

UNIVERSIDADE FEDERAL DE SANTA CATARINA  
CAMPUS ARARANGUÁ  
CENTRO DE CIÊNCIAS, TECNOLOGIAS E SAÚDE (CTS)  
COORDENADORIA ESPECIAL INTERDISCIPLINAR EM TECNOLOGIAS  
DA INFORMAÇÃO E COMUNICAÇÃO (CIT)  
TECNOLOGIAS DA INFORMAÇÃO E COMUNICAÇÃO

Ricardo Pires Maciel

**Desenvolvimento de uma Abordagem Computacional para Detecção Facial e  
Rastreamento do Sistema Visual da População Infantil**

Araranguá

2022

Ricardo Pires Maciel

**Desenvolvimento de uma Abordagem Computacional para Detecção Facial e  
Rastreamento do Sistema Visual da População Infantil**

Trabalho Conclusão do Curso de Graduação em Tecnologias da Informação e Comunicação do Centro de Ciências, Tecnologias e Saúde da Universidade Federal de Santa Catarina como requisito para a obtenção do título de Bacharel em Tecnologias da Informação e Comunicação.

Orientador: Prof. Cristian Cechinel Dr.

Coorientador: Prof. Antonio Carlos Sobieranski, Dr.

Araranguá

2022

Ficha de identificação da obra elaborada pelo autor,  
através do Programa de Geração Automática da Biblioteca Universitária da UFSC.

Maciel, Ricardo Pires

Desenvolvimento de uma Abordagem Computacional para  
Detecção Facial e Rastreamento do Sistema Visual da  
População Infantil / Ricardo Pires Maciel ; orientador,  
Cristian Cechinel, coorientador, Antônio Carlos  
Sobieranski, 2022.

47 p.

Trabalho de Conclusão de Curso (graduação) -  
Universidade Federal de Santa Catarina, Campus Araranguá,  
Graduação em Tecnologias da Informação e Comunicação,  
Araranguá, 2022.

Inclui referências.

1. Tecnologias da Informação e Comunicação. 2. Visão  
computacional. 3. Análise de imagem. 4. Tratamento  
precoce. 5. Distúrbio infantil. I. Cechinel, Cristian. II.  
Sobieranski, Antônio Carlos. III. Universidade Federal de  
Santa Catarina. Graduação em Tecnologias da Informação e  
Comunicação. IV. Título.

Ricardo Pires Maciel

**Desenvolvimento de uma abordagem computacional para detecção facial e rastreamento  
do sistema visual da população infantil**

Este Trabalho Conclusão de Curso foi julgado adequado para obtenção do Título de “Bacharelado em Tecnologias da Informação e Comunicação” e aprovado em sua forma final pelo Curso de Graduação em Tecnologias da Informação e da Comunicação do Centro de Ciências, Tecnologias e Saúde da Universidade Federal de Santa Catarina – Campus Araranguá

Araranguá, 25 de março de 2022.

---

Prof. Vilson Gruber, Dr.  
Coordenador do Curso

**Banca Examinadora:**

---

Prof. Cristian Cechinel, Dr.  
Orientador  
Instituição UFSC

---

Prof.(a) Eliane Pozzebon, Dr.a  
Avaliador(a)  
Instituição UFSC

---

Prof.(a) Roderval Marcelino, Dr.  
Avaliador(a)  
Instituição UFSC

Este trabalho é dedicado aos meus amigos e aos meus queridos pais.

## **AGRADECIMENTOS**

Agradeço primeiramente aos meus pais, Joel Soares Maciel e Rosane Pereira Pires Maciel, por sempre terem me incentivado muito em busca de estudos e novos conhecimentos.

À minha irmã Juliana Pires Maciel, pelo apoio na graduação. À minha namorada Bruna Samantha Marchi por sempre estar me apoiando e me incentivando a evoluir.

Ao Cristian Cechinel, por me orientar e me auxiliar na realização deste trabalho. Ao Antonio Carlos Sobieranski por todo o conhecimento e experiências compartilhadas comigo nos projetos de pesquisa.

Por fim, mas não menos importante, a todos meus amigos e familiares que me apoiaram durante a graduação, vocês me auxiliaram muito durante este caminho.

“Não existe um caminho para a felicidade. A felicidade é o caminho.” (Mahatma Gandhi)

## RESUMO

**Introdução:** A visão é o sentido com maior informação sensorial do ser humano, com isso deficiências visuais podem prejudicar o desenvolvimento na infância. Atualmente, os aparelhos para análise da visão possuem altos valores e uma das possibilidades são as tecnologias da informação e comunicação, amplamente utilizadas no avanço da saúde.

**Objetivos:** Este trabalho tem como objetivo desenvolver um software que auxilie na detecção facial e rastreamento visual da população infantil entre 0 e 24 meses, como ferramenta preliminar para um processo computacional de tomada de decisão acerca das atipicidades da visão infantil.

**Metodologia:** Para o desenvolvimento do software foram realizadas três fases: definição do padrão de captura, obtenção das imagens para formação de um banco de imagens da população infantil e o desenvolvimento de um algoritmo de detecção facial e do sistema visual em recém-nascidos. **Resultados preliminares:** Participaram do estudo 40 crianças. O estudo apresentou como resultado o funcionamento do software de acordo com a proposta inicial e fases desenvolvidas para conseguir alcançar a função final de rastreamento das posições percorridas pela íris. **Conclusão e discussões:** Com um algoritmo de detecção visual desenvolvido, é possível que se faça o rastreamento de objetos do sistema visual. Com isso, espera-se facilitar e difundir exames de verificação de distúrbios visuais na infância a ser realizado a posteriori por outros processos computacionais, assim facilitando o tratamento precoce.

**Palavras-chave:** Visão computacional. Análise de imagem. Tratamento precoce. Distúrbio infantil.

## **ABSTRACT**

**Introduction:** Vision is the sense with the greatest sensory information of the human being, so visual impairments can impair childhood development. Currently, vision analysis devices have high values and one of the possibilities is information and communication technologies, widely used in the advancement of health. **Objectives:** This work aims develop software that assists in facial detection and visual tracking of the child population, as a preliminary tool for a computational process to determine children's atypicalities. **Methodology:** For the development of the software, three phases were carried out: definition of the capture pattern, obtaining the images to form an image bank of the infant population and development of an algorithm for facial detection and the visual system in newborns. **Preliminary results:** 40 children participated in the study. The study presented as a result the operation of the software according to the initial proposal and phases developed to achieve the final function of tracking the positions traveled by the iris. **Conclusion and discussions:** With a developed visual detection algorithm, it is possible to track objects in the visual system. With this, it is expected to facilitate and disseminate examinations to check for visual disturbances in childhood, thus facilitating early treatment.

**Keywords:** Computer Vision. Image Analysis. Early Treatment. Child disorder.

## LISTA DE FIGURAS

Figura 1. Imagem das estruturas do olho (RAMOS, 2006).....	17
Figura 2. Representação da segmentação <i>Watershed</i> (IMAGE SEGMENTATION WITH WATERSHED ALGORITHM, 2022).....	20
Figura 3. Representação do agrupamento das ilhas (PRIESE & REHRMANN, 1993/2003).....	21
Figura 4. Representação da segmentação via Jseg (SHAN, 2009).....	22
Figura 5. Representação do operador de Canny (NAIDU <i>et al</i> , 2012).....	24
Figura 6. Exemplo de funcionamento do max pooling com filtro de 2x2 em uma imagem 4x4.....	27
Figura 7. Etapas do desenvolvimento do sistema de detecção.....	32
Figura 8. Imagem do código que desenha os valores obtidos pela aplicação.....	34
Figura 9. Imagem da estrutura ocular da criança A sem a aplicação do software.....	35
Figura 10. Imagem da resposta visual da criança A com a aplicação do software.....	36
Figura 11. Imagem da estrutura ocular da criança B sem a aplicação do software.....	36
Figura 12. Imagem da resposta visual da criança B com a aplicação do software.....	37
Figura 13. Imagem da estrutura ocular da criança C sem a aplicação do software.....	37
Figura 14. Imagem da resposta visual da criança C com a aplicação do software.....	38

## LISTA DE ABREVIATURAS E SIGLAS

API - *Application Programming Interface*

CID-10 - Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - 10ª Revisão

ICO - *International Council of Ophthalmology*

OMS - Organização Mundial da Saúde

OpenCV - *Open Source Computer Vision Library*

RN - Recém-nascido

TIC - Tecnologias da Informação e Comunicação

UTI - Unidade de terapia intensiva

UTIN - Unidade de terapia intensiva neonatal

## SUMÁRIO

<b>1</b>	<b>INTRODUÇÃO.....</b>	<b>13</b>
1.1	OBJETIVOS.....	14
<b>1.1.1</b>	<b>Objetivo Geral.....</b>	<b>14</b>
<b>1.1.2</b>	<b>Objetivos Específicos.....</b>	<b>14</b>
1.2	JUSTIFICATIVA.....	15
<b>2</b>	<b>FUNDAMENTAÇÃO TEÓRICA.....</b>	<b>15</b>
2.1	FISIOLOGIA DA VISÃO HUMANA.....	16
2.2	ATIPICIDADE VISUAL NA POPULAÇÃO INFANTIL.....	17
2.3	PROCESSAMENTO DIGITAL DE IMAGEM E VISÃO COMPUTACIONAL	18
<b>2.3.1</b>	<b>Imagem Digital.....</b>	<b>18</b>
<b>2.3.2</b>	<b>Métodos Convencionais de Processamento.....</b>	<b>19</b>
2.3.2.1	<i>Segmentação.....</i>	<i>19</i>
2.3.2.2	<i>Limiarização.....</i>	<i>22</i>
2.3.2.3	<i>Detecção de bordas.....</i>	<i>23</i>
<b>2.3.3</b>	<b>Métodos Avançados.....</b>	<b>24</b>
2.3.3.1	<i>Aprendizado de Máquina.....</i>	<i>24</i>
2.3.3.2	<i>Redes Neurais.....</i>	<i>25</i>
2.3.3.3	<i>Redes Neurais Convolucionais.....</i>	<i>26</i>
2.4	RECURSOS TECNOLÓGICOS.....	27
<b>2.4.1</b>	<b>Python.....</b>	<b>27</b>
<b>2.4.2</b>	<b>OpenCV.....</b>	<b>28</b>
<b>2.4.3</b>	<b>Face Mesh.....</b>	<b>28</b>
<b>3.</b>	<b>METODOLOGIA.....</b>	<b>28</b>
3.1	PROCEDIMENTOS DA PESQUISA.....	29

3.2	DEFINIÇÃO DE PADRÃO DE CAPTURA.....	29
3.3	OBTENÇÃO DAS IMAGENS.....	30
3.4	DESENVOLVIMENTO DO ALGORITMO DE DETECÇÃO.....	31
<b>4.</b>	<b>DESENVOLVIMENTO.....</b>	<b>31</b>
4.1	LEVANTAMENTO DE REQUISITOS.....	31
4.2	DESENVOLVIMENTO DO SISTEMA DE DETECÇÃO.....	32
4.3	DETALHES DO DESENVOLVIMENTO E CONFIGURAÇÃO DOS ALGORITMOS.....	34
<b>5.</b>	<b>RESULTADOS PRELIMINARES.....</b>	<b>35</b>
<b>6.</b>	<b>CONCLUSÕES E DISCUSSÕES.....</b>	<b>38</b>
	<b>REFERÊNCIAS.....</b>	<b>40</b>
	<b>ANEXO 1 – Termo de Consentimento Livre e Esclarecido.....</b>	<b>45</b>

## 1 INTRODUÇÃO

A visão é o sentido com maior informação sensorial, sendo responsável por 85% do contato do homem com o mundo. Através disso, consegue-se imaginar o quão mais complexo se torna a vida com distúrbios visuais. A situação das dificuldades geradas por distúrbios visuais se agrava ainda mais quando é presente desde a infância, pois traz dificuldades no aprendizado e conseqüentemente prejudica o desenvolvimento psicossocial (SILVA, et al, 2013).

No entanto, quando deficiências e distúrbios visuais são detectados precocemente, antes de a criança completar um ano de vida, tem-se uma melhora considerável no retorno dado pelos estímulos quando comparado a um tratamento feito com idade mais avançada (GAGLIARDO, NOBRE, 2001).

Uma das possibilidades de auxílio são os sistemas de informação, pois tiveram papel importante na evolução da saúde nas últimas décadas. No início auxiliaram na digitalização das instituições e no processamento de informações e nos dias atuais se aproximam para a utilização em tratamentos e diagnósticos. De acordo com a Organização Mundial da Saúde, as Tecnologias da Informação e Comunicação (TIC) aplicadas à saúde são reconhecidas como uma das áreas que crescem mais rapidamente (HAUX, 2018).

Atualmente, os valores de aparelhos de análise da visão são muito expansivos, dificultando a disseminação de testes precoces de visão. Todavia com os avanços de algoritmos de processamento de imagem, abriu-se a possibilidade do desenvolvimento do software de detecção visual, que com auxílio da visão computacional, faz a identificação facial e visual, a fim de rastrear a movimentação da íris. Possibilitando assim a identificação da possibilidade de existência de atipicidades visuais ou tendência de desenvolvimento de atipicidade na criança. Com a criança recebendo a detecção antecipada o tratamento precoce se torna algo possível, minimizando os impactos futuros (I- CREATE, 2015).

O desenvolvimento de um software que identifique as crianças com alta precisão é possível com a definição de um padrão de captura e com a criação de um banco de imagens de qualidade, diverso e grande. Pois um dos maiores problemas das *Application Programming Interface* (API) de visão computacional é o fato de que mesmo com as redes possuindo um grande volume de dados no seu treinamento, os dados geralmente são

tendenciados e de baixa qualidade. Gerando assim, um sistema com grande volume de treinamento mas com pouca redução nos erros de detecção, sendo inviável para aplicação na saúde, uma área que não tolera erros (SILVA et al, 2020).

Com isso, torna-se viável e necessário a implementação de um software acessível que detecte a movimentação visual e a face da população infantil recém nascida, contribuindo na detecção de anomalias.

## 1.1 OBJETIVOS

Nas seções abaixo estão descritos o objetivo geral e os objetivos específicos deste TCC.

### 1.1.1 Objetivo Geral

O objetivo geral é desenvolver um software que auxilie na detecção facial e rastreamento visual da população infantil até 24 meses, como ferramenta preliminar para um processo computacional de tomada de decisão acerca das atipicidades da visão infantil.

### 1.1.2 Objetivos Específicos

- Construção de um *dataset* interno;
- Estudo e implementação de método para reconhecimento facial em população infantil;
- Estudo e implementação de método de *tracking* ocular nas seleções de regiões faciais;
- Propor fluxo computacional para identificação facial e rastreamento do olhar.
- Realizar a aquisição de imagens ou vídeos do protocolo de bateria de testes visuais na população infantil para construção de um *dataset* interno, visando testar e calibrar a metodologia computacional.

## 1.2 JUSTIFICATIVA

O desenvolvimento visual tem como importância a possibilidade de interação entre diferentes sensações e o entendimento de estímulos variados. Quando prejudicada, a visão pode acarretar em limitações na adaptação da criança ao ambiente. Entretanto, por meio da intervenção precoce a criança tem a oportunidade de desenvolver suas potencialidades e integrar todas as áreas de desenvolvimento (GAGLIARDO, NOBRE, 2001).

Com os avanços recentes na área computacional, abriram-se oportunidades para o desenvolvimento de ferramentas computacionais na área da saúde, que proporcionam mais precisão e agilidade nos diagnósticos e protocolos (TIBES, DIAS, ZEM-MASCARENHAS, 2014).

Atualmente, a tecnologia mais avançada para resolver problemas envolvendo análise de visão é a visão computacional, que é a área da tecnologia que tenta se aproximar da visão humana, com o auxílio da inteligência artificial, que possui o benefício de poder ser treinada para melhorar a detecção para o cenário necessário e inferir ações a partir de resultados visuais (GONZALEZ, 2009).

Estima-se que a visão computacional seja a tecnologia do futuro, com diversas aplicações sendo desenvolvidas como a inspeção de produtos, na área da segurança através da detecção de faces de pessoas procuradas e diversas outras aplicações. Inúmeras gigantes da tecnologia estão investindo em tecnologias de detecção como a *Amazon*, *Google* e *Microsoft*, fazendo com que diversos métodos e algoritmos estejam cada vez mais acessíveis para o desenvolvimento e pesquisa quando comparado com anos atrás (PIXFORCE, 2021).

Sabendo disto, a utilização de uma ferramenta que unifique as necessidades da área da saúde, como a prevenção e intervenção precoce, e os benefícios computacionais, como precisão e mobilidade, seria de grande valia para auxiliar no diagnóstico precoce de distúrbios visuais e com isso prevenir complicações secundárias causadas por estes distúrbios.

## **2 FUNDAMENTAÇÃO TEÓRICA**

Nesta seção serão apresentadas as teorias dos principais temas do estudo.

### **2.1 FISIOLÓGIA DA VISÃO HUMANA**

A visão acontece, primeiramente, através de uma luz proveniente de um objeto que atravessa os meios transparentes do olho e chega à retina. Quando isto acontece, a luz é transformada em impulsos elétricos que são transportados por meio dos nervos e vias ópticas até o córtex occipital. No córtex acontece a decodificação dos impulsos em uma impressão visual, esta impressão apresenta efeito tridimensional devido à junção das duas imagens que são recebidas uma de cada olho (SOUZA, 1997).

Para que a visão ocorra, é necessário o funcionamento em conjunto de estruturas, exemplificado na Figura 1. Dentre elas estão:

- Córnea: é a primeira estrutura que a luz encontra, formada por cinco camadas transparentes que possuem como função a proteção do olho.
- Íris: é a estrutura colorida do olho, localizada posteriormente à córnea. Possui músculos que contraem de acordo com a luminosidade do local, aumentando ou diminuindo a pupila.
- Pupila: é a abertura central da íris pela qual a luz atravessa para atingir o cristalino.
- Cristalino: estrutura que ajusta na retina o foco da luz que vem da pupila. Em sua função de acomodação pode aumentar ou diminuir a sua curvatura para se adaptar as focalizações de imagens.
- Retina: recebe a luz focalizada pelo cristalino. Em sua estrutura possui fotorreceptores que convertem a luz em impulso elétrico, sendo dois tipos de receptores, os cones e bastonetes.
- Nervo óptico: realiza o transporte dos impulsos elétricos para o córtex.
- Esclera: estrutura fibrosa mais externa ao olho contínua a córnea, estabelece a forma do globo ocular (RAMOS, 2006).

Entretanto, a capacidade visual não está completa desde o nascimento. O desenvolvimento da visão acontece até os 7 anos de idade, como a mielinização (a bainha de mielina é formada por uma camada lipídica, que recobre os axônios que são estruturas dos neurônios, e tem como função garantir a proteção e comunicação mais rápida entre neurônios) do nervo óptico, a maturação do globo ocular, e a associação das áreas corticais (ZIMMERMANN, 2019).

O desenvolvimento visual também depende do desenvolvimento neuromotor da criança. Com as atividades exploradas no desenvolvimento infantil forma-se o repertório

visual. Os estímulos e experiências na infância são fatores essenciais para o amadurecimento do sistema visual (ZIMMERMANN, 2019).

Durante o período de maturação, o cérebro interage com a retina para aperfeiçoar as informações visuais obtidas do ambiente. Sendo assim, a visão pode aprimorar-se ou complicar-se dependendo da qualidade da informação visual. Um exemplo de complicação da maturação do desenvolvimento visual é o estrabismo, pois o desalinhamento dos eixos visuais entrega imagens diferentes para a decodificação dos impulsos em uma imagem tridimensional o que pode ocasionar em não uso do olho preterido (SOUZA, 1997).

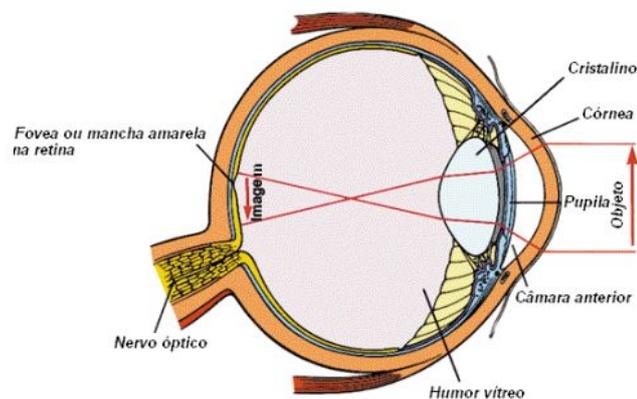


Figura 1. Imagem das estruturas do olho (RAMOS, 2006).

## 2.2 ATIPICIDADE VISUAL NA POPULAÇÃO INFANTIL

A visão é um dos sentidos mais importantes quando se trata do desenvolvimento físico e mental do indivíduo. Uma pessoa com distúrbio visual presente desde pequeno tem impacto direto no aprendizado e na vida social, pois é através da visão que se percebe diversos gestos, condutas visuais, estímulos criativos e integrações com o meio (GAGLIARDO, 2004).

O desenvolvimento visual ocorre principalmente nos primeiros 18 meses de vida, tendo os três primeiros meses de vida como principal período para detecção e intervenção precoce. Após isso a evolução visual ocorre mais lentamente, estimando-se que até os 4 anos o sistema visual já esteja completamente desenvolvido, mas podendo ser desenvolvido até os 10 anos (GAGLIARDO, 2004) (GRAZIANO e LEONE, 2005).

Quando se trata de atipicidades visuais, há diversos tipos de classificações e formas de detecção, segundo a Organização Mundial da Saúde (OMS), as deficiências visuais podem ser classificadas em categorias que abrangem desde perda visual leve até ausência total de visão . Para a categorização das anomalias, baseia-se em campos visuais, e valores de acuidade visual, usando a Classificação Estatística Internacional de Doenças e Problemas Relacionados à Saúde - 10ª Revisão (CID-10) como referência para as comparações classificativas (DIRETRIZES DE ATENÇÃO À SAÚDE OCULAR NA INFÂNCIA, 2016).

No ano de 2002, o “*International Council of Ophthalmology*” (ICO) fez a proposição de categorias de deficiência visual, baseado no critérios da CID-10, sendo as principais classificações: Cegueira, para casos onde o indivíduo possui perda total da visão e/ou necessita de auxílios especiais para repor suas habilidades visuais; Baixa visão, para indivíduos com grandes graus de perda na visão, porém que conseguem ser auxiliados por aparelhos ópticos; Incapacidade visual, para indivíduos que possuem a perda de características visuais ( perda da acuidade visual, do campo visual); Perda visual, que se refere a perda total e perda parcial da visão, com a característica da incapacidade visual (DIRETRIZES DE ATENÇÃO À SAÚDE OCULAR NA INFÂNCIA, 2016).

Sobre as anomalias visuais na infância, elas podem variar de acordo com a condição econômica e regional da criança. Dentre as principais causas de baixa visão na população infantil, pode-se citar a catarata infantil, a retinocoroidite por toxoplasmose, e o glaucoma congênito. Também se sabe que há fatores que implicam em uma grande incidência de casos de distúrbios visuais, que é em bebês prematuros, pois durante a etapa gestacional as estruturas oculares estão em pleno desenvolvimento e são muito sensíveis a malformações (DIRETRIZES DE ATENÇÃO À SAÚDE OCULAR NA INFÂNCIA, 2016).

## 2.3 PROCESSAMENTO DIGITAL DE IMAGEM E VISÃO COMPUTACIONAL

Nesta seção serão abordadas as técnicas e aplicações para processamento de imagem e visão computacional.

### 2.3.1 Imagem Digital

A imagem digital é definida como uma função bidimensional  $f(x,y)$ , onde  $x$  e  $y$  são coordenadas espaciais e o valor de  $f$  em qualquer par de coordenadas é referente a intensidade luminosa da imagem nesse ponto. Toda imagem digital é composta por um número finito de elementos de imagem, chamados de pixels, cada um com sua localização e valor de intensidade luminosa específicos. A localização dos pixels é feita através de uma matriz de duas dimensões, com  $M$  linhas e  $N$  colunas, onde  $x$  e  $y$  são as coordenadas (QUEIRÓZ, GOMES, 2001).

Para que uma imagem possa ser processada digitalmente, ela precisa ser uma imagem digital, ou seja, ou é uma imagem criada digitalmente, ou precisará ser adquirida através de um processo chamado aquisição. Para efetuar a aquisição de uma imagem analógica (algo retratado na vida real), é necessário um dispositivo óptico, como uma câmera, que faz a conversão da cena real para uma representação digital da cena (GONZALEZ, 2008).

### **2.3.2 Métodos Convencionais de Processamento**

Os métodos convencionais de processamento de imagem serão exemplificados nos tópicos: Segmentação, Limiarização e Detecção de bordas

#### *2.3.2.1 Segmentação*

O processo de segmentar consiste em subdividir a imagem em partes para a identificação entre dois ou mais objetos, que é possível graças a algoritmos e técnicas de segmentação, como a limiarização e a segmentação em bordas. Dentro da segmentação existem diversas técnicas, cada uma com seu modo de atuar específico, como por exemplo o

*watershed*, *color structure code* e *jseg*, cada técnica dessas será brevemente exemplificada a fim de contextualização (GONZALEZ, 2002).

**Watershed:** O método de segmentação *Watershed* faz o cálculo do gradiente em todos os pixels da imagem. Com os valores de gradiente, monta-se uma superfície topográfica, onde as regiões menores são compostas pelos menores gradientes, e as maiores, de maiores gradientes. O motivo da técnica ter esse nome se dá ao fato de que o crescimento das regiões com seus tamanhos é preenchido em regiões que formam uma espécie de bacia hidrográfica (PIZARRO, 2013).

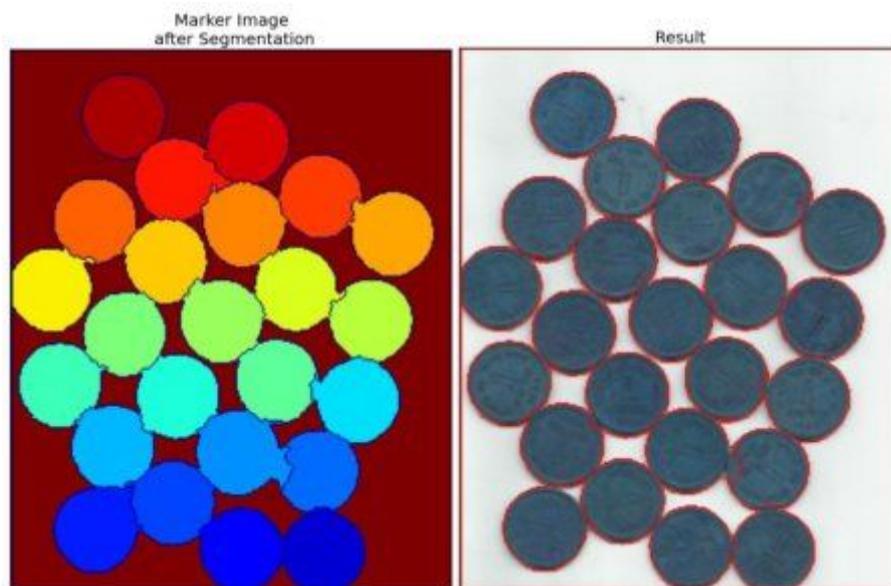


Figura 2. Representação da segmentação *Watershed* (IMAGE SEGMENTATION WITH WATERSHED ALGORITHM, 2022)

**Color Structure Code:** É um algoritmo de segmentação baseado no crescimento de regiões formadas por ilhas hexagonais, que são organizadas em níveis. Uma ilha com o nível 0 é um hexágono feito de 7 pixels, a organização dos níveis pode ser feita de tal forma que algumas ilhas sobreponham o espaço da outra. O processo de formação de ilhas é repetido até que uma única ilha consiga cobrir a imagem toda, conforme representado na Figura 3.

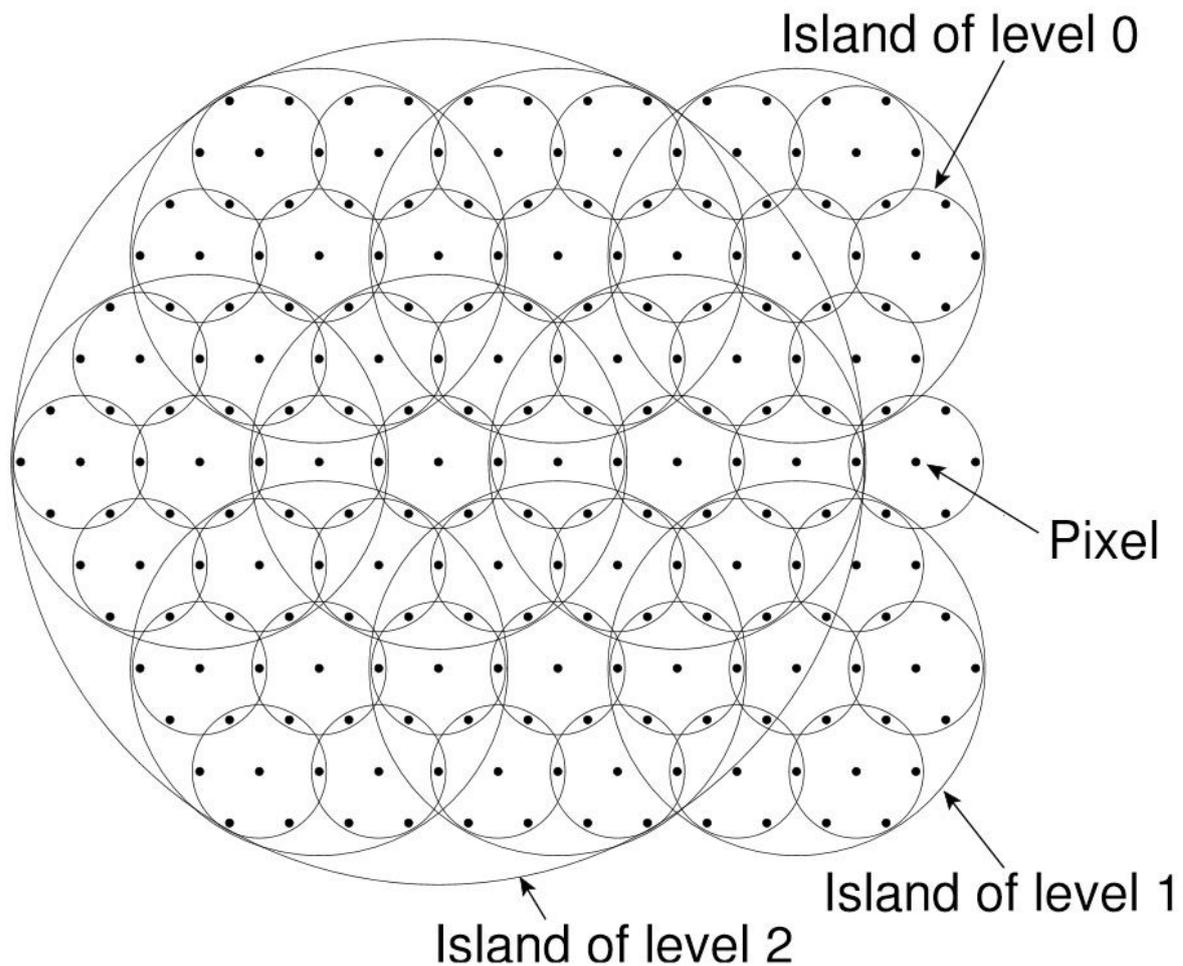


Figura 3. Representação do agrupamento das ilhas (PRIESE & REHRMANN, 1993/2003)

Para a realização do Color Structure Code, executam-se quatro etapas, a fase de pré processamento, onde são utilizadas técnicas de filtragem, com a finalidade de passar os pixels para a proximidade da vizinhança; posteriormente inicia-se a fase de inicialização, que é onde ocorre o agrupamento dos pixels em ilhas, separados por cores mútuas até certo limiar; Após isso ocorre a etapa de ligação, onde os elementos começam a se unir, formando ilhas hexagonais de pixels de cores próximas; Por fim ocorre a fase de divisão, onde é feito a segmentação pelo limiar da diferença das cores. Caso uma diferença de cor esteja acima do limiar, eles não serão unidos, apesar de estarem conectados pela cadeia de cores próximas de pixels.

**Jseg:** O foco do Jseg é trazer uma automatização para a segmentação de cores de imagens. Portanto é um método que não possui nenhuma parametrização, tendo sua criação

focada completamente em uma segmentação autônoma, que não depende de uma calibração por parte do usuário. O algoritmo atua presumindo que a imagem possui regiões com padrões de cores formado de forma uniforme. Tendo a informação das cores de uma imagem, ela é representada por faixas de cores, sendo essas cores diferenciáveis entre as regiões vizinhas, conforme representado na Figura 4. (SEGMENTAÇÃO POR QUANTIZAÇÃO CROMÁTICA BASEADA EM IMAGENS-J (J-SEG).)

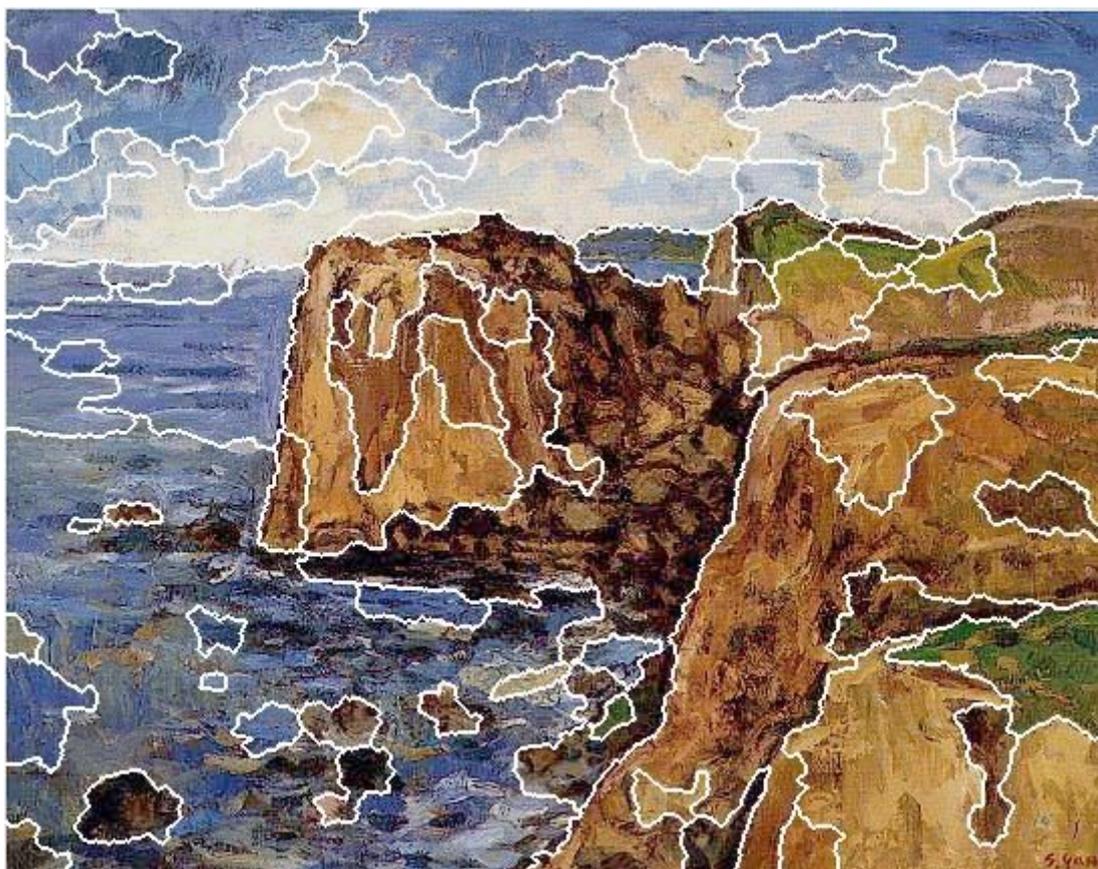


Figura 4. Representação da segmentação via Jseg (SHAN, 2009)

### 2.3.2.2 Limiarização

A limiarização ou threshold é um processo de segmentação de imagem que é baseado na diferença dos níveis de cinza, de modo a extrair os objetos de interesse pelo limiar  $T$  que separa os agrupamentos em níveis de cinza na imagem. Uma das dificuldades de segmentação pela limiarização é a de definição de qual o valor do limiar, pois em uma imagem podem haver problemas como iluminação, semelhança entre tons de cinza no objeto, ruído e contraste excessivo, dificultando o processo. Para resolver esse problema, utiliza-se métodos

adaptativos, que percorrem a imagem pixel a pixel, e separam os pixels em objeto ou fundo, o sucesso desse método depende de quão bem definidos foram os pixels de interesse segmentados (GONZALEZ, 2008).

Como exemplo de técnicas de limiarização, podemos citar três exemplos:

- Limiarização binária, onde caso a intensidade do pixel  $src(x,y)$  seja maior que o limiar, a nova intensidade do pixel é definida como  $MaxVal$ , caso o contrário, os pixels são definidos como zero.
- Limiarização para o zero, se o  $src(x,y)$  for menor que o limiar, o novo valor definido do pixel será zero
- Truncamento, onde se a intensidade máxima dos pixels passar o limiar  $src(x,y)$ , o valor será truncado no limiar (BASIC THRESHOLDING OPERATIONS).

### 2.3.2.3 Detecção de bordas

Com a limiarização, tem-se o propósito de demonstrar as principais características da imagem, porém a detecção das bordas é necessária, para identificar os grupos de pixels que formam os pontos de interesse da imagem. As bordas na imagem são os contornos dos objetos descritos na imagem, ou seja, com a detecção dos pontos de borda, encontra-se as transições de objetos distintos, assim conseguindo separar melhor os elementos para futuro tratamento (PIZARRO, 2013).

Como exemplos de algoritmos de detecção de bordas, temos Roberts, Sobel e Canny, tendo cada um suas características específicas de execução, que serão descritas individualmente (PIZARRO, 2013).

**Operador de Roberts:** É o algoritmo mais simples e antigo para detecção de bordas, e consiste em utilizar uma matriz  $2 \times 2$  para descobrir as alterações na direção de  $x$  e  $y$ . Nas regiões que os pixels estão em duas fronteiras e que o resultado dos operadores é igual a zero, sabe-se que o local é uma borda (SILVA e ALVES, 2008).

**Operador de Sobel:** Utiliza-se de duas mascaras para conseguir obter os gradientes vertical e horizontal das bordas. A fórmula para encontrar o gradiente é a mesma que a

do operador de Roberts, porém as mascarás são 3x3, fazendo que sejam menos impactadas pelo ruído da imagem do que a de Roberts (PIZARRO,2013).

**Operador de Canny:** É apontado por diversos autores como a técnica que possui resultados mais constantes na determinação de bordas. Na criação do algoritmo foi definido que um detector de bordas deveria cumprir 3 requisitos mínimos: probabilidade de erro baixíssima, devendo detectar somente bordas reais; boa localização, que consiste em a distância entre os pixels detectados e os da borda verdadeira serem muito próximos; E respostas corretas, ou seja, o detector de bordas pode encontrar apenas uma borda quando espera-se que ele detecte uma borda (SILVA e ALVES, 2008).

Para a implementação do algoritmo de Canny são necessárias quatro etapas principais, a Suavização, que ocorre com um filtro gaussiano; a computação do gradiente; a não eliminação da amplitude do gradiente; a limiarização (SILVA e ALVES, 2008).

A principal vantagem desse processo é que o número de pontos conectados é maior, melhorando os contornos determinados e trazendo uma consistência maior nos resultados quando comparado com outros métodos. Na Figura 5 consegue-se notar a detecção das bordas e das texturas quando aplicado com o operador de Canny (PIZARRO, 2013), (SILVA e ALVES, 2008).

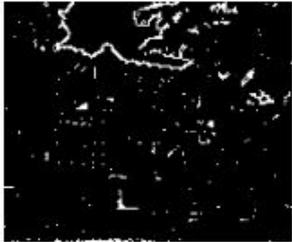
ALGORITHM	ORIGINAL IMAGE	EDGES	TEXTURES
Canny			

Figura 5. Representação do operador de Canny (NAIDU *et al*, 2012)

### 2.3.3 Métodos Avançados

Os métodos avançados computacionais serão descritos nos tópicos abaixo: Aprendizado de Máquina, Redes Neurais e Redes Neurais Convolucionais

### 2.3.3.1 *Aprendizado de Máquina*

O *Machine Learning* (Aprendizado de máquina) é uma técnica que dá a capacidade de aprendizado a uma máquina. Esse treinamento se dá através de um conjunto de objetos exemplo, que podem ser imagens, números, textos e outros métodos detectáveis. O treinamento da máquina pode ser feito tanto de forma supervisionada quanto autônoma, sendo a escolha baseada no intuito da detecção e dos dados disponibilizados para o aprendizado. No método supervisionado é criado um conjunto de regras para gerar um comportamento de treinamento a partir do processamento dos dados pelas regras. Nos métodos autônomos, os algoritmos de Aprendizado de máquina analisam os dados e os resultados esperados da análise e com isso cria as regras do aprendizado (MACHINE LEARNING E CIÊNCIA DE DADOS COM IBM WATSON).

### 2.3.3.2 *Redes Neurais*

As redes neurais são uma técnica de aprendizado de máquina que possuem seu nome e sua estrutura inspirados no cérebro humano, fazendo suas interações em forma de neurônios. As redes são treinadas com dados para aperfeiçoar a acurácia da detecção, um exemplo famoso de rede neural é o algoritmo de pesquisa do *Google* (REDES NEURAIIS, 2020).

As redes neurais funcionam com nós, que possuem seu sistema próprio de adaptação, baseado nos valores de entrada, em pesos determinados para os resultados, e saídas. Com os dados de entradas escolhidos, determina-se os pesos, que são utilizados para definir o quão impactante é o resultado da variável escolhida. Sabendo do peso determinado, multiplica-se pelos valores das entradas e depois são somados. A depender do tipo de rede, espera-se um comportamento para os nós após o processamento dos dados (REDES NEURAIIS, 2020).

Como dois exemplos de redes neurais, podemos citar as redes neurais convolucionais, que possuem maior destaque na detecção de imagens e na visão computacional, sendo esse destaque possível graças a utilização de princípios de álgebra linear, que através da multiplicação de matrizes conseguem encontrar padrões nas imagens. Outro exemplo de rede neural é a rede neural recorrente, que possuem como principal característica terem um loop

nos seus feedbacks. Com esses loops, a rede neural recorrente se torna interessante para previsões futuras, se utilizando de dados temporais em um treinamento repetitivo para prever alterações em ações no mercado financeiro, em previsões de vendas estipuladas etc (REDES NEURAS, 2020).

#### 2.3.3.3

#### *Redes Neurais Convolucionais*

As Redes Neurais Convolucionais foram baseadas no córtex visual humano, sendo especializadas em classificação de imagens e extensamente utilizadas no campo da visão computacional, pois foi desenvolvido com a capacidade de processamento em múltiplas dimensões. A arquitetura básica de uma rede neural convolucional consiste basicamente em três tipos de camadas: convolucionais, de pooling e rede totalmente conectada (CONVOLUTIONAL NEURAL NETWORKS).

**Camada convolucional:** Tem como objetivo a extração dos *features* de entrada, e para obter essas *features* ela se utiliza de filtros convolucionais, que percorrem os dados de entrada, analisando a largura, altura e profundidade da imagem, realizando assim a convolução dos dados. A cada informação de entrada processada no período de treinamento da rede, os filtros convolucionais vão se alterando a fim de se adaptar melhor às características dos dados de entrada, como as cores. Quanto mais entradas a rede analisar, mais os filtros irão aprender, e se tornarem capazes de analisar imagens e estruturas cada vez mais complexas (CONVOLUTIONAL NEURAL NETWORKS).

**Camada de pooling:** É uma camada que tem como objetivo a redução do tamanho dos dados de entrada, geralmente é utilizada após uma camada convolucional, para possibilitar que as próximas camadas de convolução sejam representadas de outra forma, a fim de que a rede consiga aprender diversas representações dos dados. Outro benefício do *pooling* é a redução da quantidade de dados para a próxima camada, pois melhora a regularização da rede e reduz o custo de processamento e memória.

Uma das técnicas mais utilizadas para o *pooling* é chamada de *max pooling*, e ela consiste em fragmentar os dados em subpartes, selecionando os maiores

valores das sub regiões, assim reduzindo o tamanho da imagem, conforme exemplificado na Figura 6.

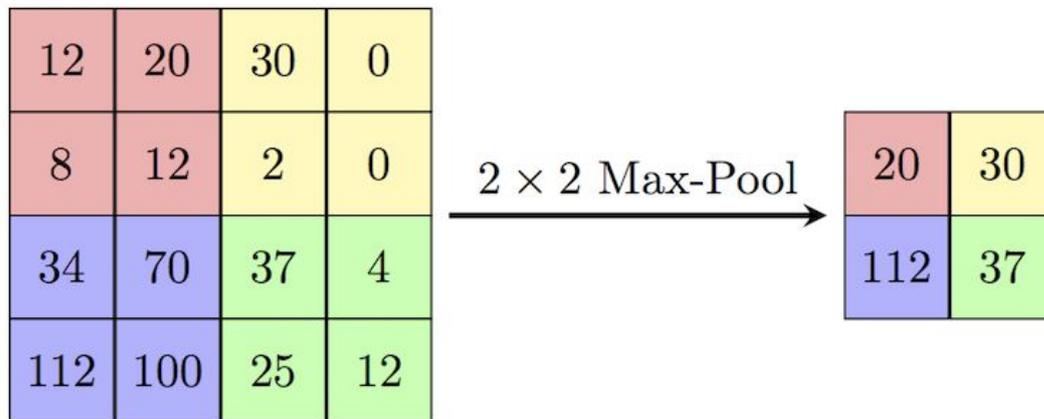


Figura 6. Exemplo de funcionamento do max pooling com filtro de 2x2 em uma imagem 4x4

**Camada totalmente conectada:** A camada totalmente conectada geralmente fica no final da rede, e recebe esse nome por ser onde ocorre a extração das características obtidas pelas camadas prévias, provendo a classificação final.

## 2.4 RECURSOS TECNOLÓGICOS

Nesta seção serão apresentadas as tecnologias utilizadas no desenvolvimento da pesquisa.

### 2.4.1 Python

O Python é uma linguagem de programação interpretada e orientada a objetos, que possui tipagem dinâmica e alto nível de tipos de dados. Como destaque possui uma capacidade de processamento altíssima com uma sintaxe limpa e intuitiva, além disso possui diversas bibliotecas e extensões com suporte nativo, tornando a linguagem simplificada para resolução de problemas para o programador. A linguagem possui utilizações diversas, sendo utilizada para conexões com protocolos na internet, para soluções de engenharia de dados, para interfaces de sistemas operacionais e vários outros usos, justamente pela alta performance aliada a bibliotecas e extensões poderosas nativas e de possuir grande apoio da comunidade (GENERAL PYTHON FAQ, 2022).

## 2.4.2 OpenCV

O OpenCV (*Open Source Computer Vision Library*) é uma biblioteca de código aberto de visão computacional, que foi desenvolvida para proporcionar uma estrutura de comum utilização para visão computacional e incentivar a disseminação e o uso da visão computacional em produtos comerciais. A biblioteca possui mais de 2500 algoritmos implementados, que abrangem as mais diversas técnicas de algoritmos, que podem ser utilizados para detecção de padrões, detecção de objetos, reconhecimento de ações e os mais diversos usos. Além disso, é uma tecnologia muito consolidada e difundida, sendo utilizada em empresas como Google, Yahoo e IBM, e também sendo utilizada por grupos de pesquisas e órgãos governamentais. O OpenCV possui suporte para C++, Python e Matlab, e pode ser utilizado em Windows, Linux e MacOS, ou seja, pode ser utilizada em praticamente qualquer computador (ABOUT OPENCV, 2022).

## 2.4.3 Face Mesh

O *Face Mesh* é uma solução de detecção facial que detecta 468 pontos faciais em tempo real. Para isso, utiliza-se de aprendizado de máquina para inferir a geometria da imagem nas três dimensões. Para fazer uma detecção facial precisa, o modelo faz a separação da face do resto do ambiente, melhorando a probabilidade do algoritmo encontrar os pontos faciais. A principal vantagem da utilização do *Face mesh* é a de possuir uma arquitetura leve, robusta e performática, fazendo com que a análise possua menos gasto de processamento e seja efetuada mais rapidamente.

Além disso, o *Face mesh* possui módulos que auxiliam na geometria facial, que são muito úteis na detecção em tempo real e facilitam a detecção facial (MEDIAPIPE FACE MESH, 2020).

## 3. METODOLOGIA

Nesta seção são descritos os procedimentos da pesquisa, a definição do padrão de captura, obtenção das imagens e desenvolvimento do algoritmo de detecção.

### 3.1 PROCEDIMENTOS DA PESQUISA

O presente estudo foi realizado durante o período de maio de 2021 a fevereiro de 2022, com 40 crianças incluídas, nenhuma criança foi excluída do estudo.

Os critérios de inclusão para participação das crianças no estudo foram:

1. Recém-nascido (RN) pré-termo, nascidos antes das 37 semanas de idade gestacional, e termo, nascidos após as 37 semanas de idade gestacional.
2. RNs internados na Unidade de tratamento intensiva (UTI) Neonatal ou alojamento conjunto do Hospital Regional Deputado Afonso Guizzo, Araranguá, Santa Catarina, ou será agendada uma consulta no ambulatório de *Follow up*.
3. Apresentar, no ato da avaliação, estabilidade dos sinais vitais, como frequência cardíaca e respiratória, e hemodinâmica, avaliada pela pressão arterial sistêmica;
4. Possuir ou não diagnóstico de alterações visuais, neurológicas e genéticas;
5. Ter o Termo de Consentimento Livre e Esclarecido assinado pelo responsável legal;
6. Apresentar o estado de alerta ativo conforme a escala de avaliação comportamental neonatal de Brazelton no momento da avaliação (BRAZELTON, 1995).

Os critérios de exclusão da participação no estudo foram:

1. RN no pós-operatório imediato de cirurgias cardíacas, neurológicas, gastrointestinais, uroginecológicas e ortopédicas;
2. RN com infecção ocular ou em uso de algum medicamento que comprometa a visão;

O estudo foi executado em três etapas, sendo elas: definição de padrão de captura, obtenção das imagens e desenvolvimento do algoritmo de detecção.

### 3.2 DEFINIÇÃO DE PADRÃO DE CAPTURA

A definição do padrão das capturas das imagens tem como propósito controlar a discrepância entre os vídeos capturados, a fim de garantir uma melhor performance no algoritmo de detecção das crianças. Para a criação do padrão de captura obteve-se apoio de uma equipe de profissionais da saúde, para que fosse estudado e testado diversas formas até se chegar no padrão final. Atualmente há dois padrões de capturas.

1. Vídeos de 30 segundos sem estímulos visuais- sendo três vídeos com *smartphone* e três vídeos com câmera infravermelho em cerca de 45 cm de distância dos olhos da criança.
2. Vídeos de 30 segundos, com o *smartphone* e a câmera infravermelho, e com a criança 45 cm de distância do avaliador sentada no colo de outro avaliador ou responsável. Primeiramente, a pessoa que está segurando a criança no colo realiza o movimento do reflexo vestibulo ocular três vezes para a direita e três vezes para a esquerda de forma alternada enquanto o avaliador que está a 45cm de distância realiza um estímulo frontal aos olhos. Após, o avaliador que está a 45cm de distância realiza estímulos nas extremidades do campo visual e por fim o avaliador realiza o movimento de perseguição com o estímulo indo da direita para a esquerda por três repetições.

### 3.3 OBTENÇÃO DAS IMAGENS

A fase de obtenção de imagens foi realizada por meio da captura de vídeos de crianças, entre 0 a 24 meses de idade, em Unidades Básicas de Saúde de Araranguá, Balneário Arroio do Silva e Sombrio e no Hospital Regional Deputado Afonso Guizzo após assinatura do Termo de Consentimento e Livre e Esclarecido (TCLE) em anexo, assinado pelos responsáveis. A filmagem foi realizada com auxílio de discentes e docentes do curso de Fisioterapia, Engenharia da Computação e TIC e aconteceu em ambiente controlado, com a criança em estado estável, acompanhada pelos pais e/ou responsáveis a fim de ter um *dataset* robusto e diverso para o desenvolvimento do algoritmo de detecção futuro treinamento de rede neural.

Para a obtenção das imagens, a avaliação foi separada em duas etapas. Na primeira etapa foi realizada a triagem das crianças, sendo realizada uma entrevista com os pais ou responsáveis sobre os dados gerais da criança e da mãe, possíveis diagnósticos médicos e questões da estabilidade clínica. Posteriormente, os pais ou responsáveis foram convidados a participarem do estudo e mediante a assinatura do TCLE, foi aplicado o protocolo de padrão de captura, especificado no item 3.2.

A coleta na Unidade de tratamento intensivo neonatal (UTIN) foi realizada no próprio leito do paciente, mais especificamente no berço aquecido. Já no alojamento conjunto a coleta

foi realizada na sala de procedimentos de enfermagem, com aquecimento interno. E no ambulatório de *follow up* foi utilizada uma sala de consultório, sendo que nestes ambientes estão disponíveis uma maca, mecanismo de aquecimento para manutenção da temperatura ambiente, cortinas para controle da luminosidade e restrição de funcionários para manter a atenção da criança durante a avaliação.

A captação das imagens relacionadas às respostas visuais das crianças foram realizadas através de webcams de alta resolução e dispositivo celular, posicionadas perpendicularmente em relação à face dos RN's utilizando para tal suportes e tripés adequados durante o processo de aquisição.

### 3.4 DESENVOLVIMENTO DO ALGORITMO DE DETECÇÃO

Com o início da obtenção das capturas, começa-se a alimentar um banco de vídeos, que será utilizado para o desenvolvimento do algoritmo de detecção. Conforme o banco de vídeos for expandindo, será possível fazer o treinamento de uma rede específica para a detecção infantil. Atualmente o *dataset* é utilizado para validação do algoritmo e para auxiliar na configuração de parâmetros do algoritmo e do OpenCV. Para a identificação facial está sendo utilizado o *FaceMesh*, que consiste em uma biblioteca que detecta os principais pontos faciais do rosto. Posteriormente o algoritmo detecta os olhos da rede facial e por fim faz-se a detecção da íris.

## 4. DESENVOLVIMENTO

Esse sistema e seu desenvolvimento foram divididos nos seguintes tópicos: levantamento de requisitos, desenvolvimento do sistema de detecção e detalhes do desenvolvimento e configuração dos algoritmos

### 4.1 LEVANTAMENTO DE REQUISITOS

Antes de o desenvolvimento do software ser iniciado de fato, ocorreram diversos testes e levantamentos de requisitos, a fim de criar um padrão de captura, para que os dados do banco de imagens tivessem a menor discrepância possível. Para isso foi-se testado diversas

câmeras, padrões de captura, padrões de estímulos visuais e técnicas de detecção visual, a fim de chegar a um padrão que garantisse a melhor qualidade o possível do banco de imagens.

Esse levantamento de requisitos foi necessário pois durante as capturas iniciais, notou-se que havia uma discrepância muito grande em diversos aspectos, como a iluminação do ambiente, a diferença de estrutura dos locais e a alta movimentação da criança, que geravam vídeos com discrepâncias elevadas, tornando difícil o desenvolvimento do algoritmo de detecção, por estar sem um ambiente de controle.

#### 4.2 DESENVOLVIMENTO DO SISTEMA DE DETECÇÃO

O sistema para detecção da visão infantil foi elaborado seguindo as etapas demonstradas na Figura 7. Estas etapas correspondem respectivamente em entrada do vídeo original no algoritmo, processamento de imagem (detecção e seleção do rosto da criança, detecção de pontos faciais específicos e a partir do reconhecimento da estrutura ocular faz-se o rastreamento das posições da íris), finalizando com o vídeo processado, contendo as informações das respostas visuais reconhecidas durante o processamento de imagem pelo *software*.



Figura 7. Etapas do desenvolvimento do sistema de detecção

A primeira etapa de detecção e seleção do rosto da criança acontece com a utilização da biblioteca *FaceMesh* aliada com o *OpenCV*. O *FaceMesh* faz a detecção do rosto e se utiliza de quatro bordas (extremidade superior direita, extremidade superior esquerda, extremidade inferior direita e extremidade inferior esquerda), formando um quadrado delimitando o rosto para alinhamento no centro da imagem com o *OpenCV*. Os parâmetros de

recorte foram aprimorados para o formato do rosto geralmente encontrado na população infantil, de acordo com o provimento de vídeos e imagens no banco das imagens capturadas.

Na próxima etapa, detecção de pontos faciais, foi utilizada a solução *FaceMesh*. Este procedimento tem como função encontrar os pontos de interesse faciais, incluindo cerca de 478 pontos, 468 para o rosto e 10 para os olhos. Os pontos detectados poderão ser utilizados para identificação de diversas estruturas faciais, como boca e olhos. Para a solução presente, foram utilizados os pontos para detecção dos olhos. O motivo de utilização da *FaceMesh* é a facilidade do seu manuseio, visto que basta importar a biblioteca para utilizar seus parâmetros, além disso, é uma solução com alta performance e robustez, que foi treinada e desenvolvida com o Google, sendo até mesmo utilizada em processos industriais.

Para a detecção dos olhos utilizaram-se os pontos de interesse recolhidos na fase anterior aliados com parâmetros de imagem do *OpenCV*. Com os parâmetros de imagem foram detectadas as bordas esquerda e direita do olho. Para encontrar a íris foi chamada uma função nativa do *FaceMesh*, que retorna o centro da íris e 4 pontos nas bordas da íris. Para obtenção do raio da íris foi utilizado a função de círculos do *OpenCV*, que calculava o raio a partir dos 5 pontos retornados (os 4 da borda e o central). Com o reconhecimento da íris, foi realizado o rastreamento das suas posições percorridas, utilizando-se dos valores salvos durante as movimentações ocorridas pela íris. Após isso, os valores obtidos são desenhados para demonstração para o usuário, como demonstrado na Figura 8.

```

def draw_iris_circles(image,left_iris,right_iris):
    cv2.circle(image, left_iris[0],left_iris[1], (255,0,255), 1, cv2.LINE_AA)
    cv2.circle(image, right_iris[0],right_iris[1], (255,0,255), 1, cv2.LINE_AA)
    return image

def draw_past_positions_iris_center(image,positions_data,max_number_draw):
    left_eye,right_eye = positions_data.get_past_n_positions(max_number_draw)

    for i in range(1,len(left_eye)):
        end_point=left_eye[i]
        start_point=left_eye[i-1]
        if i == 1:
            color = (255,0,0)
        else:
            color = (0,0,255)
        thickness =int(max_number_draw/(2*i+max_number_draw/3))+1
        image = cv2.line(image, start_point, end_point, color, thickness)

    for i in range(1,len(right_eye)):
        end_point=right_eye[i]
        start_point=right_eye[i-1]
        if i ==1:
            color = (255,0,0)
        else:
            color = (0,0,255)
        thickness =int(max_number_draw/(2*i+max_number_draw/3))+1
        image = cv2.line(image, start_point, end_point, color, thickness)
    return image

```

Figura 8. Imagem do código que desenha os valores obtidos pela aplicação.

### 4.3 DETALHES DO DESENVOLVIMENTO E CONFIGURAÇÃO DOS ALGORITMOS

O algoritmo é baseado em classes que atuam na detecção e rastreamento, cada um com sua responsabilidade. As principais características implantadas no algoritmo foram: módulo de posições percorridas, o *FaceMesh*, os ajustes faciais, e a detecção da íris.

O módulo de posições percorridas adiciona as posições percorridas e demonstra as N posições percorridas nos últimos frames, sendo o número de frames passados a ser demonstrado um parâmetro que pode ser configurado, para mais ou menos frames, assim podendo ser útil para diversas análises.

O módulo de detecção dos pontos faciais se utiliza da importação da biblioteca *FaceMesh*, que é responsável por detectar os pontos de interesse faciais.

Os ajustes faciais servem para padronizar a posição facial e dos olhos, e atuam alinhando os olhos, alinhando o rosto e separando apenas o rosto na imagem, descartando o cenário de fundo.

A detecção da íris, é feita com as posições dos 5 pontos encontrados pelo *FaceMesh* e assim calculando o raio da íris para exibição.

## 5. RESULTADOS PRELIMINARES

Este estudo apresentou como resultado o funcionamento do software de acordo com a proposta inicial e fases desenvolvidas para conseguir alcançar a função final de rastreamento das posições percorridas pela íris.

Nas Figuras 9, 10, 11, 12, 13 e 14 é apresentado a comparação das imagens sem o processamento do software na movimentação da íris e com o processamento aplicado para este rastreamento.



Figura 9. Imagem da estrutura ocular da criança A sem a aplicação do software

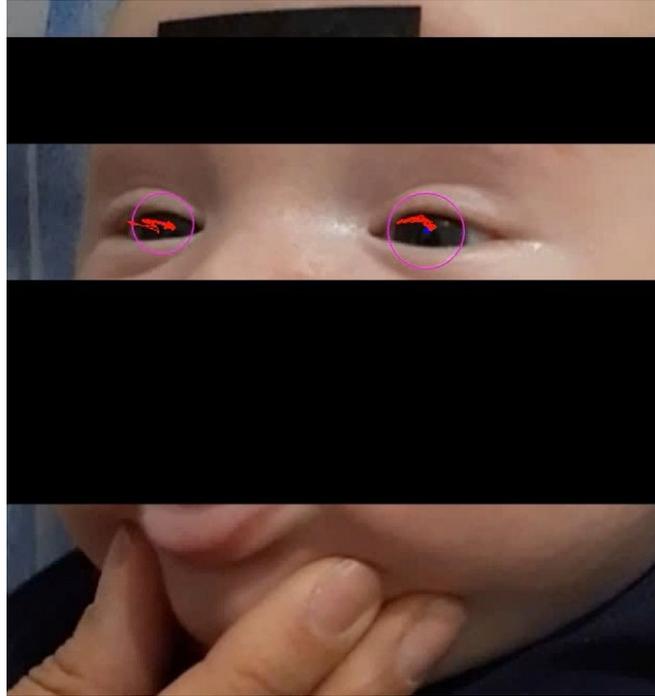


Figura 10. Imagem da resposta visual da criança A com a aplicação do software



Figura 11. Imagem da estrutura ocular da criança B sem a aplicação do software

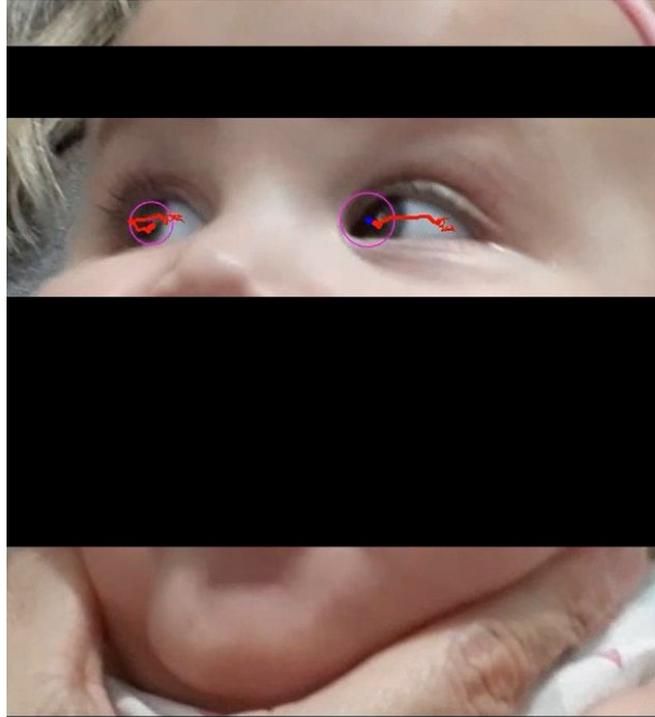


Figura 12. Imagem da resposta visual da criança B com a aplicação do software

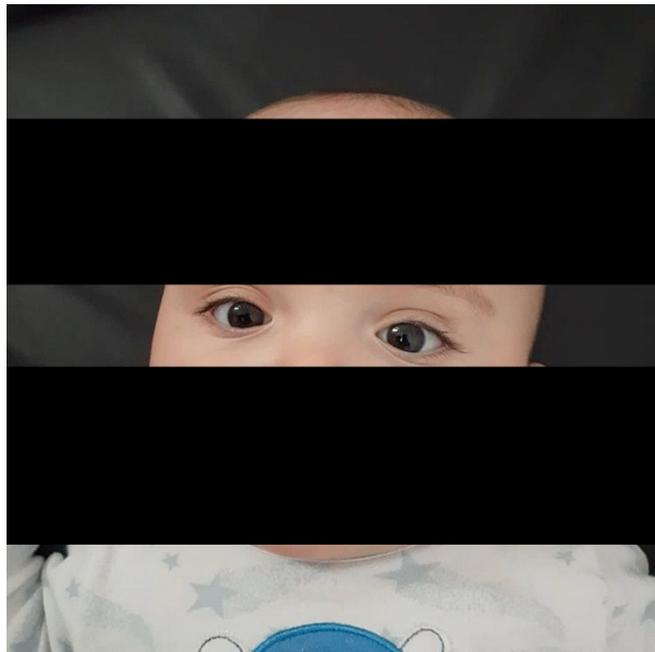


Figura 13. Imagem da estrutura ocular da criança C sem a aplicação do software



Figura 14. Imagem da resposta visual da criança C com a aplicação do software

De acordo com os exemplos das Figuras citadas acima, percebe-se o funcionamento do software em suas atribuições de detecção facial, detecção ocular e rastreamento da movimentação da íris pela linha vermelha formada na imagem, o que demonstra a qualidade do software visto que sofria com as limitações do ambiente externo, como irritabilidade da criança e iluminação do local.

## 6. CONCLUSÕES E DISCUSSÕES

Com os avanços recentes na tecnologia, em termos de hardware e de algoritmos de processamento de imagem, cada vez mais a visão computacional é viável nas mais diversas aplicações, sejam elas comerciais, pessoais ou em pesquisa.

O presente estudo traz como conclusão a execução eficaz do desenvolvimento de um software que auxilie na detecção facial e rastreamento visual da população infantil, além da aquisição de imagens ou vídeos do protocolo de bateria de testes visuais na população infantil para construção do *dataset* interno. Com o *dataset* interno, irá se desenvolver futuramente um modelo computacional para a tomada de decisões sobre possíveis atipicidades

visuais na população infantil. Portanto o projeto ainda está sendo executado e possui diversas novas etapas em concepção.

Com o rastreamento da movimentação da íris nos estímulos externos realizados, será possível, com a análise das respostas visuais pelos profissionais de saúde especializados na área visual infantil, detectar possíveis alterações visuais e desta forma encaminhar a criança para o tratamento de forma precoce.

O estudo, no futuro, pode aprofundar o funcionamento do software com o aumento na quantidade de vídeos no *dataset* e dessa forma aumentar a robustez nas respostas da íris e também a identificação e rastreamento de outras estruturas oculares, aumentando o número de atipicidades possíveis de serem detectadas. Ademais, o *software* pode evoluir para a versão *mobile*, tornando-se viável o seu uso na rotina dos profissionais de saúde, tanto no âmbito hospitalar, como ambulatorial, facilitando o diagnóstico funcional da visão de forma precoce e assim também o seu tratamento.

Os pontos limitantes do estudo incluem a dificuldade do software em delimitar o rosto da criança quando em algum momento do vídeo aparecia o rosto de um adulto, como os pais ou o avaliador, visto que os pontos extremos do rosto são diferentes e limitam esta detecção, além da necessidade de um maior tempo viável para robustez do software.

## REFERÊNCIAS

ABOUT OPENCV. **OPENCV**, 2022. Disponível em: <https://opencv.org/about/>. Acesso em: 12 de fevereiro de 2022.

BASIC THRESHOLDING OPERATIONS. **OPENCV**. Disponível em: <[https://docs.opencv.org/3.4/db/d8e/tutorial\\_threshold.html](https://docs.opencv.org/3.4/db/d8e/tutorial_threshold.html)>.

BRAZELTON, T.B. NUGENT, J.K. Neonatal Behavioral Assessment Scale: 3ed. Londres, **Clinics Developmental Medicine** n°137, 1995.

CONFERENCE: 9TH INTERNATIONAL CONVENTION ON REHABILITATION ENGINEERING AND ASSISTIVE TECHNOLOGY (I-CREATE). **Development of a low-cost, portable, tablet-based eye tracking system for children with impairments**, 2015, Singapore.

CONVOLUTIONAL NEURAL NETWORKS. **CS231N**. Disponível em: <<https://cs231n.github.io/convolutional-networks/>>. Acesso em 20 de fevereiro de 2022.

DIRETRIZES DE ATENÇÃO À SAÚDE OCULAR NA INFÂNCIA: DETECÇÃO E INTERVENÇÃO PRECOCE PARA A PREVENÇÃO DE DEFICIÊNCIAS VISUAIS. **Ministério da Saúde**, 2ª edição Brasília – DF 2016. Disponível em: <[https://bvsmms.saude.gov.br/bvs/publicacoes/diretrizes\\_atencao\\_saude\\_ocular\\_infancia.pdf](https://bvsmms.saude.gov.br/bvs/publicacoes/diretrizes_atencao_saude_ocular_infancia.pdf)>.

DONG, Chuan-Zhi; CATBAS, F Necati. A review of computer vision–based structural health monitoring at local and global levels. **Structural Health Monitoring**, v. 20, n. 2, 2021. Disponível em: <<https://journals.sagepub.com/doi/full/10.1177/1475921720935585>. Acesso em: 28 ago. 2021.

GAGLIARDO, Heloisa. G. R. Gardon; NOBRE, Maria. Ines. R. S. **Intervenção Precoce na Criança com Baixa Visão**. **Revista Neurociências**, v. 9, n. 1, p. 16–19, 2001. Disponível

em: <https://periodicos.unifesp.br/index.php/neurociencias/article/view/8928>. Acesso em: 1 set. 2021.

GAGLIARDO, Heloisa; GONÇALVES, Vanda; LIMA, Maria Cecília. Método para avaliação da conduta visual de lactentes. **Arq.Neuropsiquiatr**, v.62, n.2-A, pag.300-306, 2004.

GENERAL PYTHON FAQ. **Python**, 2022. Disponível em: < <https://docs.python.org/3/faq/general.html#id2>>. Acesso em 8 de fevereiro de 2022.

GONZALEZ, Rafael; WOODS, Richard. **Digital Image Processing**. Prentice Hall, 2 edição, 2002.

GONZALEZ, Rafael C.; WOODS, Richard. **Digital Image Processing**. 3. ed. [S.L]: Pearson Universidades, 2008.

GONZALEZ, Rafael. **Processamento Digital de Imagens**. Editora Pearson, 2009.

GRAZIANO, Rosa Maria; LEONE, Cléa Rodrigues. Problemas oftalmológicos mais freqüentes e desenvolvimento visual do pré-termo extremo. **J Pediat (Rio J)**, v.81, n.1, pag. 95-100, 2005.

HAUX, Reinhold. Health Information Systems - from Present to Future?. *Methods of Information in Medicine*. **The German Medical Informatics Initiative**, [S. l.], p. 43-45, 2 jun. 2018.

IMAGE SEGMENTATION WITH WATERSHED ALGORITHM. OPENCV, 2022. Disponível em: < [https://docs.opencv.org/4.x/d3/db4/tutorial\\_py\\_watershed.html](https://docs.opencv.org/4.x/d3/db4/tutorial_py_watershed.html)>.

MACHINE LEARNING E CIÊNCIA DE DADOS COM IBM WATSON. **IBM**. Disponível em: <<https://www.ibm.com/br-pt/analytics/machine-learning>>. Acesso em 14 de janeiro de 2022.

MEDIAPIPE FACE MESH. **MEDIAPIPE**, 2020. Disponível em: < [https://google.github.io/mediapipe/solutions/face\\_mesh.html](https://google.github.io/mediapipe/solutions/face_mesh.html)>. Acesso em: 20 de fevereiro de 2022.

NAIDU, B. Ramesh; RAO, P. Lakshman; BABU, M.S. Prasad, *et al.* Framework for Efficient Edge Detection Techniques– Comparison Among Robert, Prewitt, Sobel, Robinson, Kirsch and Canny. **IJCST**, v.3, n.2, 2012. Disponível em: < <http://ijest.com/vol32/6/bramesh.pdf>>.

Acesso em: 20 de março de 2022.

PIXFORCE. Os olhos das máquinas superando a visão humana. Disponível em: <https://pixforce.com.br/visao-computacional-os-olhos-das-maquinas-superando-a-visao-humana/>. Acesso em: 6 set. 2021.

PIZARRO, Paulo; COMUNELLO, Eros; MANTELLI, Sylvio. **Introdução a visão computacional**, 2013. Disponível em: <<https://www.inf.ufsc.br/~aldo.vw/visao/regiongrow.pdf>>.

PRIESE, L; REHRMANN, V. A Fast Hybrid Color Segmentation Method, **In: S.J. Pöppel and H.Handels, editors**, Mustererkennung 1993, pages 297-304. Springer Verlag, 1993.

QUEIRÓZ, José Eustáquio Rangel de; GOMES, Hermes Martins. Introdução ao Processamento Digital de Imagens. **Revista RITA**, v.8, n.1, 2001.

RAMOS, André. Fisiologia da visão: um estudo sobre o “ver” e o “enxergar”. **PUC- RIO**, 2006. Disponível em: [https://sites.unifoa.edu.br/portal/plano\\_aula/arquivos/04054/Fisiologia%20da%20visao%20-%20MODULO%20I.pdf](https://sites.unifoa.edu.br/portal/plano_aula/arquivos/04054/Fisiologia%20da%20visao%20-%20MODULO%20I.pdf). Acesso em: 5 de março de 2022.

REDES NEURAIAS. **IBM**, 2020. Disponível em: <

<https://www.ibm.com/br-pt/cloud/learn/neural-networks>>. Acesso em 2 de fevereiro de 2022.

SEGMENTAÇÃO POR QUANTIZAÇÃO CROMÁTICA BASEADA EM IMAGENS-J (J-SEG). **Image processing and computer graphics lab**. Disponível em: < <https://lapix.ufsc.br/ensino/visao/segmentacao-por-quantizacao-cromatica-baseada-em-imagens-j-jseg/>>. Acesso em: 13 de janeiro de 2022.

SHAN, M-W. Discovering color styles from fine art images of impressionism. **International Journal of Computer Science and Security**, v.3, n.4, 2009. Disponível em: <[https://www.researchgate.net/publication/41822925\\_Discovering\\_Color\\_Styles\\_from\\_Fine\\_Art\\_Images\\_of\\_Impressionism](https://www.researchgate.net/publication/41822925_Discovering_Color_Styles_from_Fine_Art_Images_of_Impressionism)>.

SILVA, F.J.V; ALVES, C.H.F. Aplicação de técnicas de processamento de imagens digitais em imagens geradas por ultra-som. **8 o Encontro Regional de Matemática Aplicada e Computacional**, Natal, 2008.

SILVA, Cibele Maria Ferreira da; ALMEIDA, Driellen Rodrigues de; BERNARDES, Rafael Ribeiro et al. **School performance: visual acuity interference**. Rev Bras Oftalmol., v. 72, n. 3, p. 168-171, Jun. 2013.

SILVA, Tarcízio et al. APIs de Visão Computacional: investigando mediações algorítmicas a partir de estudo de bancos de imagens. **Logos**, [S.l.], v. 27, n. 1, jun. 2020. Disponível em: <<https://www.e-publicacoes.uerj.br/index.php/logos/article/view/51523>>. Acesso em: 08 set. 2021.

SONKSEN, Patricia M; PETRIE, Aviva; DREW, Kristina J. Promotion of visual development of severely visually impaired babies: evaluation of a developmentally based programe. **DevMed Child Neurol**, v.33, p.320-35, 1991.

SOUSA, Sidney Júlio de Faria e. Fisiologia e desenvolvimento da visão physiology and development of vision. **Ribeirão Preto**, v. 30, pág. 16-19, jan./mar. 1997.

TIBES, Chris Mayara dos Santos; DIAS, Jéssica Davis; ZEM-MASCARENHAS, Silvia Helena. Aplicativos móveis desenvolvidos para a área da saúde no Brasil: revisão integrativa da literatura. **Revista Mineira de Enfermagem**, [S. l.], p. 471-478, 1 abr. 2014. Disponível em: <https://cdn.publisher.gn1.link/reme.org.br/pdf/v18n2a16.pdf>. Acesso em: 31 ago. 2021.

TOLEDO, Carolina Cumani et al. Detecção precoce de deficiência visual e sua relação com o rendimento escolar: study in A. **Revista da Associação Médica Brasileira [online]**, v. 56, n.

4, p. 415-419, 2010. Disponível em: <<https://doi.org/10.1590/S0104-42302010000400013>>. Acesso em: 28 ago. 2021.

ZIMMERMANN, Anita et al. Visual development in children aged 0 to 6 years. **Arquivos Brasileiros de Oftalmologia [online]**. 2019, v. 82, n. 03 , pag. 173-175, 2019. Disponível em:< <https://www.scielo.br/j/abo/a/pRR9rBksv6DRZWTP9WypFFs/?lang=en#>>. Acesso em: 05 de março de 2022.

## ANEXO 1 – Termo de Consentimento Livre e Esclarecido

### Termo de Consentimento Livre e Esclarecido

Você como representante legal do seu filho está sendo convidado a autorizar a participação dele no projeto de pesquisa da Universidade Federal de Santa Catarina, Universidade de São Paulo e Universidade Federal do Rio Grande do Norte conforme os dados abaixo.

1 – Título do projeto: IMPACTO DA COVID-19 GESTACIONAL NO DESENVOLVIMENTO SENSORIO-MOTOR E VISUAL DE RECÉM-NASCIDOS ATRAVÉS DE TÉCNICAS COMPUTACIONAIS DE ANÁLISE DE IMAGEM

2 – Desenho do estudo e objetivo(s): Estudo de coorte prospectiva, com o objetivo de avaliar o crescimento e desenvolvimento de crianças nascidas de mães que tiveram COVID-19 durante a gestação; comparar o desenvolvimento entre crianças brasileiras nas regiões sul, sudeste e nordeste, desenvolvimento de uma técnica computacional para avaliar a visão por meio de imagens.

3 – Descrição dos procedimentos: As crianças irão participar deste estudo através de convite direto aos pais ou responsáveis legais, durante as primeiras horas após o nascimento. A participação será voluntária, a amostragem por conveniência e o recrutamento ocorrerá a partir do consentimento dos pais depois da assinatura do Termo de Consentimento Livre e Esclarecido. As famílias de crianças nascidas no período da pandemia (a partir do dia 11/03/2020 até data a ser definida pela OMS), de gestantes positivas para COVID-19, serão convidadas para participar deste estudo. Esta coorte será recrutada a partir dos dados do prontuário de internação e será comparada com as crianças incluídas no grupo caso e controle, de faixa etária semelhante, nascidas na mesma região e de nível socioeconômico semelhante.

As avaliações serão realizadas em todos os participantes do estudo. As consultas para o acompanhamento do crescimento e desenvolvimento serão realizadas nos tempos de até 72 horas após o nascimento, durante o 1º, 3º, 6º, 12º, 18º e 24º mês da data do parto. As fichas de Avaliação e a avaliação da funcionalidade por meio da aplicação da Classificação Internacional da Funcionalidade (CIF) serão preenchidas apenas uma vez no início da pesquisa e as consultas serão agendadas e realizadas no ambulatório de seguimento das maternidades com vínculo com as instituições de ensino.

A ficha de Avaliação Inicial será composta por itens relacionados com os dados clínicos e possíveis intercorrências referentes ao recém-nascido, dados socioeconômicos e demográficos dos responsáveis como o Nível de Atividade Física e dados antropométricos da mãe.

O exame neurocomportamental de recém-nascidos, será realizado pela avaliação neurológica de Hamermith, composta por 34 itens, e dividida em 6 categorias que avalia os músculos, a postura, comportamento e movimento da criança.

Já a escala qualitativa dos movimentos espontâneos generalizados (GM) será utilizada para complementar o diagnóstico precoce de sequelas neurológicas, sendo indicado para crianças de 36 horas vida até 20 semanas de idade cronológica.

Para a avaliação do desenvolvimento infantil será utilizada a Escala Bayley de desenvolvimento infantil, que busca identificar atrasos de desenvolvimento infantil de crianças entre 1 a 42 meses, e avalia a linguagem, os movimentos, a coordenação e o comportamento do recém-nascido.

Para as avaliações o recém-nascido será filmado de barriga para cima com apenas uso de fraldas.

Os riscos podem envolver inicialmente a irritabilidade dos recém-nascidos que serão avaliados pelos especialistas; caso isto ocorra, medidas de conforto serão realizadas pelo avaliador juntamente do responsável (acalmar no colo, balanceio, sucção não nutritiva entre outras medidas). Ainda, bebês extremamente prematuros poderão apresentar alterações de frequência cardíaca e/ou respiratória diante dos estímulos motores e/ou visuais. Caso ocorra alterações nos sinais vitais (frequência cardíaca, frequência respiratória, saturação de oxigênio), a avaliação será interrompida até que o bebê retorne à sua normalidade e, se necessário, a avaliação será remarcada. Os pesquisadores utilizarão equipamentos de proteção individual (EPIs) com o objetivo de não proporcionar risco de contaminação pelo COVID-19.

Os benefícios do projeto incluem a saúde e bem-estar por possibilitar o diagnóstico precoce de possíveis alterações sensório-motoras e visuais de recém-nascidos de mães positivas para o COVID19 e favorecer a aplicação de terapêuticas possibilitando a redução de déficits futuros. E também proporcionar o treinamento de qualidade aos pesquisadores, a cooperação entre universidades e a troca de conhecimento entre especialistas, com o objetivo de fomentar a saúde da região da AMESC.

Os pesquisadores serão os únicos a ter acesso aos dados e tomarão todas as providências necessárias para manter a privacidade e o sigilo dos dados dos participantes da pesquisa, mas sempre existe a remota possibilidade da quebra do sigilo, mesmo que involuntariamente e não intencional, cujas consequências serão tratadas nos termos da lei.

Os resultados deste trabalho poderão ser apresentados em encontros ou revistas científicas da área, sendo que serão mostrados apenas os resultados obtidos como um todo, sem revelar seu nome, o nome do seu filho, instituição ou qualquer informação relacionada à sua privacidade.

A legislação brasileira não permite que você tenha qualquer compensação financeira pela sua participação em pesquisa e você não terá nenhuma despesa advinda da participação na pesquisa. Caso alguma despesa extraordinária associada à pesquisa venha a ocorrer, você será ressarcido nos termos da lei, pois caso você tenha algum prejuízo material ou imaterial em decorrência da pesquisa poderá solicitar indenização, de acordo com a legislação vigente e amplamente consubstanciada por via judicial.

A participação do participante é voluntária, não sendo ofertado ônus ou benefício financeiro pela participação. Assim como todos os gastos relacionados a pesquisa são de responsabilidade dos pesquisadores, não sendo responsabilidade do participante as despesas ou gastos.

Acredito ter sido suficientemente informado a respeito das informações que li ou que foram lidas para mim, descrevendo o estudo "Impacto da covid-19 gestacional no desenvolvimento sensório-motor e visual de recém-nascidos através de técnicas computacionais de análise de imagem".

Eu discuti com os pesquisadores sobre a minha participação no estudo. Ficaram claros para mim quais são os objetivos do estudo, as avaliações sobre a saúde, os desconfortos e riscos, as garantias de confidencialidade e de esclarecimentos permanentes. Ficou claro também que a participação é voluntária e gratuita. Concordo voluntariamente em participar deste estudo e poderei mudar de ideia a qualquer momento, antes ou durante o estudo, sem penalidades ou prejuízo ou perda de qualquer benefício que eu possa ter ganho. Desta forma, estou ciente que posso deixar de participar da pesquisa a qualquer momento, sem ter que apresentar qualquer justificativa.

O pesquisador responsável, que também assina esse documento, compromete-se a conduzir a pesquisa de acordo com o que preconiza a Resolução 466/12 de 12/06/2012, que trata dos preceitos éticos e da proteção aos participantes da pesquisa. Além disso, duas vias do termo de consentimento devem ser rubricadas e assinadas por você e pelo pesquisador responsável. Você deve guardar cuidadosamente a sua via, pois é um documento que traz importantes informações de contato e garante os seus direitos.

Você poderá entrar em contato com a pesquisadora Dra Cristiane Aparecida Moran – Universidade Federal de Santa Catarina. Departamento de Ciências da Saúde – Campus Jardim das Avenidas – Rodovia Governador Jorge Lacerda, 3201 – Jardim das Avenidas – CEP 88900-000. Araranguá – Santa Catarina – Brasil, E-mail: cristiane.moran@ufsc.br. Você também poderá entrar em contato com o Comitê de Ética em Pesquisa com Seres Humanos da UFSC pelo telefone 55 48 3721-6094 e-mail cep.propesq@contato.ufsc.br ou pessoalmente na rua R: Desembargador Vitor Lima, nº 222, 4º andar, sala 401 Trindade 88040-400 – Florianópolis – Santa Catarina.

Eu \_\_\_\_\_, documento de identidade (RG) \_\_\_\_\_, responsável legal pela criança \_\_\_\_\_, participarei do projeto de livre e espontânea vontade, de acordo com a autorização assinada. Declaro que obtive de forma apropriada e voluntária o Consentimento Livre e Esclarecido do participante para a participação neste estudo.

\_\_\_\_\_  
Assinatura do responsável pelo estudo

Data / /

\_\_\_\_\_  
Assinatura do responsável legal

Data / /