pattern Recognition Workshops*. [S.l.: s.n.], 2014. p. 132–137. ISSN 2160-7516. Citado na página 107.

SUNARYONO, D.; SISWANTORO, J.; ANGGORO, R. An android based course attendance system using face recognition. *Journal of King Saud University - Computer and Information Sciences*, 2019. ISSN 1319-1578. Citado na página 106.

SUNG, K.-k. *Learning and Example Selection for Object and Pattern Recognition*. Tese (Doutorado) — MIT, Artificial Intelligence Laboratory and Center for Biological and Computational Learning, Cambridge, MA, 1996. Citado na página 46.

SZKUDLAREK, M.; PIETRUSZKA, M. Fast gpu and cpu computing for head position estimation. In: *2015 Federated Conference on Computer Science and Information Systems (FedCSIS)*. [S.l.: s.n.], 2015. p. 231–240. Citado na página 107.

TAKAHASHI, K.; MITSUKURA, Y. Head pose tracking based on optimizing normalized cross-correlation. In: *The SICE Annual Conference 2013*. [S.l.: s.n.], 2013. p. 2513–2517. Citado na página 106.

TAN, X.; TRIGGS, B. Enhanced local texture feature sets for face recognition under difficult lighting conditions. *IEEE Transactions on Image Processing*, v. 19, n. 6, p. 1635–1650, 2010. Citado 3 vezes nas páginas 21, 22 e 34.

TANG, S.; XU, Y.; LIU, X.; CHEN, Z.; ZHOU, Y.; NIE, L.; HE, L. Quantitative susceptibility mapping shows lower brain iron content in children with autism. *European Radiology*, v. 31, n. 4, p. 2073–2083, Apr 2021. ISSN 1432-1084. Disponível em: ⟨https://doi.org/10.1007/s00330-020-07267-w⟩. Citado na página 20.

TAO, Q.; VELDHUIS, R. Biometric authentication system on mobile personal devices. *IEEE Transactions on Instrumentation and Measurement*, v. 59, n. 4, p. 763–773, April 2010. ISSN 1557-9662. Citado na página 107.

TAPU, R.; MOCANU, B.; ZAHARIA, T. Face recognition in video streams for mobile assistive devices dedicated to visually impaired. In: *2018 14th International Conference on Signal-Image Technology Internet-Based Systems (SITIS)*. [S.l.: s.n.], 2018. p. 137–142. Citado na página 108.

TELLEZ, N.; JIMENO, M.; SALAZAR, A.; NINO-RUIZ, E. D. Container-based architecture for optimal face-recognition tasks in edge computing. In: *Proceedings of the 4th ACM/IEEE Symposium on Edge Computing*. New York, NY, USA: Association for Computing Machinery, 2019. (SEC '19), p. 3. ISBN 9781450367332. Citado na página 105.

TIAN, Q.; ARBEL, T.; CLARK, J. J. Structured deep fisher pruning for efficient facial trait classification. *Image and Vision Computing*, v. 77, p. 45 – 59, 2018. ISSN 0262-8856. Citado 2 vezes nas páginas 37 e 107.

TRIANTAFYLLIDOU, D.; TEFAS, P. N.; ANASTASIOS. Fast deep convolutional face detection in the wild exploiting hard sample mining. *Big Data Research*, v. 11, p. 65 – 76, 2018. ISSN 2214-5796. Selected papers from the 2nd INNS Conference on Big Data: Big Data & Neural Networks. Citado na página 106.

V, S.; G, V.; A, B.; P, C.; KP, B.; M, P. A fast and robust emotion recognition system for real-world mobile phone data. In: *2015 IEEE International Conference on Multimedia Expo Workshops (ICMEW)*. [S.l.: s.n.], 2015. p. 1–6. Citado na página 107.

VANITHA, A.; VAIDEHI, V.; VASUHI, S. Liveliness detection in real time videos using color based chromatic moment feature. In: *2018 International Conference on Recent Trends in Advance Computing (ICRTAC)*. [S.l.: s.n.], 2018. p. 162–167. Citado 3 vezes nas páginas 31, 33 e 108.

VEZZETTI, E.; MARCOLIN, F. 3d human face description: landmarks measures and geometrical features. *Image and Vision Computing*, v. 30, n. 10, p. 698–712, 2012. ISSN 0262-8856. 3D Facial Behaviour Analysis and Understanding. Citado na página 20.

VIOLA, P. A.; JONES, M. J. Rapid object detection using a boosted cascade of simple features. In: *CVPR (1)*. [S.l.]: IEEE Computer Society, 2001. p. 511–518. ISBN 0-7695-1272-0. Citado 5 vezes nas páginas 23, 31, 32, 33 e 35.

WANG, H.; LI, S.; WANG, Y. Face recognition under varying lighting conditions using self quotient image. In: . [S.l.: s.n.], 2004. p. 819– 824. ISBN 0-7695-2122-3. Citado na página 22.

WANG, J.; AMOS, B.; DAS, A.; PILLAI, P.; SADEH, N.; SATYANARAYANAN, M. A scalable and privacy-aware iot service for live video analytics. In: *Proceedings of the 8th ACM on Multimedia Systems Conference*. New York, NY, USA: Association for Computing Machinery, 2017. (MMSys'17), p. 38–49. ISBN 9781450350020. Citado 3 vezes nas páginas 33, 36 e 108.

WANG, P.; LIN, W.-H.; CHAO, K.-M.; LO, C.-C. A face-recognition approach using deep reinforcement learning approach for user authentication. In: *2017 IEEE 14th International Conference on e-Business Engineering (ICEBE)*. [S.l.: s.n.], 2017. p. 183–188. Citado na página 107.

WANG, Z.; YAN, J.; PANG, C.; CHU, D.; AGHAJAN, H. Who is here: Location aware face recognition. In: *Proceedings of the Third International Workshop on Sensing Applications on Mobile Phones*. New York, NY, USA: Association for Computing Machinery, 2012. (PhoneSense '12), p. 5. ISBN 9781450317788. Citado 2 vezes nas páginas 28 e 107.

WASHINGTON, P.; VOSS, C.; KLINE, A.; HABER, N.; DANIELS, J.; FAZEL, A.; DE, T.; FEINSTEIN, C.; WINOGRAD, T.; WALL, D. Superpowerglass: A wearable aid for the at-home therapy of children with autism. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.*, Association for Computing Machinery, New York, NY, USA, v. 1, n. 3, set. 2017. Citado 4 vezes nas páginas 29, 30, 37 e 107.

WEICHBROTH, P. Usability of mobile applications: A systematic literature study. *IEEE Access*, v. 8, p. 55563–55577, 2020. Citado 2 vezes nas páginas 64 e 65.

WEINBERG, S.; NAIDOO, S.; GOVIER, D.; MARTIN, R.; KANE, A.; MARAZITA, M. Anthropometric precision and accuracy of digital three-dimensional photogrammetry. *The Journal of craniofacial surgery*, v. 17, p. 477–83, 05 2006. Citado na página 16.

WELIVITA, A.; NIMALSIRI, N.; WICKRAMASINGHE, R.; PATHIRANA, U.; GAMAGE, C. Quantitative evaluation of face detection and tracking algorithms for head pose estimation in mobile platforms. In: *2017 Moratuwa Engineering Research Conference (MERCon)*. [S.l.: s.n.], 2017. p. 310–315. Citado na página 106.

WIJAYA, L.; SAVVIDES, M.; KUMAR, B. The effect of low-bit rate JPEG2000 wavelet images on face authentication using advanced correlation filters. In: *Conference Record of the Thirty-Eighth Asilomar Conference on Signals, Systems and Computers, 2004*. [S.l.: s.n.], 2004. v. 1, p. 592–596 Vol.1. Citado na página 105.

WILHELM, T.; BöHME, H.-J.; GROSS, H.-M. A multi-modal system for tracking and analyzing faces on a mobile robot. *Robotics and Autonomous Systems*, v. 48, n. 1, p. 31 – 40, 2004. ISSN 0921-8890. European Conference on Mobile Robots (ECMR '03). Citado 2 vezes nas páginas 33 e 107.

WONG, R.; POH, N.; KITTLER, J.; FROHLICH, D. Towards inclusive design in mobile biometry. In: *3rd International Conference on Human System Interaction*. [S.l.: s.n.], 2010. p. 267–274. ISSN 2158-2254. Citado 2 vezes nas páginas 30 e 108.

WU, X.; XU, J.; WANG, J.; LI, Y.; LI, W.; GUO, Y. Identity authentication on mobile devices using face verification and ID image recognition. *Procedia Computer Science*, v. 162, p. 932 – 939, 2019. ISSN 1877-0509. Citado na página 108.

XIONG, Q.; LIN, J.; YUE, W.; LIU, S.; LIU, Y.; DING, C. A deep learning approach to driver distraction detection of using mobile phone. In: *2019 IEEE Vehicle Power and Propulsion Conference (VPPC)*. [S.l.: s.n.], 2019. p. 1–5. ISSN 1938-8756. Citado 2 vezes nas páginas 31 e 109.

XU, L.; LI, J.; WANG, K. Real-time and multi-view face tracking on mobile platform. In: *2011 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. [S.l.: s.n.], 2011. p. 1485–1488. Citado na página 106.

YANG, S.; LUO, P.; LOY, C. C.; TANG, X. WIDER FACE: A face detection benchmark. In: *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*. [S.l.: s.n.], 2016. Citado 2 vezes nas páginas 16 e 21.

ZHANG, C.; ZHANG, Z. *A Survey of Recent Advances in Face Detection*. [S.l.: s.n.], 2010. 1-17 p. Citado na página 34.

ZHANG, D.; HU, M.; JI, Q. Financial markets under the global pandemic of COVID-19. *Finance Research Letters*, v. 36, p. 101528, 2020. ISSN 1544-6123. Citado na página 16.

ZHANG, H.; WANG, X.; ZHU, J.; KUO, C.-C. J. Fast face detection on mobile devices by leveraging global and local facial characteristics. *Signal Processing: Image Communication*, v. 78, p. 1–8, 2019. ISSN 0923-5965. Citado na página 106.

ZHANG, K.; ZHANG, Z.; LI, Z.; QIAO, Y. Joint face detection and alignment using multitask cascaded convolutional networks. *IEEE Signal Processing Letters*, v. 23, n. 10, p. 1499–1503, Oct 2016. ISSN 1558-2361. Citado na página 16.

ZHANG, X.; JING, K.; DAI, Y.; XU, X. Face biometric identity authentication system. In: *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*. [S.l.: s.n.], 2018. p. 1473–1477. Citado 3 vezes nas páginas 32, 33 e 108.

ZHANG, Y.; XIAO, X.; YANG, L.-X.; XIANG, Y.; ZHONG, S. Secure and efficient outsourcing of PCA-Based face recognition. *IEEE Transactions on Information Forensics and Security*, v. 15, p. 1683–1695, 2020. ISSN 1556-6021. Citado na página 105.

Apêndice A – Tabela dos artigos utilizados na revisão sistemática

Tabela 7 – Estudos selecionados

Referência	Área	Finalidade	Métricas	Base de dados
(TELLEZ *et al.*, 2019)	Computação	Reconhecimento facial	TG	Privado
(CHEN; CHEN, 2012)	Computação	Reconhecimento facial	EER	Público
(SHEN *et al.*, 2017)	Computação	Reconhecimento facial	A	Privado e Público
(ZHANG *et al.*, 2020)	Computação	Reconhecimento facial	Des.	Privado
(LI; MORARIU; DAVIS, 2015)	Computação	Reconhecimento facial	EER	Público
(CRUZ; SUCAR; MORALES, 2008)	Computação	Reconhecimento facial	P; R	Público
(BERTóK; FAZEKAS, 2016)	Computação	Reconhecimento facial	TG	Público
(WIJAYA; SAVVIDES; KUMAR, 2004)	Computação	Reconhecimento facial	FAR	Público
(AN *et al.*, 2009)	Computação	Reconhecimento facial	P; R	Público
(RAO *et al.*, 2015)	Computação	Reconhecimento facial	Speed	Público
(REN; JIANG; YUAN, 2013)	Computação	Reconhecimento facial	NF; TG	Público
(CHOI; TOH; BYUN, 2011)	Computação	Reconhecimento facial	A; TG	Público
(SMITH-CREASEY; ALBALOOSHI; RAJARAJAN, 2018)	Computação	Reconhecimento facial	EER	Público
(CHENG *et al.*, 2012)	Computação	Experiência de usuário	UO; NF	Privado
(CHIA *et al.*, 2015)	Computação	Experiência de usuário	UO; Des.	Privado
(PHAM; WANG, 2019)	Computação	Experiência de usuário	UO; A	Privado

Referência	Área	Finalidade	Métricas	Base de dados
(HELLMUND *et al.*, 2018)	Computação	Experiência de usuário	BU; TG	Privado
(POWELL *et al.*, 2014)	Computação	Experiência de usuário	IAS	Público
(LUNERTI *et al.*, 2017)	Computação	Experiência de usuário	UO	Privado
(SHI *et al.*, 2019)	Computação	Experiência de usuário	Freq.	Privado
(JOHANSSEN; BERNIUS; BRUEGGE, 2019)	Computação	Experiência de usuário	Sens.; A	Privado
(MA *et al.*, 2012)	Computação	Experiência de usuário	Brilho	Privado
(MY; ZELL, 2013)	Computação	Experiência de usuário	Speed	Público
(SUNARYONO; SISWANTORO; ANGGORO, 2019)	Computação	Experiência de usuário	A; TT	Privado
(RAHMAN; KEHTARNAVAZ; REN, 2009)	Computação	Detecção facial	NF	Privado
(CHEN *et al.*, 2015)	Computação	Detecção facial	A	Privado
(SARKAR; PATEL; CHELLAPPA, 2016)	Computação	Detecção facial	F1	Público
(TRIANTAFYLLIDOU; TEFAS; ANASTASIOS, 2018)	Computação	Detecção facial	P; A; R; F1	Público
(ZHANG *et al.*, 2019)	Computação	Detecção facial	P; R; TG	Público
(LEE; PARK; PARK, 2006)	Computação	Detecção facial	NF	Público
(MAHBUB; SARKAR; CHELLAPPA, 2019)	Computação	Detecção facial	P; A; R; F1	Privado e Público
(WELIVITA *et al.*, 2017)	Computação	Rastreamento facial	P; R; F1; TG	Público
(TAKAHASHI; MITSUKURA, 2013)	Computação	Rastreamento facial	EA	Privado
(XU; LI; WANG, 2011)	Computação	Rastreamento facial	P; R	Privado

Referência	Área	Finalidade	Métricas	Base de dados
(SZKUDLAREK; PIETRUSZKA, 2015)	Computação	Rastreamento facial	TG	Privado
(PASCHALAKIS; BOBER, 2004)	Computação	Rastreamento facial	EER; FAR	Privado
(WILHELM; BöHME; GROSS, 2004)	Computação	Rastreamento facial	NF	Privado
(V *et al.*, 2015)	Computação	Reconhecimento de expressão	A; R; F1	Público
(SUK; PRABHAKARAN, 2014)	Computação	Reconhecimento de expressão	A	Público
(WASHINGTON *et al.*, 2017)	Computação	Reconhecimento de expressão	UO; A	Público
(LIU *et al.*, 2019)	Computação	Outros	NF; Des.	Público
(TIAN; ARBEL; CLARK, 2018)	Computação	Outros	P; A; R; F1	Público
(NI; LI, 2020)	Computação	Reconhecimento de comportamento	A; TG; Speed	Público
(WANG *et al.*, 2012)	Segurança	Reconhecimento facial	A; TG	Privado
(WANG *et al.*, 2017b)	Segurança	Reconhecimento facial	A	Público
(MUSLIM; ISLAM, 2017)	Segurança	Reconhecimento facial	TG	Público
(NAMBIAPPAN; DATTA, 2018)	Segurança	Reconhecimento facial	A; Speed	Privado
(TAO; VELDHUIS, 2010)	Segurança	Reconhecimento facial	NF; EER	Público
(SCHNEIDER *et al.*, 2006)	Segurança	Reconhecimento facial	NF	Privado
(HU *et al.*, 2012)	Segurança	Reconhecimento facial	TRF	Privado
(FATHY; PATEL; CHELLAPPA, 2015)	Segurança	Reconhecimento facial	TRF	Privado
(SMITH; ALBALOOSHI; RAJARAJAN, 2018)	Segurança	Reconhecimento facial	ROC; EER	Público

Referência	Área	Finalidade	Métricas	Base de dados
(RIESCH *et al.*, 2015)	Segurança	Reconhecimento facial	A	Privado, Público
(PENG; SHEN, 2013)	Segurança	Reconhecimento facial	EER	Privado
(WANG *et al.*, 2017a)	Segurança	Reconhecimento facial	P	Público
(HULSEBOSCH; EBBEN, 2008)	Segurança	Reconhecimento facial	EER	Privado
(BARRA *et al.*, 2013)	Segurança	Reconhecimento facial	UO	Privado
(ZHANG *et al.*, 2018)	Segurança	Reconhecimento facial	A; FFR; FAR	Público
(MARIA *et al.*, 2014)	Segurança	Reconhecimento facial	UO; EER	Privado
(SAMANGOUEI; PATEL; CHELLAPPA, 2017)	Segurança	Reconhecimento facial	ROC; EER	Público
(WU *et al.*, 2019)	Segurança	Reconhecimento facial	A	Público
(KODIROV *et al.*, 2013)	Segurança	Rastreamento facial	DC	Privado
(ELREFAEI *et al.*, 2017)	Segurança	Rastreamento facial	TG	Privado
(MUKHERJEE *et al.*, 2019)	Segurança	Outros	ST; EC	Público
(VANITHA; VAIDEHI; VASUHI, 2018)	Segurança	Reconhecimento de comportamento	P; A; R; F1	Público
(JIN *et al.*, 2015)	Saúde	Reconhecimento facial	TS	Privado
(TAPU; MOCANU; ZAHARIA, 2018)	Saúde	Reconhecimento facial	Des.	Privado
(WONG *et al.*, 2010)	Saúde	Reconhecimento facial	EER	Público
(CHAUDHRY; CHANDRA, 2017)	Saúde	Reconhecimento facial	NF; A	Público
(GHAZAL *et al.*, 2018)	Saúde	Reconhecimento de comportamento	A	Público

Referência	Área	Finalidade	Métricas	Base de dados
(XIONG *et al.*, 2019)	Saúde	Reconhecimento de comportamento	A; FAR; TG	Privado
(ABULKHAIR *et al.*, 2015)	Saúde	Reconhecimento de comportamento	A; L	Privado
(JEON *et al.*, 2016)	Saúde	Reconhecimento de expressão	A	Público
(KALANTARIAN *et al.*, 2019)	Saúde	Reconhecimento de expressão	A; ROC	Privado
(FOGGIA *et al.*, 2019)	Saúde	Outros	A	Público
(HUANG; YANG; CHENG, 2016)	Saúde	Outros	ST; A	Público
(MITTAL; AGARWAL; NIGAM, 2018)	Educação	Reconhecimento facial	P; A	Público

Abreviações: A - Acurácia; Des. - Desempenho; DC - Distância ao Centro; EC - Energia consumida; EA - Erro absoluto; EER - *Equal Error Rate*; F1 - *F1-Score*; FAR - *False Acceptance Rate*; FFR - *False Frequency Rate*; FRI - Taxa de Reconhecimento Facial; Freq. - Frequência; IAS - Indíce de Ataques com Sucesso; L - Latência; NF - Número de faces; P - Precisão; R - Revocação; ROC - Curva ROC; TG - Tempo Gasto; TS - Taxa de sucesso; Sens. - Sensitividade; TT - Tempo de Treinamento; US - Uso de bateria

Apêndice B – Questionário de pesquisa.

Experimento de Usabilidade

Este experimento faz parte do trabalho de Mestrado intitulado: "Detecção e processamento de imagens faciais em dispositivos móveis para auxílio ao diagnóstico de Transtorno de Espectro Autista" do mestrando Henrique Silva Bortoletti no Programa de Pós-graduação em Sistemas de Informação (PPgSI) da Escola de Artes, Ciências e Humanidades da Universidade de São Paulo sob a orientação da professora Fátima L. S. Nunes e coorientação da professora Ariane Machado Lima .

O objetivo deste experimento é validar a facilidade de uso do aplicativo de captura de imagens faciais. O sistema visa a auxiliar no diagnóstico de crianças com Transtorno de Espectro Autista.

Os riscos dessa pesquisa estão relacionados ao desconforto e/ou cansaço em responder às questões: talvez o Senhor(a) se sinta cansado em responder as perguntas do formulário. Caso isso ocorra, o(a) Senhor(a) poderá interromper o preenchimento do formulário imediatamente e avisar aos responsáveis pelo experimento. Vale ressaltar que o(a) Senhor(a) não está sendo avaliado, o(a) Senhor(a) está avaliando sua experiência de uso com o aplicativo. É esperado que o experimento seja realizado até o final, mas em qualquer momento o(a) Senhor(a) poderá parar a execução se desejar.

Para participar do experimento o(a) Senhor(a) precisará:
- ter um smartphone com o sistema Android, versão mínima 9.0 (Pie);
- possuir acesso a internet;
- o dispositivo móvel possuir pelo menos 2GB de RAM;

Informações sobre a pesquisa ou dúvidas sobre a utilização do sistema poderão ser solicitadas, a qualquer momento, ao pesquisador responsável (Henrique Silva Bortoletti) pelo e-mail: henrique.bortoletti@usp.br

Este experimento consiste em realizar 8 tarefas com o tempo total de aproximadamente 30 minutos.

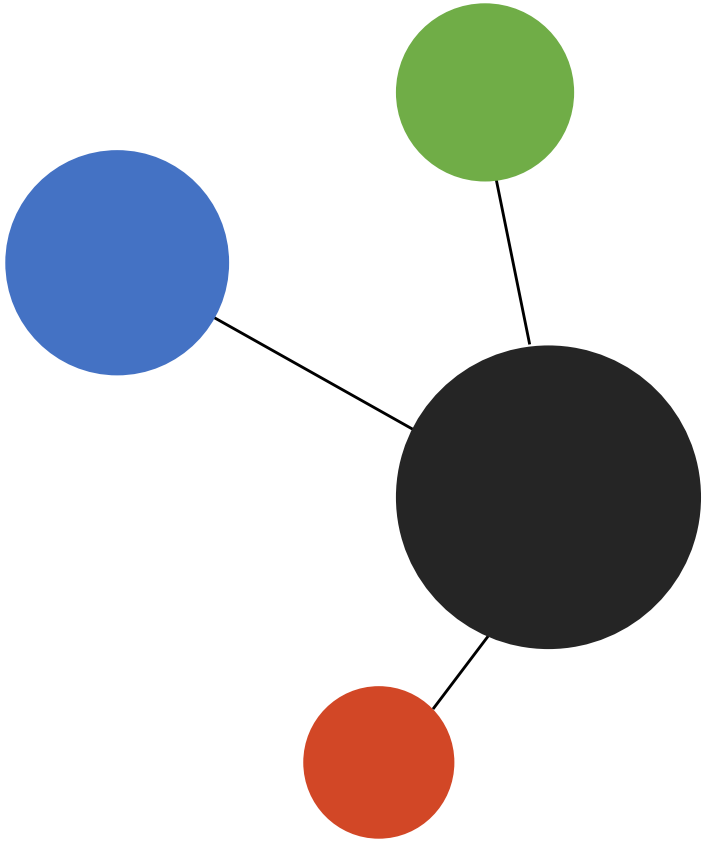
* Required

1. Email *

2. Qual o modelo do smartphone e versão do Android? *

Apêndice C – Manual da aplicação.

Child Facial Detection



Manual de instruções sobre como utilizar o aplicativo de detecção facial para crianças.

Requisitos

Para utilizar o aplicativo será necessário possuir um dispositivo móvel com as seguintes configurações mínimas:

1. Sistema operacional Android versão 9
2. Memória RAM de 2GB+
3. Acesso à internet
4. 60MB de espaço livre

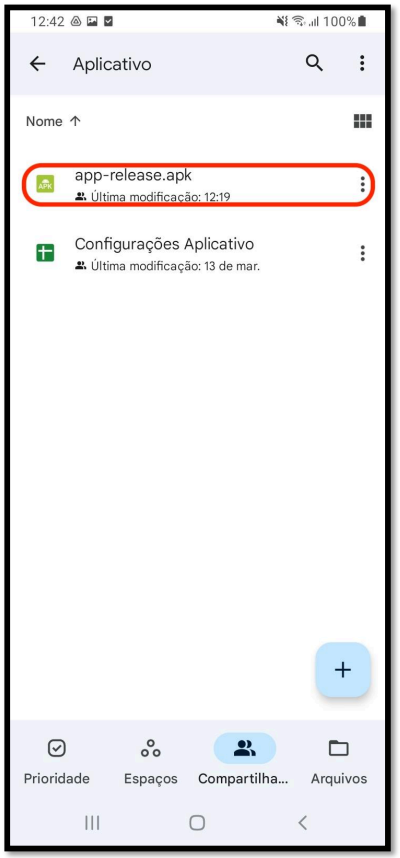
Como acessar o aplicativo?

Parte 1

1. Para realizar o download do aplicativo acesse o link:

2. https://drive.google.com/drive/u/1/folders/1iQqo_dTV3eB0kC9airwOcDP-dV0G7Wdm

3. Após acessar o link selecione o ícone com nome **app-release.apk**, conforme a imagem ao lado:

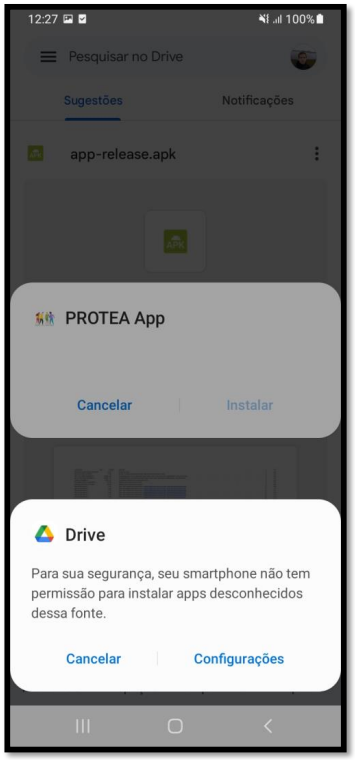


Como acessar o aplicativo?

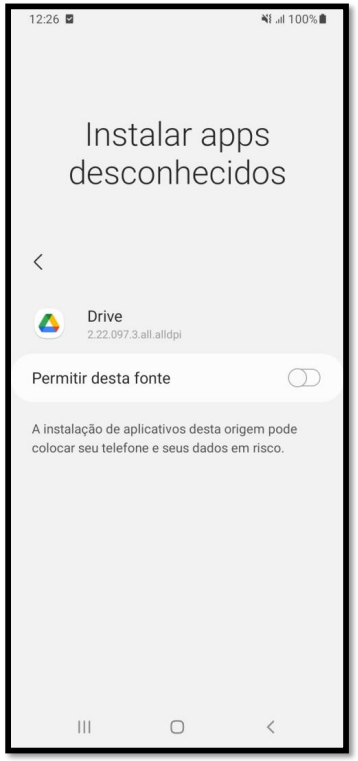
Parte 2

1. Após clicar no ícone mostrado anteriormente, autorize a instalação de aplicativos desconhecidos conforme as imagens ao lado.

2. Esta etapa deverá ocorrer automaticamente após clicar no link.



Etapa 1



Etapa 2

Como acessar o aplicativo?

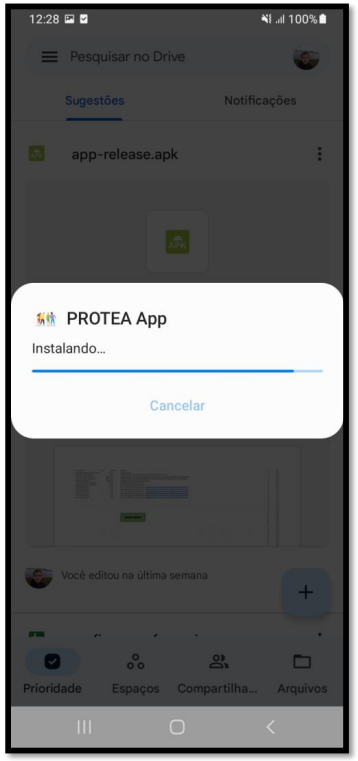
Parte 3

1. Após permitir a instalação de aplicativos desconhecidos, o aplicativo será instalado automaticamente.

2. Caso o download não inicie automaticamente, então retorne para o link inicial e clique novamente no ícone conforme descrito na parte 1 desde manual.

3. Após clicar no ícone aceite instalar o aplicativo.

Etapa 1



Etapa 2

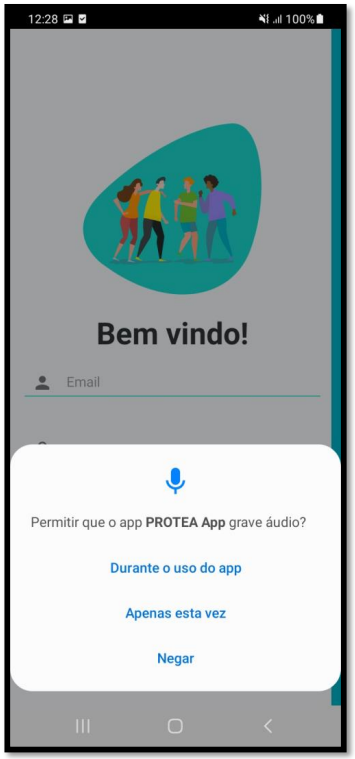
Como acessar o aplicativo?

Parte 4

1. Após o aplicativo ser instalado procure em seu dispositivo móvel pelo aplicativo **PROTEA App.**

2. Clique no link do aplicativo e aceite as permissões necessárias para o correto funcionamento.

3. Para se cadastrar clique no ícone **+**



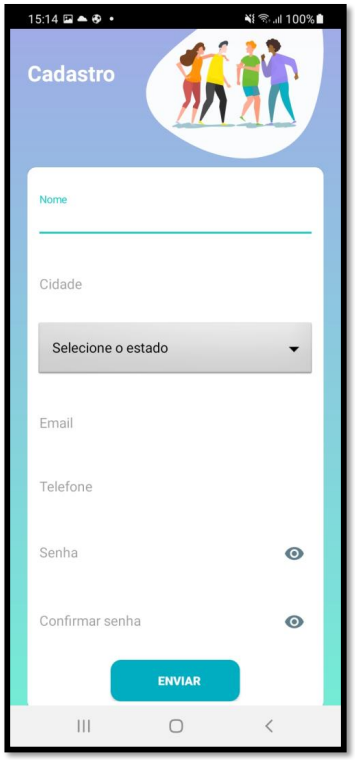
Etapa 1

Etapa 2

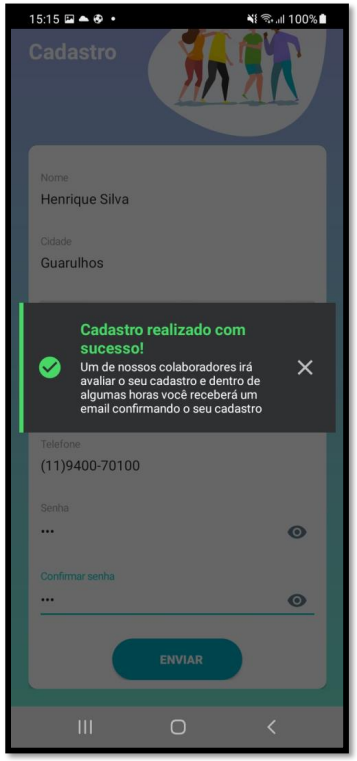
Como usar o aplicativo?

Parte 5

1. Após clicar no ícone, você será levado para a tela de cadastro.
2. Preencha os seus dados para poder realizar capturas pelo aplicativo. (Etapa 1)
3. Após ter preenchido os dados clique no botão próximo localizado na parte inferior.
4. Uma mensagem de sucesso irá aparecer indicando que o cadastro foi realizado com sucesso. (Etapa 2)



Etapa 1



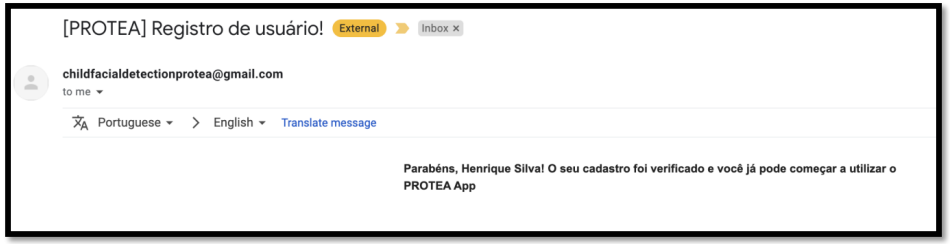
Etapa 2

Como usar o aplicativo?

Parte 6

1. O login somente poderá ser realizado após um de nossos pesquisadores aprovar o seu cadastro.

2. Quando um de nossos pesquisadores aprovar o seu cadastro, então uma mensagem será enviada para o seu email



Email de confirmação

Tela de login

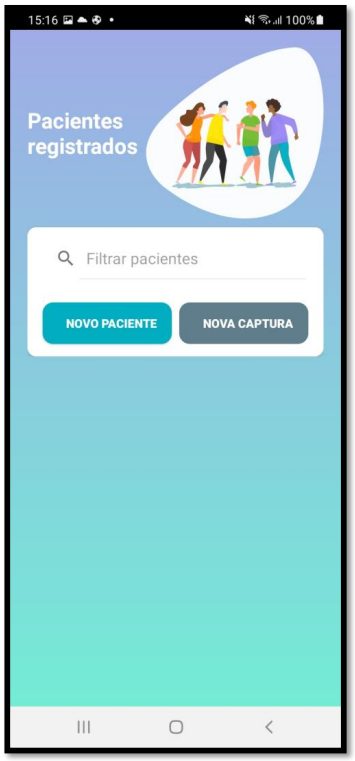
Como usar o aplicativo?

Parte 7

1. Após realizar o login uma tela com os pacientes registrados por você irá aparecer (etapa 1).

2. Clique em **NOVO PACIENTE** para cadastrar um novo paciente.

3. Após clicar em **NOVO PACIENTE** insira os dados do seu paciente (criança) e clique em **REGISTRAR** (etapa 2)



Etapa 1

Etapa 2

Como usar o aplicativo?

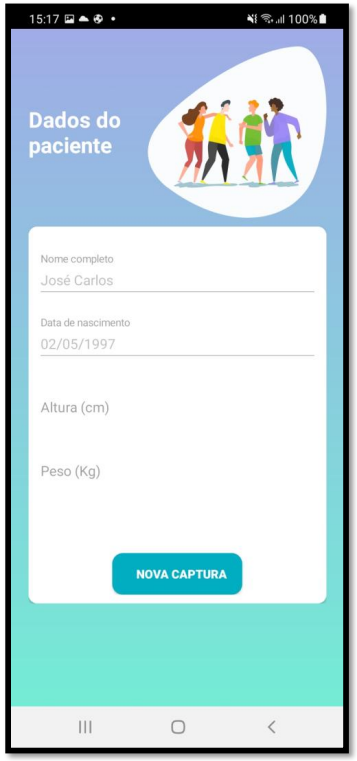
Parte 8

1. Após cadastrar o paciente este irá aparecer cadastrado na tela seguinte.

2. Selecione o paciente e clique em **NOVA CAPTURA** (etapa 1)

3. Após clicar em **NOVO CAPTURA** insira os dados do seu paciente (criança) e clique em **NOVO CAPTURA** (etapa 2)

Etapa 1



Etapa 2

Como usar o aplicativo?

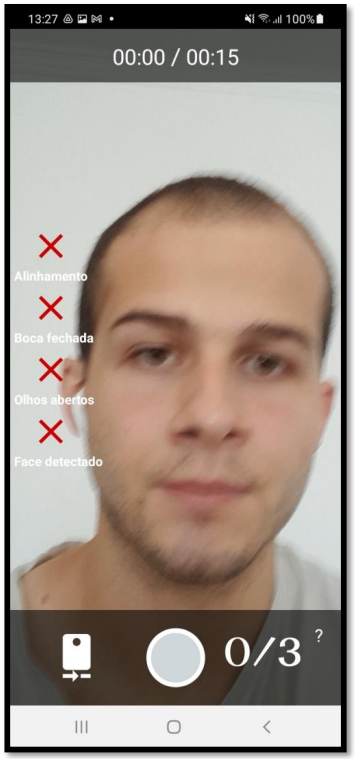
Parte 9

1. Para realizar a captura clique no círculo ao centro da parte inferior da tela.

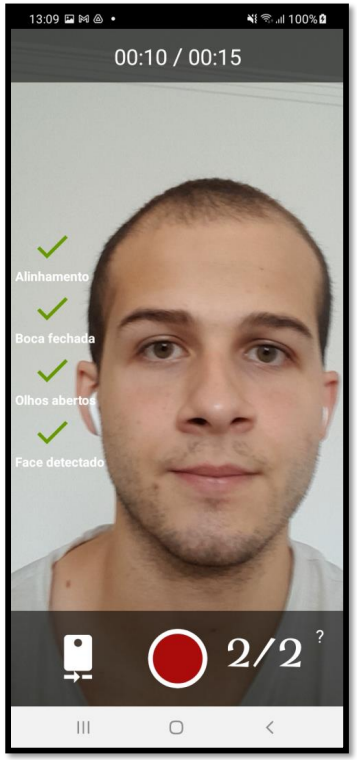
2. Para alterar a câmera para a captura clique no ícone ao lado esquerdo do círculo na parte inferior da tela.

3. Os números ao lado direito do círculo de captura indicam quantas imagens ideais foram capturadas/ quantas imagens são necessárias para finalizar a captura.

4. Os ícones no canto esquerdo da tela indicam o status da captura, caso todos fiquem verde, então uma imagem ideal foi encontrada.



Tela de captura inicial



Captura em andamento

Como usar o aplicativo?

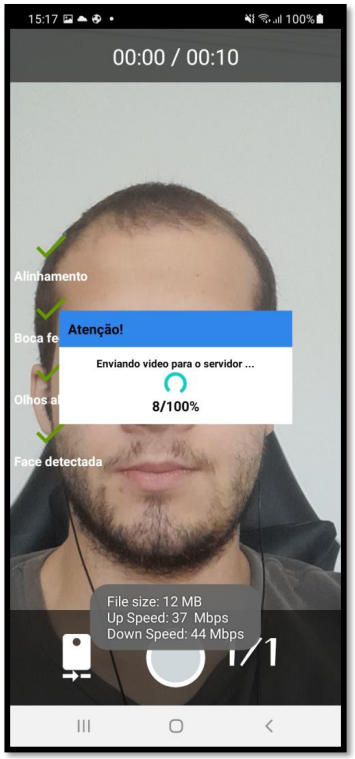
Parte 10

1. Após as imagens ideais do paciente forem encontradas, então estas serão enviadas para a análise.

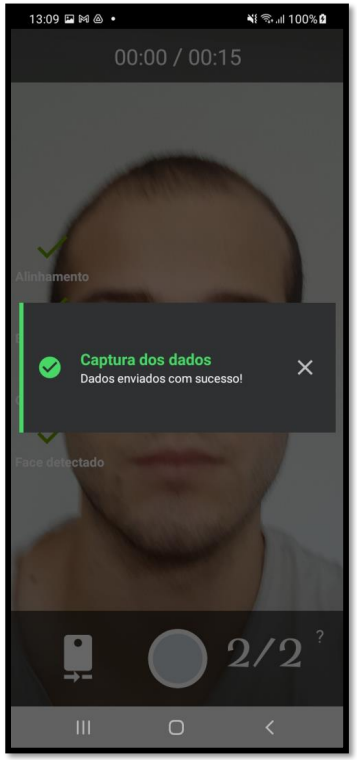
2. Durante o envio dos dados é extremamente necessário que o aplicativo não seja fechado.

3. A média de tempo de envio dos dados dura entre 10 segundos e 3 minutos.

4. Após os dados serem enviados, uma tela de sucesso será exibida.



Dados sendo enviados



Captura realizada com sucesso

Como usar o aplicativo?

Parte 11

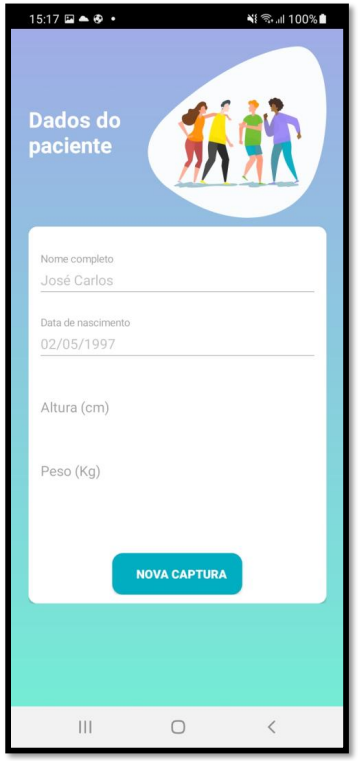
1. Caso você queira realizar uma nova captura para um paciente já cadastrado, então clique no nome do paciente. (Etapa 1)

2. Após clicar no nome do paciente, então clique no botão **NOVA CAPTURA.**

3. Para realizar a nova captura é necessário inserir apenas a altura e peso atualizados. (Etapa 2)

4. Após inserir os dados, clique em **CAPTURAR VIDEO** e repita os passos descritos na parte 8 e 9.

Etapa 1



Etapa 2

Como usar o aplicativo?

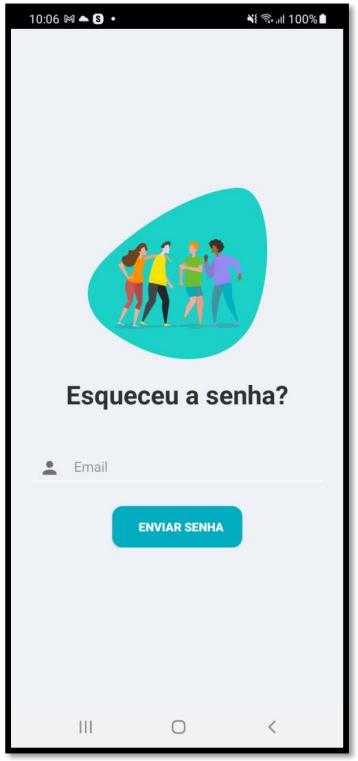
Parte 12

1. Para realizar a troca de senha clique no elemento "Esqueceu a senha?"

2. Na etapa 2 adicione a sua senha e clique em enviar senha, dessa forma você receberá um email com uma nova senha conforme a figura abaixo:

Email com nova senha

Etapa 1



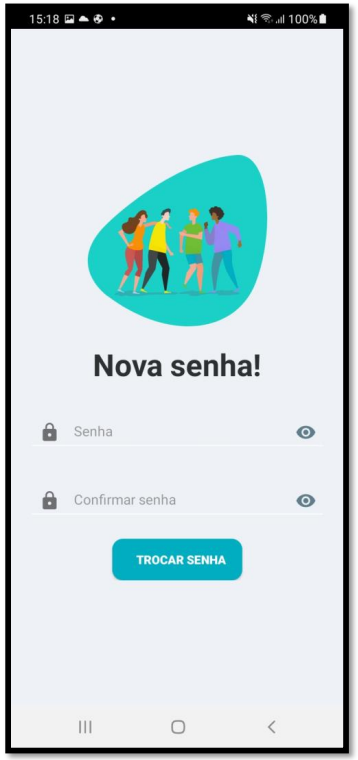
Etapa 2

Como usar o aplicativo?

Parte 13

1. Após receber sua nova senha, realize o login com a nova senha.
2. Ao realizar o login cadastre uma nova senha.

Etapa 1



Etapa 2

Como acessar os dados capturados?

Parte 14

1. Para acessar os dados capturados acesse o link abaixo:

➢

Caso você queira realizar o download do vídeo ou imagem de um paciente, então apenas clique em cima do vídeo ou imagem.

Caso você queira realizar o download de todos os vídeos ou imagens, então basta clicar no ícone ao lado da palavra Vídeo ou Imagem no cabeçalho.

Capturas realizadas pelo aplicativo

email | Pesquisar | paciente | Pesquisar | dd/mm/yyyy | dd/mm/yyyy | Pesquisar data da coleta

Email	Paciente	Data da coleta	Data de nascimento	Peso	Altura	Video	Imagem
henriquebortoletti@gmail.com	testes a	2022-03-25 09:17:59.343	25/03/2022	64.0 cm	64.0 Kg	Baixar	Baixar
henriquebortoletti@gmail.com	testes a	2022-03-25 09:18:00.098	25/03/2022	64.0 cm	64.0 Kg	Baixar	Baixar
henriquebortoletti@gmail.com	testes a	2022-03-25 09:18:00.986	25/03/2022	64.0 cm	64.0 Kg	Baixar	Baixar

Dados para consulta

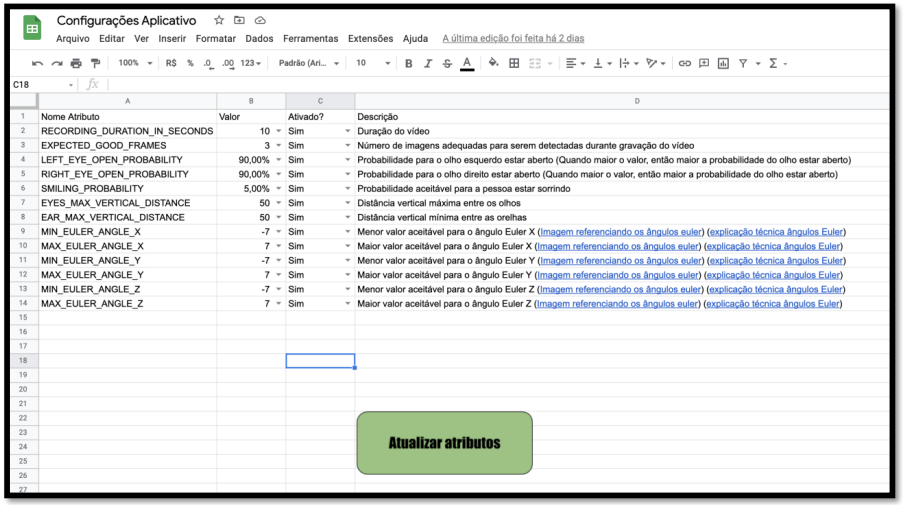
Como configurar?

Parte 15

1. Para alterar os parâmetros de captura acesse o link abaixo:

➢

Ao lado de cada parâmetro é possível ver uma descrição do que significa.



Configurações Aplicativo

Arquivo Editar Ver Inserir Formatar Dados Ferramentas Extensões Ajuda A última edição foi feita há 2 dias

Nome Atributo	Valor	Ativado?	Descrição
RECORDING_DURATION_IN_SECONDS	10	Sim	Duração do vídeo
EXPECTED_GOOD_FRAMES	3	Sim	Número de imagens adequadas para serem detectadas durante gravação do vídeo
LEFT_EYE_OPEN_PROBABILITY	90,00%	Sim	Probabilidade para o olho esquerdo estar aberto (Quando maior o valor, então maior a probabilidade do olho estar aberto)
RIGHT_EYE_OPEN_PROBABILITY	90,00%	Sim	Probabilidade para o olho direito estar aberto (Quando maior o valor, então maior a probabilidade do olho estar aberto)
SMILING_PROBABILITY	5,00%	Sim	Probabilidade aceitável para a pessoa estar sorrindo
EYES_MAX_VERTICAL_DISTANCE	50	Sim	Distância vertical máxima entre os olhos
EAR_MAX_VERTICAL_DISTANCE	50	Sim	Distância vertical mínima entre as orelhas
MIN_EULER_ANGLE_X	-7	Sim	Menor valor aceitável para o ângulo Euler X (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
MAX_EULER_ANGLE_X	7	Sim	Maior valor aceitável para o ângulo Euler X (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
MIN_EULER_ANGLE_Y	-7	Sim	Menor valor aceitável para o ângulo Euler Y (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
MAX_EULER_ANGLE_Y	7	Sim	Maior valor aceitável para o ângulo Euler Y (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
MIN_EULER_ANGLE_Z	-7	Sim	Menor valor aceitável para o ângulo Euler Z (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)
MAX_EULER_ANGLE_Z	7	Sim	Maior valor aceitável para o ângulo Euler Z (Imagem referenciando os ângulos euler) (explicação técnica ângulos Euler)

Atualizar atributos

Configurações para captura