

**MIRIAN DE CESARO REVERS BIASÃO**

**Classificação da gravidade do transtorno do espectro autista baseada  
no padrão de rastreamento do olhar**

Dissertação apresentada à Faculdade de Medicina da  
Universidade de São Paulo para obtenção do título de Mestre  
em Ciências Programa de Psiquiatria Orientadora: Profa.  
Dra. Helena Paula Brentani

(Versão corrigida. Resolução CoPGr 6018/11, de 1 de novembro de 2011. A versão original  
está disponível na Biblioteca da FMUSP)

**SÃO PAULO 2019**

**Dados Internacionais de Catalogação na Publicação (CIP)**

Preparada pela Biblioteca da  
Faculdade de Medicina da Universidade de São Paulo

©reprodução autorizada pelo autor

Biasão, Mirian de Cesaro Revers

Classificação da gravidade do transtorno do espectro autista baseada no padrão de rastreamento do olhar / Mirian de Cesaro Revers Biasão. -- São Paulo, 2019.

Dissertação(mestrado)--Faculdade de Medicina da Universidade de São Paulo.

Programa de Psiquiatria.

Orientadora: Helena Paula Brentani.

Descritores: 1.Transtorno do espectro autista  
2.Fixação ocular 3.Índice de gravidade de doença  
4.Aprendizado de máquina supervisionado  
5.Classificação 6.Rastreamento do olhar

USP/FM/DBD-305/19

Responsável: Erinalva da Conceição Batista, CRB-8 6755

Para Beto, Mateus e Miguel,

*home is where you are.*

## **Agradecimentos:**

À Deus, por guiar essa experiência e me cercar de pessoas excepcionais.

À minha orientadora, Professora Helena Brentani, por mudar minha forma de pensar e perceber o mundo. Obrigada por acreditar e confiar em meu potencial antes de mim mesma. Obrigada pela paciência, pelas conversas, pelo afeto.

Ao Protea, com vocês aprendi ver de forma multi, e que a colaboração traz os melhores resultados.

Ao PsysBio, pela recepção, paciência e auxílio nas minhas tentativas de programação.

Ao grupo da EACH, especialmente a Jéssica, pela parceria desde a captação dos dados até a análise.

À Eliza e Isabel pelo suporte e cuidado durante todo o mestrado, vocês tornaram tudo menos complicado.

Aos pacientes e famílias, obrigada pela confiança e carinho.

À minha família, principalmente ao Beto, por embarcar nessa aventura comigo, Mateus e Miguel, por mudarem minha perspectiva e mostrar que tudo pode ser mais intenso e divertido e a minha mãe, maior e primeira incentivadora. Sem vocês nada disso teria sentido.

Aos meus amigos, pelos ouvidos, colos e risadas. A vida é muito melhor com vocês por perto.

O presente trabalho foi realizado com apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - Brasil (CAPES) - Código de Financiamento 001.

*“ O essencial é invisível aos olhos. Só se vê bem com o coração.”*

*Antoine de Saint Exupery*

## **Normalização Adotada**

Esta dissertação ou tese está de acordo com as seguintes normas, em vigor no momento desta publicação:

Referências: adaptado de International Committee of Medical Journals Editors (Vancouver) e APA quinta edição.

Universidade de São Paulo. Faculdade de Medicina. Divisão de Biblioteca e Documentação. Guia de apresentação de dissertações, teses e monografias. Elaborado por Anneliese Carneiro da Cunha, Maria Julia de A. L. Freddi, Maria F. Crestana, Marinalva de Souza Aragão, Suely Campos Cardoso, Valéria Vilhena. 3a ed. São Paulo: Divisão de Biblioteca e Documentação; 2011.

Abreviaturas dos títulos dos periódicos de acordo com List of Journals Indexed in Index Medicus.

## Sumário:

Lista de abreviaturas

Lista de figuras

Lista de tabelas

Resumo

Abstract

<b>1. INTRODUÇÃO</b>	<b>1</b>
1.1. Importância do olhar	1
1.2. Rastreamento do olhar	3
1.3. Processamento do olhar	7
1.4. Modelo computacional de processamento de imagens	11
1.5. Transtorno do espectro autista	12
1.6. Rastreamento do olhar no transtorno do espectro autista	17
<b>2. OBJETIVOS</b>	<b>20</b>
2.1. Objetivo principal	20
2.2. Objetivos específicos	20
<b>3. MATERIAL E MÉTODOS</b>	<b>21</b>
3.1. Tipo de estudo	21
3.2. População de estudo	21
3.3. Critérios de inclusão	21
3.4. Critérios de exclusão	21
3.5. Instrumentos	21
3.5.1. Childhood Autism Rating Scale	21
3.5.2. Rastreamento do olhar	24
3.5.3. Paradigma de preferência visual	24
3.6. Análise dos dados	25
3.6.1. Análise ROIs	26
Construção de mapas de saliência e Análise com aprendizado de máquina supervisionada	27
<b>4. RESULTADOS</b>	<b>32</b>
4.1. Caracterização da amostra	32
4.2. Gravidade X resultados do paradigma social processado como feito na literatura, ou seja, usando ROIs.	33
4.3. Resultados do classificador a partir dos mapas de saliência	35
<b>5. DISCUSSÃO</b>	<b>41</b>
<b>6. CONCLUSÃO</b>	<b>45</b>
<b>7. ANEXOS</b>	<b>46</b>
<b>8. BIBLIOGRAFIA</b>	<b>54</b>

## Lista de abreviaturas

<b>TEA</b>	<b>Transtorno do espectro autista</b>
<b>ROI</b>	<b>do inglês, <i>Regions of interest</i></b>
<b>FIT</b>	<b>do inglês, <i>Feature Integration Theory</i></b>
<b>ADOS</b>	<b>do inglês, <i>Autism Diagnostic Observation Schedule</i></b>
<b>WASI</b>	<b>do inglês, <i>Wescher Abbreviated scale of intelligence</i></b>
<b>PROTEA</b>	<b>Programa de pesquisa em Transtorno do Espectro Autista da faculdade de medicina da Usp</b>
<b>CARS</b>	<b>do inglês, <i>Childhood Autism Rating Scale</i></b>
<b>RGB</b>	<b>do inglês, Red, Green, Blue</b>
<b>SVM</b>	<b>do inglês, <i>Support Vector Machine</i></b>
<b>ROC</b>	<b>do inglês, <i>Receiver Operating Characteristic</i></b>



## Lista de Figuras

<b>Figura 1</b> - Lente para captação do rastreamento do olhar .....	<b>4</b>
<b>Figura 2</b> - Esquema da captação remota utilizando reflexo da fonte de luz....	<b>5</b>
<b>Figura 3</b> - Exemplo de ROIs.....	<b>7</b>
<b>Figura 4</b> - Esquema dos circuitos cerebrais responsáveis pelo processamento visual.....	<b>8</b>
<b>Figura 5</b> - Foto do guarda roupa com casaco saliente.....	<b>9</b>
<b>Figura 6</b> - Modelo da Teoria das características integradas (FIT).....	<b>10</b>
<b>Figura 7</b> - Modelo de Saliência Visual.....	<b>12</b>
<b>Figura 8</b> - Paradigma de preferência visual.....	<b>25</b>
<b>Figura 9</b> - ROIs utilizadas.....	<b>26</b>
<b>Figura 10</b> - Esquema das etapas realizadas no processamento de dados...	<b>28</b>
<b>Figura 11</b> - Mapa de fixação de cada grupo.....	<b>36</b>
<b>Figura 12</b> - Mapas de saliência.....	<b>37</b>
<b>Figura 13</b> - Mapa de saliência com fixações sobrepostas.....	<b>38</b>
<b>Figura 14</b> - Curvas ROC com os dados do classificador.....	<b>39</b>

## **Lista de Tabelas**

<b>Tabela 1 - Nível de gravidade do TEA.....</b>	<b>16</b>
<b>Tabela 2 - Caracterização da amostra .....</b>	<b>32</b>
<b>Tabela 3 - Correlação entre os escores de ROIs.....</b>	<b>34</b>
<b>Tabela 4 - Regressão da associação dos escores e a CARS.....</b>	<b>35</b>

## Resumo

Bisão MCR. *Classificação da gravidade do transtorno do espectro autista baseada no padrão de rastreamento do olhar* [dissertação]. São Paulo: Faculdade de Medicina, Universidade de São Paulo; 2019.

O transtorno do espectro autista cursa com alterações precoces na percepção visual, culminando com déficits na comunicação social e comportamentos restritos e estereotipado. Dados objetivos sobre o padrão visual dos indivíduos são obtidos através da técnica de rastreamento do olhar. Sabe-se que a técnica é eficaz para identificar indivíduos com TEA quando comparados a controles, mas, ainda não há trabalhos que utilizem esses dados a fim de classificar subtipos do transtorno. O objetivo deste estudo é preencher essa lacuna, e utilizar os dados de rastreamento do olhar associado a padrões de aprendizado de máquina a fim de classificar subgrupos de TEA, quanto a gravidade. Para tanto, foi utilizado modelo baseado em mapas de atenção visual, que utiliza os dados da captação sem filtros. O classificador foi testado pelo método de validação cruzada. Os resultados mostraram que foi possível classificar em TEA grave e não grave com média de 85% de precisão, atingindo o máximo de 88% de precisão, 87% de sensibilidade e 60% de especificidade. Espera-se que novos estudos, envolvendo número maior de indivíduos e outras características fenotípicas, possam ser desenvolvidos utilizando esta técnica, a fim de identificar biomarcadores para o transtorno.

**Descritores:** Transtorno do espectro autista; Fixação ocular; Índice de gravidade de doença; Aprendizado de máquina supervisionado; Classificação; Rastreamento do olhar.

## Abstract

Bisão MCR. *Classification of the severity of autistic spectrum disorder based on the eye tracking pattern* [dissertation]. São Paulo: “Faculdade de Medicina, Universidade de São Paulo”; 2019.

Autism spectrum disorder presents with early alterations in visual perception, culminating with deficits in social communication and restricted and stereotyping behaviors. Objective data on the visual pattern of individuals are obtained through the technique of eye tracking. It is known that the technique is effective to identify individuals with ASD in comparison with controls, but there are still no studies that use these data to classify subtypes of the disorder. The purpose of this study is to fill this gap, and to use the data from the eye tracking associated with machine learning patterns in order to classify subgroups of ASD as the severity. For this, a model based on visual attention maps was used, which uses the captured data without filters. The classifier was tested by the cross-validation method. The results showed that it is possible to classify in severe and non-severe ASD with an average of 85% accuracy, reaching a maximum of 88% accuracy, 87% sensitivity and 60% specificity. It is hoped that new studies, involving larger numbers of individuals and other phenotypic characteristics, could be developed using this technique in order to identify biomarkers for the disorder.

**Descriptors:** Autism spectrum disorder; Fixation, ocular; Severity of illness index; Supervised machine learning; Classification; Eye tracking.

## **1. Introdução**

### **1.1- Importância do olhar**

Humanos são considerados mamíferos sociais e, desde o nascimento, o contato social é crucial para o desenvolvimento e manutenção da espécie (Insel & Fernald, 2004). A vida em sociedade traz muitos benefícios: variabilidade genética ao propiciar maior número de encontros, facilita o ataque a predadores e favorece a aprendizagem entre indivíduos. Por outro lado, viver em grandes grupos também demanda habilidades específicas, como reconhecer hierarquias e a capacidade de prever comportamentos dos pares. (Emery, 2000)

A percepção do mundo, envolve, além de rastrear o ambiente, observar o sistema social. Ou seja, reconhecer e prever o comportamento do outro, bem como o do grupo, e é essencial para que a resposta aos eventos seja adequada (Lorenz, 1970). Os sistemas sensoriais são os instrumentos de entrada para a percepção, e são usados pelas diferentes espécies de forma única. Assim, cada uma capta o mundo da forma mais apropriada para seu grupo. O desenvolvimento do fenótipo das espécies está relacionado a necessidade e intensidade dos diversos sistemas sensoriais (Kaas, 1989).

Um marco evolutivo dos primatas foi o uso do sistema visual como guia para o comportamento (Ghazanfar & Santos, 2004). No cérebro dos primatas, há mais de 30 regiões associadas a estímulos visuais, em detrimento de outros estímulos, como olfativo e tátil (Felleman & Van, 1991). Sinais visuais são menos ambíguos e podem informar sobre o estado mental e intenções do outro. Esse tipo de

sinalização, tão complexa, não é encontrada em outras espécies de vertebrados(Senju & Johnson, 2009).

Circuitos neurais visuais traduzem informação sensorial, principalmente a proveniente dos outros membros do grupo, em informações usadas para orientar as futuras ações. Isso ocorre em vias diferentes e integradas, uma mais primitiva, subcortical e rápida, que identifica objetos em movimento e faces e outra mais refinada e cortical que direciona a atenção visual relacionada ao contexto(Klein et al, 2009). A face é uma região prioritária, por oferecer grande quantidade de informação social. Através dela, podemos identificar gênero, idade, emoções e inferir o estado mental.

Apesar dos olhos serem muito menores que a face, são eles os responsáveis por transmitir a maior parte das informações. Para isso, é necessário, captar, integrar, associar as mensagens e por fim, usar o próprio olhar para informar(Emery, 2000). Alterações na anatomia e musculatura da cabeça, que ocorreram durante o processo evolutivo, vem favorecendo maior destaque aos olhos e maior amplitude de expressões faciais(Perrett & Emery, 1994).

Em humanos, as maçãs do rosto proeminentes, nariz protuso e sobrancelhas funcionam como molduras para os olhos, indicando o foco atencional. A musculatura facial, que é quase inexistente em não primatas, desenvolveu-se e especializou-se, sendo capaz de produzir diferentes expressões faciais(Huber, 1931).

A anatomia dos olhos, não acompanhou essas mudanças, permanecendo praticamente estável entre primatas e humanos. A única diferença, é o tamanho na íris em relação a esclera. Humanos são os únicos animais em que é possível

visualizar esclera dos dois lados da íris a maior parte do tempo. Propiciando maior clareza e efetividade da comunicação visual(Kobayashi & Kohshima, 2001).

Sabe-se, que o olhar é uma das primeiras formas de interação do bebê com a mãe, e, durante toda a vida, é responsável pela percepção e orientação dos indivíduos no mundo social(Haith et al, 1977). O contato ocular indica interesse na interação, enquanto a ausência do contato pode significar rejeição social(Mason et al, 2005), além disso, o olhar é um sinalizador das intenções do outro, transmitindo segurança(Bayliss et al, 2006) e acessibilidade(Bayliss et al, 2007). Apesar do estímulo e resposta visual estarem diretamente relacionados ao ambiente, fatores genéticos influenciam a preferência atencional, tempo e direcionamento dos movimentos oculares individuais. Estudos com gêmeos mostram que a preferência do direcionamento visual para os olhos e boca é a característica altamente herdável, e está diminuída nos indivíduos com transtorno do espectro autista(TEA)(Constantino et al, 2017, Kliemann et al, 2010).

A detecção da direção do olho é parte fundamental e antecessora do mecanismo de atenção compartilhada e teoria da mente. Para perceber em que direção as outras pessoas estão olhando, compartilhar informações e associar aos possíveis estados mentais envolvidos, é necessário detectar a face e os olhos e associar experiências prévias, para, então, definir qual a resposta adequada naquele momento(Baron-Cohen, 1997). Captar o trajeto do olhar do indivíduo nos permite identificar qual é o objeto de sua atenção e inferir sobre a sua percepção da cena apresentada(Duchowski, 2007). Características como a percepção do movimento biológico e ritmo, também percebidas pelo olhar, nos dão dicas importantes quanto a predição do comportamento do outro.



## 1.2 - Rastreamento do olhar

A técnica de *Eye Tracking* permite que o rastreamento ocular seja registrado de forma objetiva e acurada. Há mais de trinta anos, é utilizada em diversas áreas com diferentes objetivos. Ciência da computação, engenharia, *marketing*, entretenimento, neurociência, psicologia são algumas especialidades que utilizam a captação ocular para obter dados sobre a interação visual e os estímulos oferecidos(Duchowski, 2002).

Apesar da técnica ser utilizada desde a década de 70, foi o aparecimento da tecnologia associada ao *eye tracking* que permitiu a expansão e garantiu a utilidade da técnica para pesquisa clínica em diversos campos, sendo um meio de acessar a cognição através do comportamento visual (Duchowski, 2007). Através do movimento ocular é possível avaliar atenção, processamento de informações, planejamento e execução de tarefas(Koch & Ullman, 1987).

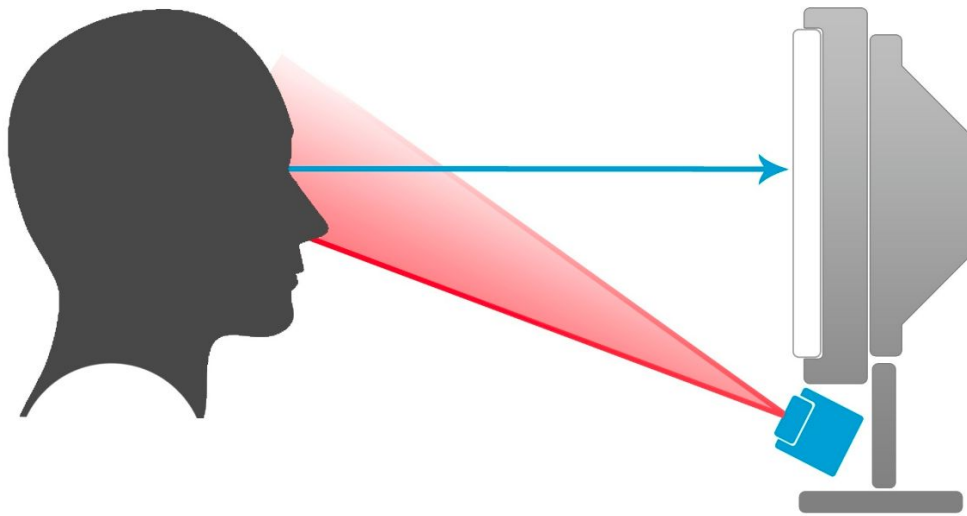
Na década de 50, as medidas de rastreamento do olhar eram obtidas através de um dispositivo acoplado a lentes de contato (Figura 1). Apesar de muito acuradas, tinham a desvantagem de ser invasivos.



Fonte:Skalar Medical, Delft, The Netherlands (<http://www.skalar.nl>)

**Figura 1** - Lente para captação do rastreamento do olhar. Exemplo de dispositivo intra ocular antigamente utilizado para captação do rastreamento do olhar.

Para resolver essa questão, os métodos não invasivos, ou remotos, utilizam como medida características visíveis do olho, como a pupila, fronteira entre a íris e a esclera ou reflexo da córnea a fonte de luz (Figura 2)(Duchowski, 2007).



Fonte: <http://mescla.cc>

**Figura 2** - Esquema da captação remota utilizando reflexo da fonte de luz. O aparelho emite luz infravermelha que é refletida pela pupila indicando a direção do olhar.

Hoje rastreio do movimento ocular é uma técnica não invasiva, de fácil aplicação que possibilita captação de dados objetivos e quantitativos sobre o processo visual e atencional dos indivíduos (Kienzle et al, 2007).

O dispositivo de captação do olhar pode ser ajustado no próprio indivíduo, através de óculos ou capacete. Outra possibilidade é usar o mesmo acoplado a uma tela onde são apresentados os estímulos desejados.

Neste caso, para captar as informações visuais, os olhos são iluminados por uma luz infravermelha que reflete a córnea e a pupila. Esses dados são captados por câmeras. A fim de minimizar erros de captação causados por movimentos da cabeça, diversas câmeras captam os dados que são processados de forma unificada. O uso de softwares específicos permite a integração das informações e o processamento das mesmas. A partir disso, é possível determinar os movimentos oculares de acordo com a imagem apresentada(Duchowski, 2007, Peters & Itti, 2007).

Apesar de realizar a captação com acurácia, os dispositivos remotos podem gerar dados com ruídos, causados, principalmente, pelos movimentos de cabeça. O software tem estratégias para lidar com esses ruídos, mas, isto pode ser um problema em populações que têm dificuldades em permanecer paradas ou obedecer a comandos, como crianças e pessoas com transtornos psiquiátricos(Holmqvist et al, 2011).

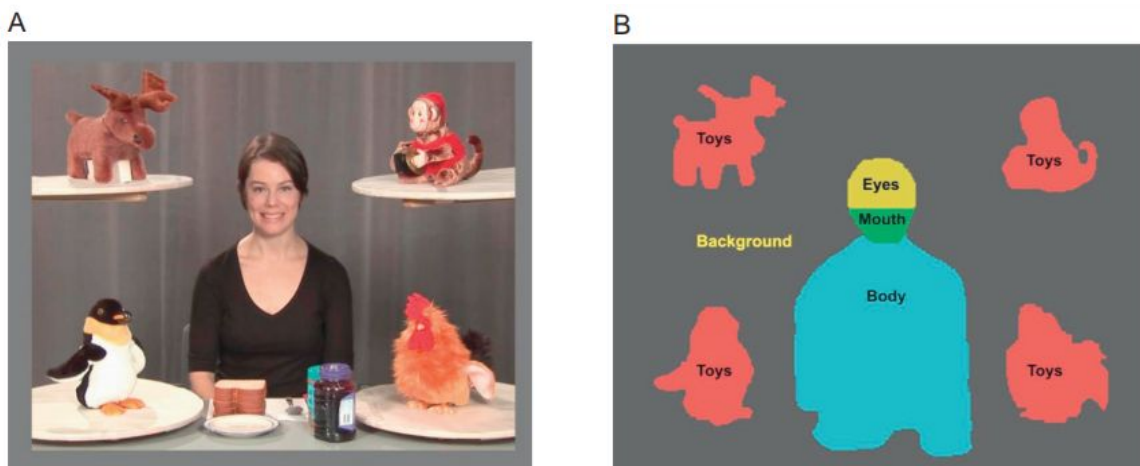
Dependendo do objetivo do estudo, o estímulo apresentado pode ser estático ou dinâmico. Os paradigmas variam desde estímulos que exigem apenas que o sujeito olhe para tela ou podem ser acompanhados a perguntas ou indicações para tarefas específicas. Essa técnica pode ser aplicada em qualquer idade e para qualquer nível de funcionamento intelectual, uma vez que exige apenas o

engajamento de olhar para a tela, uma tarefa relativamente simples (Gredeback et al, 2009).

Os dados captados pelo aparelho podem ser extraídos na forma bruta, em que as fixações são dadas por coordenadas x e y por intervalo de tempo, ou podem ser filtrados, neste caso, relacionados a regiões específicas, determinadas previamente pelo pesquisador.

Diversos trabalhos que utilizam dados de rastreamento do olhar, utilizam essas regiões de interesse específicas para análise dos dados, chamadas de “Regions of interest” (ROI) . Para tanto, é necessário, marcar, com ajuda do software, os locais sobre os quais os dados serão fornecidos. (Munn et al, 2008)

A figura 3 é um exemplo da marcação das regiões de interesse.



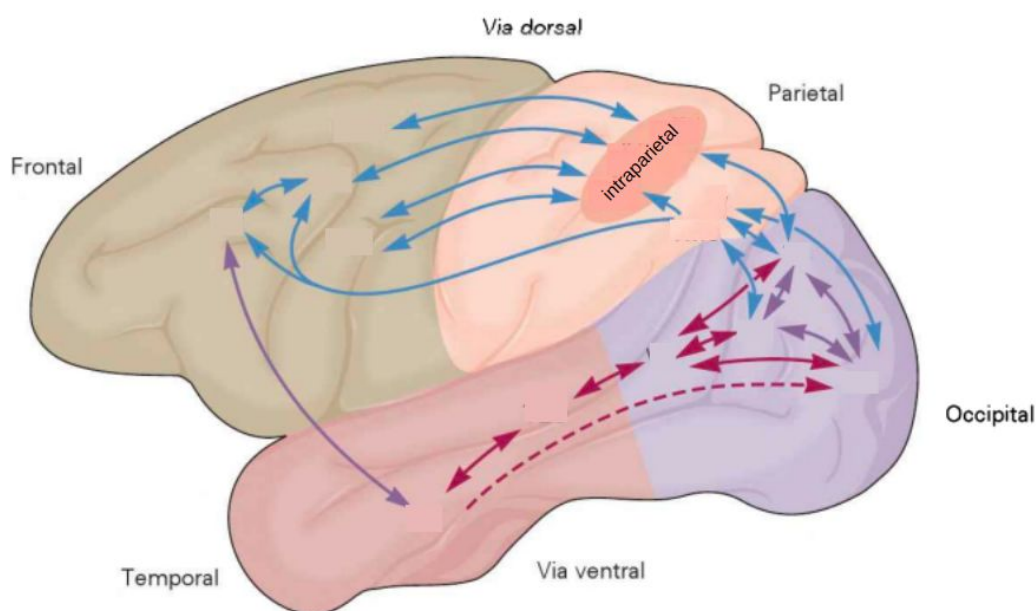
Fonte: Chawarska, 2013

**Figura 3** - Exemplo de ROIs. A. Estímulo apresentado ao indivíduo, B. Regiões de interesses marcadas manualmente com a ajuda do software específico. (Chawarska et al, 2013)

As ROIs são definidas manualmente, sendo necessário portanto assumir antecipadamente regiões de interesse para o objetivo do trabalho, limitando a análise dos dados a esses locais, em detrimento da imagem como um todo. O uso de ROIs leva em conta a duração das fixações, o local das mesmas e as sacadas que podem dar uma ideia do trajeto percorrido, porém não levam em conta modelos de atenção visual, baseados no conhecimento do processamento de imagens e redes cerebrais da neurofisiologia do olhar.

### **1.3 - Processamento do olhar**

Os estímulos visuais são processados por redes paralelas do córtex cerebral. A corrente ventral, tem como função analisar e codificar as informações referentes a forma e estrutura da cena e dos objetos, enviando a informação para o córtex para-hipocampal e córtex pré-frontal. Outra corrente, paralela a ventral é a dorsal, que analisa localização e movimento dos estímulos, enviando a informação a áreas motoras do córtex, que controlam o movimento do olhar e do corpo. As conexões entre as áreas são recíprocas, e envolvem circuitos de pró ação e retroalimentação (Figura 4).



Fonte: Adaptado de Kandel, 2014

**Figura 4** - Esquema dos circuitos cerebrais responsáveis pelo processamento visual. As setas bordô representam a corrente ventral, enquanto as setas azuis representam a corrente dorsal. Os lobos frontais são representados pelas cores: Marrom: Lobo frontal, Rosa: Lobo Temporal, Roxo: Lobo Occipital, Salmão: Lobo Parietal. A região em destaque no lobo parietal representa regiões intraparietais, que não estariam visíveis no plano do esquema (Kandel et al, 2014).

O olho humano recebe grande quantidade de informação visual, cerca de  $10^9$  bits por segundo (Koch et al, 2006), e para processá-la de forma eficiente e rápida, utiliza mecanismos cognitivos e complexos para selecionar o que é mais relevante no momento (Borji & Itti, 2012). Evolutivamente, esse mecanismo foi essencial, pois permite que predadores, parceiros e alimentos sejam identificados rapidamente, mesmo em ambientes com sobrecarga visual (Treisman & Gelade, 1980). A retina é guiada pela atenção visual a fim de captar os estímulos do ambiente de forma organizada e dando relevância ao que é mais importante.

Isso só acontece graças a integração de sistemas cerebrais conhecidos como *bottom-up* e *top-down*. A atenção visual pode ser orientada pelos sistemas *bottom-up*, rápidos e involuntários, baseado nas características vindas do objeto como cor, orientação, profundidade da cena, movimento. Outra forma de orientação é *top-down*, realizada por circuitos cognitivos, dependentes de experiências prévias, conhecimento, expectativas e recompensas. Estes são orientados a tarefa, mais lentos e voluntários (Itti, 2005).

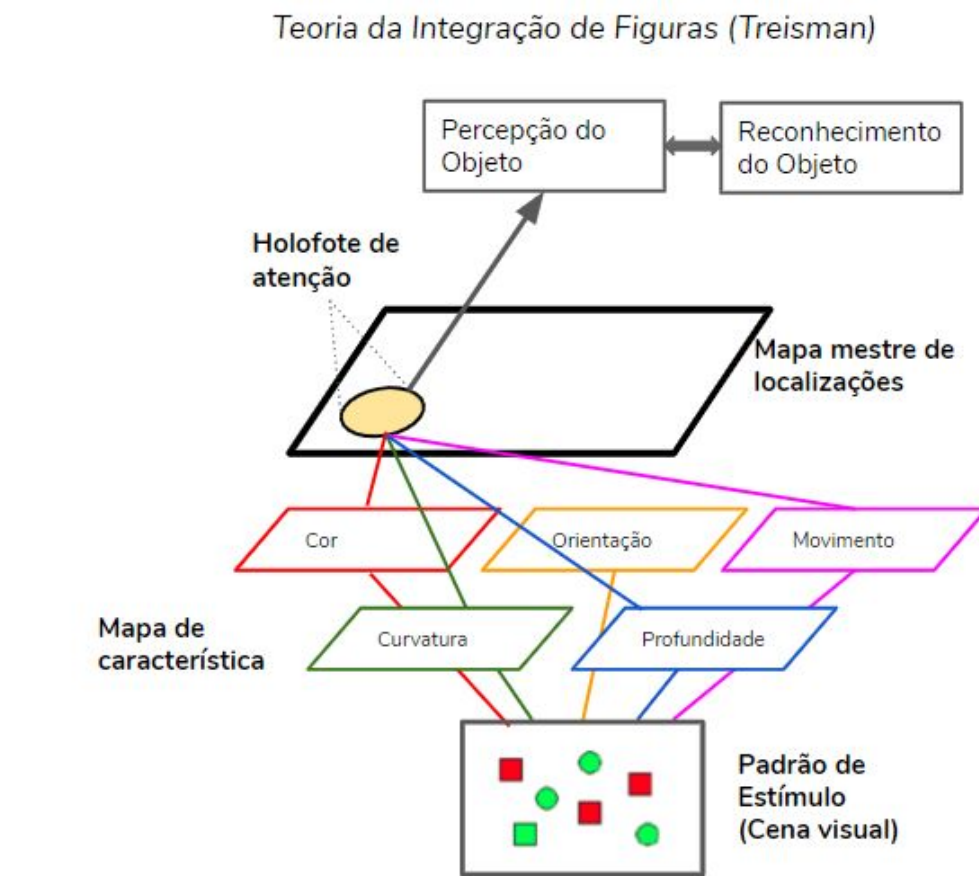
Assim, por exemplo, ao olhar a cena da figura 5, o sistema *bottom-up* direciona a atenção para o casaco vermelho. Já para responder a pergunta: “Qual a cor do armário?” o sistema *top-down* é o responsável por direcionar a atenção ao móvel.



Fonte: Rigo, 2010

**Figura 5** - Foto do guarda roupa com casaco saliente. Usada para exemplificar como o olhar é direcionado para regiões específicas, dependendo da tarefa a ser executada. (Rigo, 2010)

A teoria das características integradas (*Feature Integration Theory – FIT*) representado na figura 6, afirma que os estímulos captados pelo campo visual são processados em dois passos. O primeiro, avalia características básicas da imagem, como: cor, forma e tamanho. A partir dessas características, é formado um mapa da imagem, que indica onde estão os objetos, mas não identifica o que são. A partir desse mapa, é localizado o ponto de maior relevância, onde deve ser focada a atenção, e então, os objetos desse local específico são reconhecidos (Treisman & Gelade, 1980).



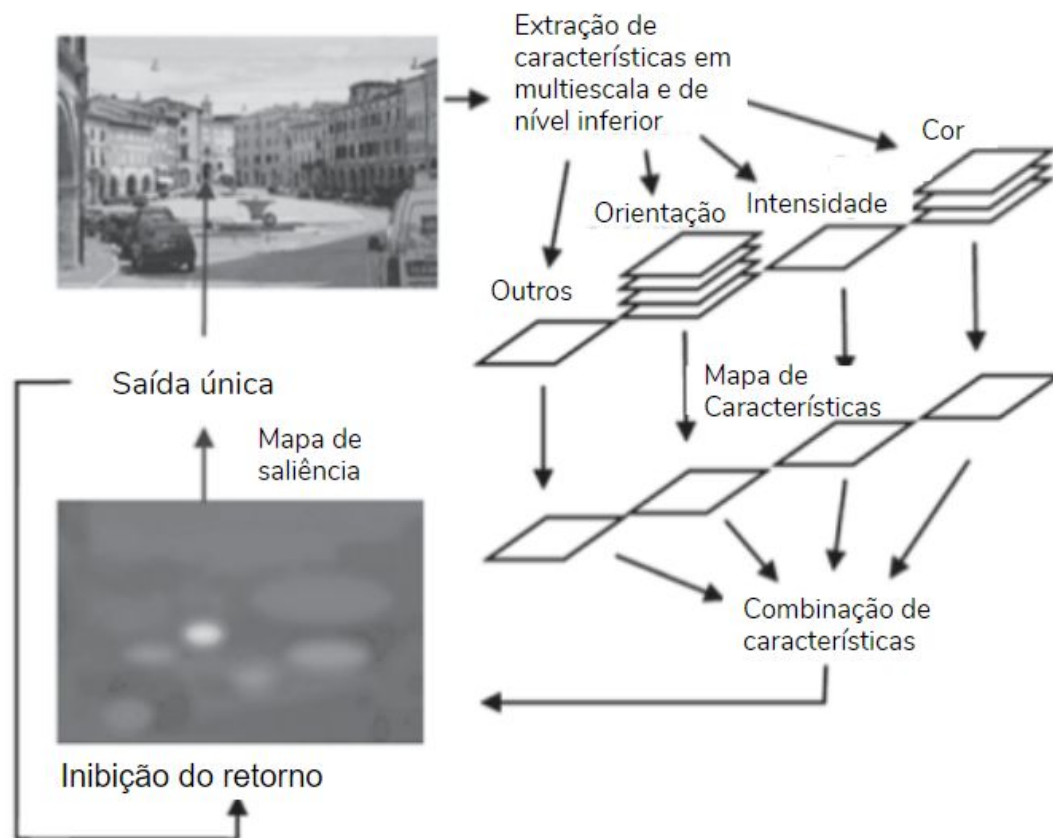
Fonte: adaptado de Koch, C., & Ullman, S. (1987)

**Figura 6** - Modelo da Teoria das características integradas (FIT). Representação da percepção visual através de mapas de características.



#### **1.4 - Modelo computacional de processamento de imagens**

Considerando a FIT, foi desenvolvido um modelo computacional de processamento de imagens, em que primeiro se avalia as características básicas, para, então, gerar um mapa de saliência, que indica o local mais relevante da imagem(Koch & Ullman, 1987). Hoje, existem diferentes categorias de modelos computacionais de atenção visual. Uma delas, o modelo de mapas de saliência, esquematizado na figura 7, utiliza técnicas de aprendizado de máquina para aprender a saliência dos estímulos, considerando dados de rastreamento do olhar ou rótulos em regiões salientes. Diversas técnicas podem ser utilizadas para aprender a saliência das cenas: *patches* de imagens(Kienzle et al, 2007), essência da cena(Peters & Itti, 2007), vetor de características por pixel(Judd et al, 2009), ou mapa de saliências(Borji, 2012). O modelo de mapa de saliências utiliza dados do movimento do olhar combinados com outras características, para predizer a atenção visual humana.



Fonte: Itti and Koch's, 2010

**Figura 7** - Modelo de Saliência Visual. A imagem é decomposta em características menores que juntas formam mapas de características. A combinação desses resulta no mapa de saliência, em que há uma saída única, informando o local mais relevante da cena. (Itti et al, 2008)

## 1.5 - Transtorno do Espectro Autista

Desde as primeiras descrições de Autismo, em 1943, Kanner indica alterações na percepção visual, com pouca ou nenhuma saliência aos estímulos sociais, e prejuízo no contato ocular nos indivíduos diagnosticados com o transtorno(Kanner, 1943).

A prevalência do TEA é de 1-69(Christensen et al, 2018), ocorrendo cerca de quatro vezes mais em meninos do que em meninas. Essa diferença é menor nos casos mais graves e aumenta quanto menor a gravidade do transtorno(Lord, 2982). No Brasil, um estudo piloto identificou prevalência de 1 caso a cada 370 crianças (Paula et al, 2011).

Hoje é descrito como um transtorno do neurodesenvolvimento poligênico e multifatorial(Chauhan & Chauhan, 2015), caracterizado por alterações qualitativas na comunicação social, como, falta de reciprocidade social, dificuldade em fazer e manter relacionamentos sociais, além de comportamentos restritos e estereotipados, que incluem, comportamentos ritualizados, forte aderência a rotina e interesses fixos(Brentani et al, 2013).

Alterações precoces no neurodesenvolvimento, a partir da oitava semana de gestação, no período de migração neuronal, podem levar a um déficit inibitório no cérebro de pessoas com TEA(Rubenstein & Merzenich, 2003). Sabe-se que, a função cerebral, depende da organização dos neurônios em minicolunas (Casanova & Casanova, 2019). Essa organização, garante que o impulso elétrico siga uma direção específica, sem estimular neurônios vizinhos de forma aleatória. Estudos de histopatologia, com tecido cerebral de indivíduos diagnosticados com TEA mostraram minicolunas menores, com neurônios em maior número e mais dispersos(Casanova et al, 2006).

O desenvolvimento cerebral, que se reflete na complexidade de comportamentos, apresenta inicialmente alta conectividade de curtas distâncias e

ao longo do tempo conectividade de longas distâncias vão sendo favorecidas, com desenvolvimento da mielinização das mesmas. Assegurando assim, maior eficiência na condutividade dos estímulos elétricos. Desta forma a integração de sistemas sensoriais, motores e cognitivos sustenta o desenvolvimento da percepção interna do próprio organismo como também da percepção externa do ambiente.

Um possível resultado das alterações anátomo patológicas observadas em TEA relaciona-se com um cérebro com maior conectividade a curtas distâncias e pouca conectividade a longa distância, que discrimina muito os estímulos, mas generaliza pouco(Casanova et al, 2006, Geschwind & Levitt, 2007). E é essa falha na integração dos sistemas sensoriais, motores e cognitivos que resulta, posteriormente, em déficits sociais e comportamentos restritos característicos do transtorno, com possível falha, ou desajuste, do sistema rudimentar de orientação aos estímulos sociais(Elsabbagh & Johnson, 2016). Um padrão complexo de interação entre os sistemas sociais e de percepção deve estar na base dos sintomas observados nessa condição(Zikopoulos & Barbas, 2013).

Estudos mostram que crianças posteriormente diagnosticadas com TEA mostram menor orientação para estímulos sociais, como atenção reduzida ao nome antes do primeiro ano de vida(Chawarska et al, 2013; Dawson et al, 2004; Yirmiya et al, 2006; Dawson et al, 1998). Uma possibilidade é que o interesse aumentado em objetos seja consequência do desinteresse em pessoas(Pellicano & Burr, 2012).

Essa hipótese corroborada por dados de neuroimagem que demonstram que em pacientes com TEA, regiões relacionadas a percepção da voz

humana(Saitovitch et al, 2012), identificação de faces, pessoas e movimento biológico, são acionadas de forma atípica(Ecker et al, 2015).

O sulco temporal superior (STS) tem papel fundamental na cognição social, e no TEA pode apresentar diminuição na concentração de substância cinzenta, hipoperfusão em repouso e ativação anormal durante tarefas sociais(Zilbovicius et al, 2006). Crianças com TEA podem apresentar hipoconectividade do STS com hemisfério esquerdo, áreas bilaterais ventrais tegmentares e núcleo accumbens, responsáveis pelo sistema de recompensa. Assim como menor conectividade com o hemisfério direito, córtex orbitofrontal e amígdala, regiões críticas para aprendizagem associativa relacionada a emoção, sugerindo que a percepção social não ativa os circuitos de recompensa que trariam reforço positivo aos estímulos, e também não aciona as regiões de aprendizados necessárias para o processamento da informação social(Abrams et al, 2013).

Apesar de compartilharem os mesmos sintomas nucleares e alterações fisiopatológicas, indivíduos com TEA apresentam grande heterogeneidade clínica e diferentes respostas ao tratamento(Masi et al, 2017). Ainda não existem medicações específicas para o tratamento dos sintomas nucleares do transtorno, e, mesmo as terapias bem estabelecidas de intervenção comportamental precoce intensiva e treino comportamental naturalístico, tem diferentes graus de efetividade(Zwaigenbaum et al, 2015; Wong et al, 2015; Howes et al, 2018). Além de impactar na prática clínica, a heterogeneidade dificulta as pesquisas, dada a impossibilidade de estabelecer uma população de estudo homogênea para *clinical trials*, por exemplo. Em razão disso, diversos estudos visam encontrar

biomarcadores para o transtorno, a fim de identificar subtipos e melhorar o entendimento sobre o TEA e a indicação terapêutica adequada(Loth et al, 2018).

Atualmente, existem cinco especificadores do diagnóstico(APA, 2013): 1- com ou sem comprometimento intelectual; 2- com ou sem comprometimento de linguagem; 3- associado a condição médica, genética ou fator ambiental conhecido; 4- presença de comorbidades e 5- presença de catatonia.

O transtorno também pode ser classificado em três níveis de gravidade, descritos na tabela 1: 1, 2 e 3; de acordo com a necessidade de suporte que o indivíduo apresente durante suas atividades cotidianas(Lord & Bishop, 2015).

**Tabela 1 - Nível de gravidade do TEA**

<b>Gravidade</b>	<b>Comunicação social</b>	<b>Comportamentos restritos e estereotipados</b>
<b>Nível 1</b>  Exige suporte	Déficit na comunicação social causam prejuízos notáveis. Pode ter desinteresse nas relações interpessoais, dificuldade em iniciar o contato, respostas atípicas ou sem resposta a abertura social	Inflexibilidade causa interferência em um ou mais contextos. Problemas de organização, planejamento, dificuldade em trocar de atividades.

	de outros.	
<b>Nível 2</b> Exige suporte substancial	Déficit grave na comunicação social que causam limitações mesmo com apoio. Resposta reduzida ou anormal ao contato do outro.	Inflexibilidade, dificuldade de lidar com mudanças, comportamentos restritos e repetitivos que interferem em diversos contextos e são óbvios para o observador casual. Sofrimento para mudar o foco das ações.
<b>Nível 3</b> Exige suporte muito substancial	Déficits graves na comunicação verbal e não verbal com limitação para iniciar contato interpessoal e resposta mínima ao contato de outros.	Inflexibilidade extrema, dificuldade de lidar com mudanças, comportamentos restritos e repetitivos que interferem de forma acentuada em todas as esferas. Grande sofrimento para mudar o foco das ações.

FONTE: APA, 2013

Legenda: Níveis de diagnóstico do TEA de acordo com os dois domínios diagnósticos.

## 1.6. Rastreamento do olhar no transtorno do espectro autista

Um dos sintomas característicos do TEA é a alteração da atenção visual, menos voltada para estímulos sociais quando comparado com pessoas sem o transtorno(Wang et al, 2015). Essa alteração está presente, precocemente, nos indivíduos afetados e persiste ao longo da vida. Crianças que posteriormente foram diagnosticadas com TEA apresentavam dificuldades de atenção para cenas sociais aos seis meses de idade quando comparadas com lactantes saudáveis(Pelphrey et al, 2002).

A comparação do padrão do olhar de adultos com TEA do nível 1 e controles ao analisarem imagens com expressões emocionais foi um dos primeiros trabalhos realizados com *Eye Tracking* e TEA, e o resultado foi que os indivíduos com TEA olham menos para a região dos olhos, nariz e boca que os controles(Naber et al, 2008).

Estudos com crianças mostraram que, os que tinham diagnóstico de TEA apresentaram atenção social reduzida(Sacrey et al, 2013) e dificuldade de mudar o foco atencional(Swanson & Siller, 2013), além de demorar mais tempo para dirigir o olhar para essas regiões(Freeth et al, 2010). As diferenças iniciam-se aos seis meses de idade, quando as crianças afetadas apresentavam menor atenção voltada a cenas sociais e as faces(Kou et al, 2019).

O paradigma de preferência visual mostrou ser o mais adequado para estudos com essa população(Kou et al, 2019). Esse paradigma também foi utilizado



para associar o padrão de rastreamento do olhar e a gravidade dos sintomas do TEA, independente da complexidade das cenas sociais apresentadas no vídeo(Moore et al, 2018). Além de separar as crianças com TEA de outros transtornos do neurodesenvolvimento, a preferência pelos padrões geométricos foi maior nas crianças com maior comprometimento da linguagem e comportamento social e pior nível cognitivo(Pierce et al, 2016).

Em adultos, utilizando o paradigma de observação livre de cenas sociais e também de movimento biológico, os dados do rastreamento apontaram associação entre o padrão do olhar e os domínios de comunicação da “*Autism Diagnostic Observation Schedule*” (ADOS), o teste padrão ouro para diagnóstico de TEA, e nos testes de performance do “*Wescher Abbreviated scale of intelligence*”(WASI)(Del Valle Rubido, 2018).

O uso do rastreamento ocular é uma excelente ferramenta para investigar como o indivíduo com TEA direciona a sua atenção visual diante de cenas cotidianas ou na execução de determinada tarefa. Mostra-se promissor como uma medida objetiva e quantitativa que pode auxiliar a tomada de decisão quanto ao diagnóstico e indicação de tratamento(Frazier et al, 2018).

Alguns trabalhos já mostraram associação do rastreamento do olhar com sintomas relacionados a comunicação e cognição de indivíduos com TEA(Pierce et al, 2016; Del Valle Rubido, 2018; Frazier et al, 2018), mas nenhum utilizou essa técnica a fim de classificar subtipos do transtorno assim como nenhum usou mapas de saliência visual e nem ROIs para classificação de TEA.

A lacuna que esse projeto visa preencher está relacionada tanto ao processamento dos sinais do movimento ocular utilizando os dados brutos através da implementação de modelos de atenção visual, como o uso de técnicas computacionais (Koch & Ullman, 1987; Borji & Itti, 2012; Rayner 1998) para classificação de TEA baseados na gravidade da apresentação clínica.

## **2. OBJETIVOS**

### **2.1. Objetivo principal:**

Classificar pacientes TEA quanto a gravidade usando modelos de atenção visual e rastreamento ocular.

### **2.1. Objetivos específicos:**

1. Avaliar a associação entre movimento biológico (Eye tracking) usando ROIs e a gravidade da apresentação do TEA (descrita pela CARS);
2. Utilizar técnicas de aprendizagem de máquina supervisionada para construção dos mapas de saliência;
3. Utilizar técnicas de aprendizagem de máquina supervisionada para classificação do TEA.

### **3 Material e métodos**

#### **3.1 Tipo de estudo**

Trata-se de estudo retrospectivo e transversal. O projeto foi aprovado pelo comitê de ética, CAAE, 17556019.8.0000.0068.

#### **3.2 População de estudo**

Foram incluídos 90 indivíduos em acompanhamento no ambulatório de autismo do Instituto de Psiquiatria do Hospital das Clínicas de São Paulo (PROTEA).

#### **3.3 Critérios de inclusão**

- Diagnóstico de TEA pelo DSM V
- CARS > 30
- Idade entre 3-16 anos
- Quadro estável (sem uso de medicação ou sem alterações na medicação há 3 meses)

#### **3.4 Critérios de exclusão**

- Impossibilidade de calibração adequada pelo aparelho de *Eye Tracking*
- Captação de dados pelo *Eye Tracking* menor de 80%

#### **3.5 Instrumentos:**

**3.5.1: Childhood Autism Rating Scale (CARS):** Escala que rastreia sintomas do TEA. Quinze itens compõem o instrumento de observação comportamental. Categoriza o TEA em leve, moderado e grave (Pereira et al, 2008).

A CARS é uma escala em que o responsável pela criança deve responder sobre quinze itens, cuja pontuação varia de 1 a 4, sendo 1 sem evidência de dificuldade e 4 gravemente anormal. Há a possibilidade de resposta intermediária entre os itens, por exemplo, se ficar entre o item 2 e 3, a criança pontua 2,5 pontos. Sendo assim, são 4 itens, mas 7 possibilidades de pontuação: 1, 1,5, 2, 2,5, 3, 3,5, 4.

O aplicador lê o item para o responsável e em seguida as alternativas, uma a uma, a fim de que o responsável aponte qual assemelha-se mais à criança. Mesmo que haja uma resposta positiva antes do último item, o aplicador deve ler o item seguinte, para confirmar que a afirmativa apontada é a que mais se aplica a criança.

Os itens são:

1. Relações pessoais: como a criança realiza o contato com o outro, se engaja e inicia a relação. Presença de timidez, estranhamento;
2. Imitação: se ocorre a imitação espontânea e esperada de comportamentos, de acordo com o nível de habilidade da criança;
3. Resposta emocional: avalia a mudança de expressão facial, postura e conduta frente a situações de estresse ou prazer;
4. Uso corporal: refere-se a agilidade e coordenação durante movimentos comuns à idade;
5. Uso de objetos: se a criança utiliza os objetos e brinquedos da forma esperada para habilidade e idade;

6. Resposta a mudança: como a criança reage, emocional e cognitivamente, a mudanças de rotina;
7. Resposta visual: refere ao modo da criança explorar o ambiente através da visão; exploração visual de objetos e uso da visão junto aos outros sentidos;
8. Resposta auditiva: considera tanto respostas diminuídas ou ausentes como respostas exageradas a estímulos auditivos, além do uso da audição relacionada aos outros sentidos;
9. Resposta e uso do paladar, olfato e tato: avalia o uso do paladar e olfato para exploração, se acontece, quando é adequado e na intensidade esperada ou não. Também considera a resposta dolorosa;
10. Medo ou nervosismo: se a resposta é adequada ao evento desencadeante da emoção;
11. Comunicação verbal: uso da comunicação verbal relacionado a idade e situação;
12. Comunicação não verbal: uso da comunicação não verbal relacionado a idade e situação;
13. Nível de atividade: quão ativa é a criança em diferentes ambientes;
14. Nível e consistência da resposta intelectual: avalia se a inteligência é consistente e compatível com a idade da criança;
15. Impressões gerais: impressão do aplicador sobre a criança, a partir do relato do responsável.

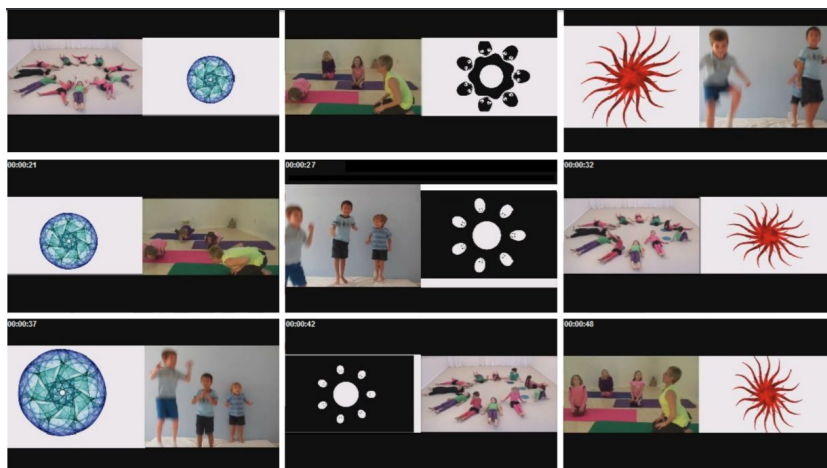
Os valores da escala CARS, determinam que o intervalo entre 30-36 corresponde a TEA leve/moderado, e acima de 36 TEA grave, abaixo de 30 são crianças que estão dentro do esperado para a idade.

**3.5.2: Rastreamento do olhar:** O aparelho utilizado para captação do rastreamento do olhar foi o *Tobii1 Pro TX300* que é capaz de capturar o olhar em até 300Hz. Neste caso, a unidade de rastreamento do olhar está acoplada ao monitor de 23". O aparelho *Tobii1 Pro TX300*, apesar de fixo e remoto, consegue capturar os movimentos visuais com boa acurácia permitindo liberdade de movimentos de cabeça a uma velocidade máxima de 50cm/s, 37x17 cm de espaço, com ângulo máximo do olhar a 35 graus. O observador deve permanecer a uma distância de 50cm a 80cm do aparelho.

**3.5.3: Paradigma de preferência visual:** Escolhemos o paradigma de atenção social, que é amplamente utilizado na literatura e mostrou-se ideal para uso em indivíduos com TEA. Este paradigma, mostrou-se, em comparação a outros dois, ser o que apresenta melhores resultados na distinção das crianças com TEA em comparação a controles(Kou et al, 2019). Além disso, tem a vantagem de ser rápido e não precisar de respostas voluntárias, possibilitando o uso em crianças verbais e não verbais, e em qualquer faixa etária.

O estímulo consiste em quadros curtos de aproximadamente 6 segundos, com dimensão de 1920 x 1080 *pixels*. A tela é apresentada dividida ao meio sendo que em metade são apresentados fractais e na outra metade cenas sociais. Na metade social da tela, aparecem crianças brincando ou fazendo ioga, enquanto na

metade geométrica, aparecem figuras abstratas, coloridas e em movimento. As imagens sociais e geométricas aparecem em lados diferentes, de forma aleatória, para evitar a acomodação visual (Figura 8). Três vídeos de movimento biológico e três vídeos de movimento geométrico foram combinados, compondo nove vídeos exibidos sequencialmente, com tempo total de 54 segundos.

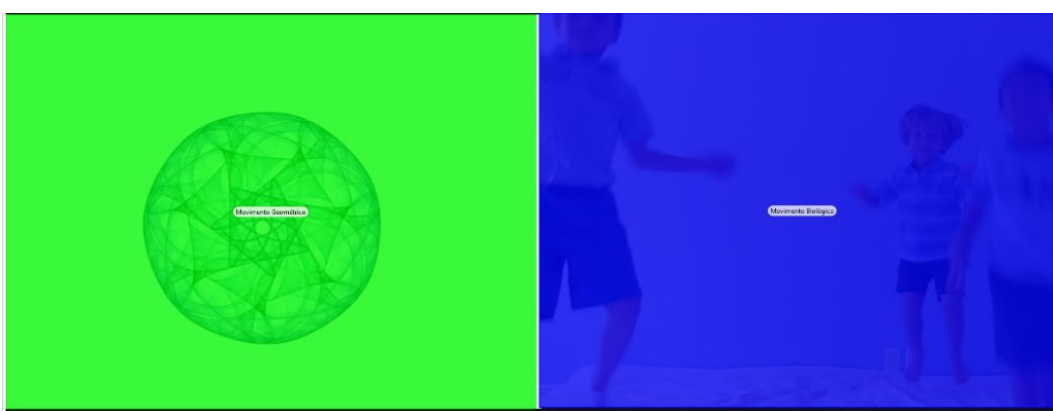


**Figura 8** - Paradigma de preferência visual. Exemplo das imagens exibidas durante o paradigma de preferência visual, de um lado da tela cenas dinâmicas e sociais e de outro fractais variando de forma aleatória.

### 3.6 Análise dos dados

Para processamento dos dados, utilizamos o software estatístico R, versão 3.5 e o Matlab (MATLAB, 2015). Os dados foram extraídos através do pacote *tobiistudium* do software do *Eye Tracking (Tobii)*. Os dados brutos indicam as coordenadas x e y das fixações em relação a cada frame, com duração de 3ms. O software também gera dados filtrados, que relacionam a captação do rastreamento do olhar às regiões de interesse (ROIs). As ROIs foram demarcadas considerando o lado da tela com movimentos biológicos ou geométricos, como mostra a Figura 9.





**Figura 9** - ROIs utilizadas. marcação das regiões de interesse, geométrica em verde e biológica em azul.

### 3.6.1 Análise ROI

Sobre os dados filtrados, o programa informa características das fixações: contagem, que registra a soma do número de fixações, e duração, que é a soma do tempo que o indivíduo fixou o olhar. Também obtivemos os dados relacionados às visitas, ou seja, o número e a duração de tempo que o indivíduo passou em determinada ROI, independente do número de fixações. As fixações são registros da estabilização da retina em um objeto de interesse, e optamos por analisar essa característica porque ela reflete o local da cena a que a atenção do indivíduo foi direcionada. Como o paradigma utilizado dividia a tela ao meio, e a informação

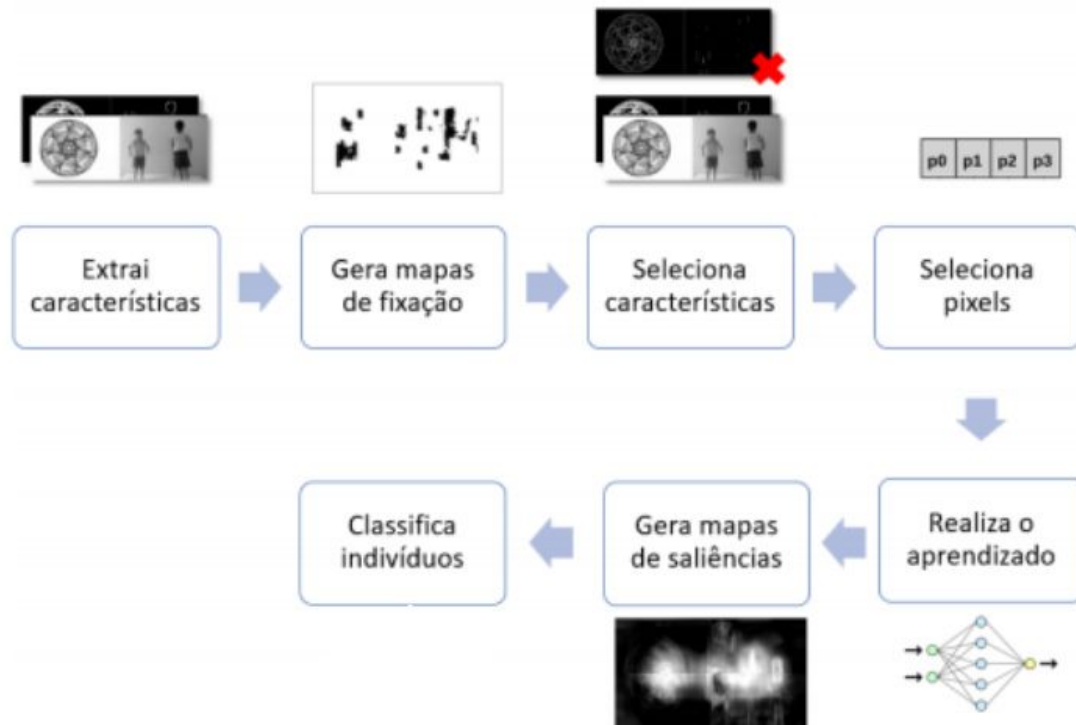
relevante é em que lado da tela o indivíduo focou atenção, as informações sobre as sacadas, que são os movimentos rápidos utilizados para reposicionar a fóvea em uma nova localização no ambiente, não foram analisadas.

### **3.6.2 Construção de mapas de saliência e análise com aprendizado de máquina supervisionada**

O projeto foi realizado na ferramenta Matlab(MATLAB, 2015) com auxílio da ferramenta mexopencv(Yamaguchi, 2013) para extração das características, em colaboração com os alunos Jessica dos Santos de Oliveira e Thiago Victor Cardoso, da professora Dra. Fátima de Lourdes dos Santos Nunes Marques da EACH.

As etapas de processamento e programas de análise de dados foram estabelecidos no projeto “Processamento do rastreamento de olhar com base em modelo de atenção visual para auxílio ao diagnóstico de autismo” de mestrado da aluna Jéssica dos Santos de Oliveira orientada pela prof. Dra. Fátima de Lourdes dos Santos Nunes Marques(Oliveira, 2018), onde está descrito de forma detalhada as etapas computacionais envolvidas neste projeto.

As etapas desse processamento estão exemplificadas no diagrama (Figura 10) e serão detalhadas a seguir :



Fonte: Oliveira, 2018

**Figura 10** - Esquema das etapas realizadas no processamento de dados. O software realiza as etapas sequencialmente, para, por fim, classificar os indivíduos em um dos grupos.

1) Análise de qualidade: os dados com menos de 80% de captação do rastreamento do olhar foram excluídos. Os selecionados para análise foram divididos em *frames*, cada um corresponde a 3ms do vídeo. As bordas foram cortadas, eliminando regiões apenas com pixels pretos. Retiramos os 10 frames do início e 10 frames do final de cada vídeo, para que o movimento dos olhos durante a transição dos vídeos não atrapalhasse a análise. Por fim, os frames, com resolução de 1920 x 720 pixels, foram analisados e redimensionados pelo método bicúbico, que reescala os pixels com média ponderada dos 4 x 4 vizinhos mais próximos, servindo de base para a etapa de seleção de características. Os frames com menos de 0,33 ms de

movimento entre eles, foram agrupados, sendo que a característica referente ao grupo é a média das características originais.

2) Extração de características dos vídeos para formação do vetor de características: Os principais modelos de aprendizado encontrados na literatura foram os guias das características inicialmente utilizadas (Judd et al, 2009; Borji 2012). São elas:

- 13 pirâmides orientadas com quatro escalas e tres orientações (Simoncelli et al, 1995);
- intensidade e orientação da imagem, presença de pele(Walter & Koch, 2006);
- 3 cores, vermelho, verde e preto, seguindo o modelo RGB (*modelo red, green, blue*);
- linha do horizonte(Judd et al, 2009);
- distância do centro às bordas da imagem;
- distância ao centro da cena biológica;
- distância ao centro da cena geométrica;
- presença de face(Viola & Jones, 2001):
- presença de pessoa(Viola & Jones, 2001);
- movimento;

Usando os modelos de características foram extraídas vinte e sete características de cada grupo de frames.

3) construção dos mapas de fixação: os mapas de fixação são a somatória das fixações dos indivíduos de cada grupo. Portanto, para cada frame foram gerados dois mapas de fixação, um para o grupo grave e outro para o grupo não grave.

Depois, foi aplicado um filtro gaussiano que realiza um espalhamento da imagem, deixando mais claras as regiões com mais fixações e escuras as com menos.

4) Seleção das características extraídas: para seleção de características, foi usado o algoritmo *ReliefF*. Que identifica os  $k$  vizinhos próximos da mesma classe e  $k$  vizinhos de classe distinta, aumentando a qualidade do estimador de atributos. O valor da média de pesos foi definido como limiar, assim, qualquer característica que ficar acima desse limiar será selecionada. O algoritmo olhará para sessenta instâncias. Esse valor foi definido de forma empírica, uma vez que aumentar o número de instâncias não gerava resultados melhores. Para cada grupo, foram selecionadas algumas características que formam o vetor de características que será a entrada para o método de treinamento.

5) Seleção dos *pixels* dos frames utilizados no treinamento: os *pixels* selecionados para treinamento foram escolhidos de duas formas: considerando quaisquer 350 *pixels* que foram olhados, chamados de positivos, e 350 que não foram fixados, negativos (Rigo, 2010; Judd et al, 2009) e considerando os 350 *pixels* mais fixados como positivos e aleatoriamente, qualquer *pixel* não olhado como negativo. Após a seleção, os *pixels* positivos são categorizados como pertencentes a classe 1 e os negativos pertencentes a classe 0. Essas classes serão preditas pelo algoritmo de classificação.

6) Método de aprendizado para geração dos mapas de saliência: utilizamos dois métodos para o aprendizado de máquina: Rede Neural Artificial e Máquina de Vetores de Suporte (*SVM - Support Vector Machine*). O método Rede Neural Artificial foi utilizado para o aprendizado de máquina, com função de ativação sigmóide na camada oculta e linear na camada de saída e treinamento com *backpropagation*

com regularização bayesiana. Havia 10 camadas ocultas e as condições de parada foram atingir 1000 épocas ou erro menor que 1 e - 7. O algoritmo de SVM utilizado foi o linear, da biblioteca liblinear(Fan et al, 2008). Por fim, para cada um dos frames dos dois grupos, foram gerados mapas de saliência, quanto mais claro o local do mapa, mais chance esse *pixel* tem de ser fixado pelo grupo específico.

7) Método de classificação: Os mapas de saliência gerados foram normalizados entre 0 e 1, binarizados com limiar de 0,75. Ou seja, apenas valores acima de 0,75 são considerados como pixels fixados. Na comparação das fixações individuais com os mapas de saliência, são considerados verdadeiros positivos os pixels que foram fixados no mapa de saliência e pelo indivíduo, e verdadeiros negativos os pixels que não foram fixados no mapa de saliência nem pelo indivíduo. Se a soma de verdadeiros positivos e negativos for maior ao comparar o mapa individual com o mapa de saliência modelo Grave do que com o mapa modelo não Grave em um determinado frame, isso é contado como um voto para a classe de Grave. Indivíduos que pontuaram acima de 20 como Grave, são classificados como tal, o mesmo processo ocorre com o grupo não Grave.

8) Avaliação dos resultados: Para treinamento e teste do classificador, foi utilizado o método de *cross-validation* com 5 *folds*. Para avaliar o desempenho de classificação, criou se uma curva *Receiver operating characteristic* (ROC), utilizando como parâmetro a quantidade de frames utilizados como limiar para classificação.

## 4-Resultados

**4.1-Characterização da Amostra:** Na tabela suplementar 1 podemos observar os dados clínicos de todos os pacientes avaliados. Foram excluídos três indivíduos por não conseguirem realizar a captação do rastreamento do olhar. Destes, dois eram do grupo não Grave e do sexo masculino e um do grupo Grave e sexo feminino. Nossa amostra final foi de oitenta e oito indivíduos, que, de acordo com o escore da CARS, foram formados dois grupos. Na tabela 2 abaixo podemos ver a distribuição de sexos, idade e CARS entre meninos e meninas em cada grupo.

**Tabela 2 - Caracterização da amostra**

	<b>Leve/Moderado</b>	<b>Grave</b>
<b>Número</b>	49	39
<b>Idade</b>	3-16 (M: 8,8 DP: 3,2)	3-16 (M: 8,7 DP: 3,9)
<b>Sexo Masc/Fem</b>	27/22	12/27
<b>CARS</b>	30-36 (M: 33,1 , DP: 2)	36,5 - 46,5 (M: 39,4 ,DP: 2,6)

Caracterização da amostra em relação a cada grupo. Idade em anos, CARS = Childhood Autism Rating Scale M.= média, DP = Desvio Padrão.

Os grupos mostraram-se homogêneos em relação a idade, mas diferiam quanto ao sexo, sendo que o grupo grave tem mais meninos do que meninas. A literatura mostra que o autismo é mais prevalente no sexo masculino, mas que a diferença entre sexos diminui quanto mais grave o quadro(Lord et al, 1982). Levando-se em conta que meninas podem apresentar um quadro mais grave que

meninos foi feito um Teste-t de duas amostras presumindo variâncias equivalentes (p-valor= 0,01, alpha de 5%) para comparação da CARS entre sexos. Não foi observada diferença de idade entre sexos (p-valor =0,11, alpha de 5%). Também não foi observada associação entre a CARS e idade independente do sexo (p-valor=0,68, alpha de 5%).

#### **4.2- Gravidade X Resultados do paradigma social processado como feito na literatura, ou seja, usando ROIs.**

O software do Tobbi permite a extração dos dados filtrados, ou seja, referentes às ROIs previamente marcadas, neste caso, região geométrica e região biológica. As informações são referentes às características das fixações: contagem, que registra a soma do número das fixações, e duração, que é a soma do tempo que o indivíduo fixou o olhar. Também obtivemos os dados relacionados às visitas, ou seja, o número e a duração de tempo que o indivíduo passou em determinada ROI, independente do número de fixações. A tabela suplementar 2 apresenta todos os dados filtrados dos pacientes.

A fim de normalizar os dados, fizemos um escore: o valor da medida referente a região geométrica, dividido pela soma dos valores correspondentes nas duas regiões (geométrica + biológica). A partir da matriz de correlações entre os escores ficou claro que os escores de Duração da Visita e Duração das Fixações estão correlacionados, o que é esperado, já que, apesar de serem duas características diferentes, elas avaliam o mesmo elemento: quanto tempo o olhar permaneceu na ROI de interesse. Os escores de Contagem das Visitas e das fixações não são correlacionados, como mostra a tabela 3.



**Tabela 3 - Correlação entre os escores de ROIs**

Escores	DF	CF	DV	CV
DF	1			
CF	0,73	1		
DV	0,99	0,74	1	
CV	0,50	0,58	0,52	1

Escores dos dados de rastreamento do olhar. DF: duração das fixações, CF: contagem das fixações, DV: duração das visitas, CV: contagem das visitas.

Fizemos análise de regressão para observar a associação de cada um dos escores e a CARS, usando como covariáveis idade e sexo. O único modelo significativo foi com o escore da contagem das fixações, no entanto, a explicação do modelo foi baixa, de acordo com a tabela 4.

**Tabela 4 - Regressão da associação dos escores e a CARS**

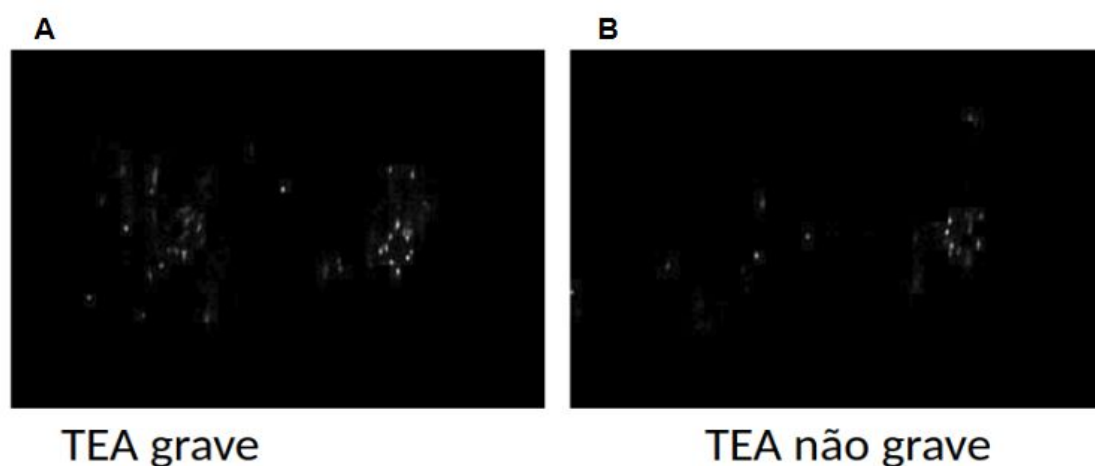
	Coeficiente	Erro padrão	R <sup>2</sup> ajustado	p-valor	Intervalo de confiança (95%)	
Score DF	3,871	2,441	0,055	0,05	-,98	8,72
Score CF	6,954	2,688	0,099	0,008	1,6	12,30
Score VD	4,01	2,38	0,059	0,043	-,73	8,76
Score VC	10,5	8,62	0,044	0,078	-6,67	27,64

Resultado da regressão utilizando os escores do rastreamento do olhar e a CARS, covariando por sexo e idade. DF = Duração das fixações, CF = contagem das fixações, VD = Duração das visitas, VC= Contagem das visitas

### 4.3-Resultados do classificador feito a partir dos mapas de saliência

Após análise de qualidade realizada obtivemos 1590 frames, que foram redimensionados pelo método bicúbico. *Frames* semelhantes foram agrupados gerando 198 grupos de frames. Para cada *frame* foram gerados dois mapas de fixação, um para o grupo Grave e outro para o grupo não Grave e aplicado filtro gaussiano.

A figura 11 mostra um exemplo de mapa de fixação de cada um dos grupos.



**Figura 11** - Mapa de fixação de cada grupo. Em A o mapa de fixação do grupo grave e em B para o grupo não grave, considerando o mesmo frame.

Para cada grupo, foram selecionadas algumas características para formar o vetor de características usado na entrada para o método de treinamento. As doze características do grupo grave foram: linhas horizontais, centro da tela e centro das cenas, cinco de pirâmide, as três de cores *RGB*, e característica de intensidade de cor. Para o grupo não grave, foram dezesseis características, linhas horizontais,

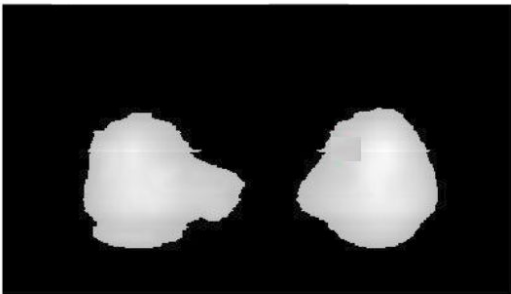
centro da tela e centro das cenas, seis de pirâmide, as três de cores *RGB*, característica de intensidade de cor, pele e pessoas. Isso indica que o grupo não Grave, fixa mais o olhar em *frames* com pessoas e com cor de pele.

Os *pixels* foram selecionados para treinamento e depois os *pixels* positivos foram categorizados como pertencentes a classe 1 e os negativos pertencentes a classe 0. Essas classes serão as previstas pelo algoritmo de classificação.

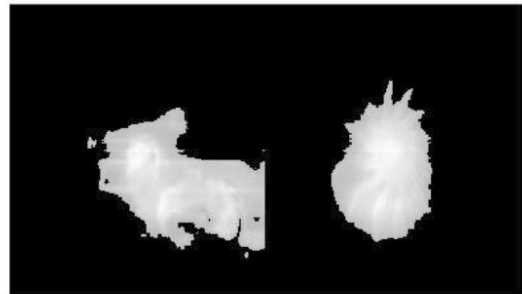
O método Rede Neural Artificial foi utilizado para o aprendizado de máquina. Por fim, para cada um dos *frames* dos dois grupos, foram gerados mapas de saliência, quanto mais claro o local do mapa, mais chance esse *pixel* tem de ser fixado pelo grupo específico.

A figura 12 ilustra o mapa de saliência de cada grupo.

Mapa de Saliência grupo Grave



Mapa de Saliência grupo não-Grave

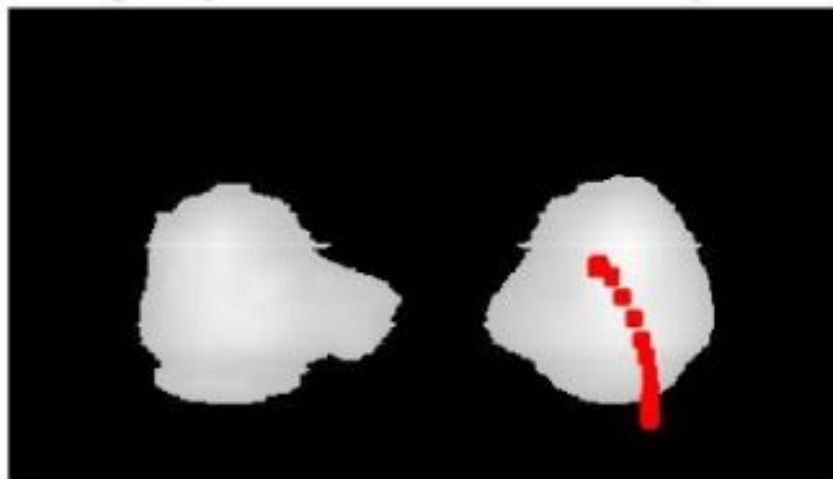


**Figura 12** - Mapas de saliência. A imagem da direita é o mapa de saliência para o grupo não grave e a da esquerda para o grupo grave para o mesmo frame.

Os mapas de saliência gerados foram normalizados entre 0 e 1, binarizados com limiar de 0,75. Ou seja, apenas valores acima de 0,75 são considerados como *pixels* fixados. Na comparação das fixações individuais com os mapas de saliência, foram considerados verdadeiros positivos os *pixels* que foram fixados no mapa de

saliência e pelo indivíduo, e verdadeiros negativos os *pixels* que não foram fixados no mapa de saliência nem pelo indivíduo. Se a soma de verdadeiros positivos e negativos for maior ao comparar o mapa individual com o mapa de saliência modelo Grave do que com o mapa modelo não Grave em um determinado frame, isso é contado como um voto para a classe de Grave. Indivíduos que pontuaram acima de 20 como Grave, são classificados como tal, o mesmo processo ocorre com o grupo não Grave.

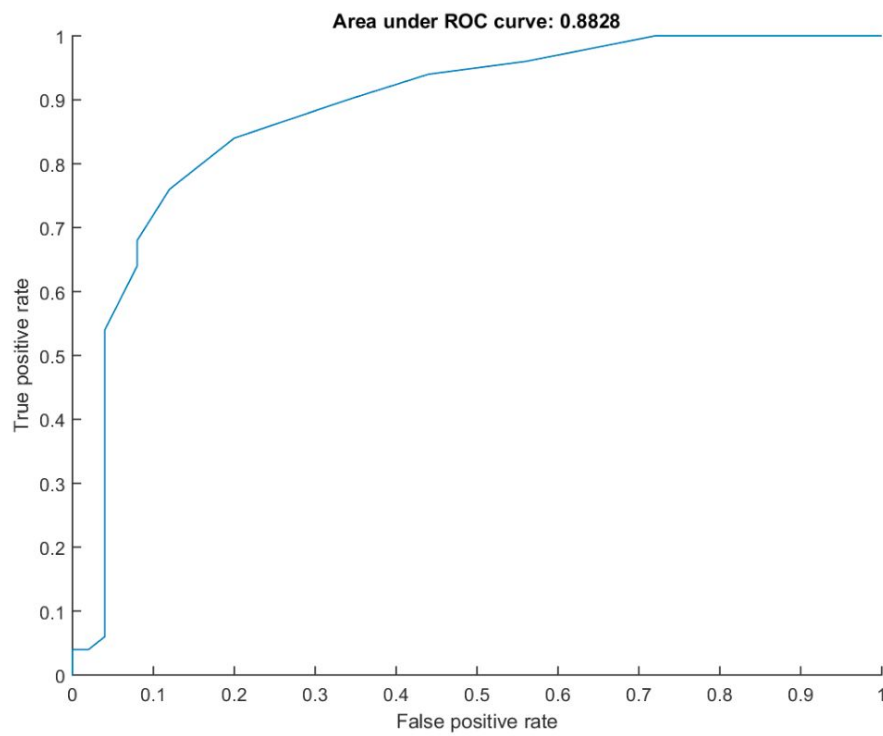
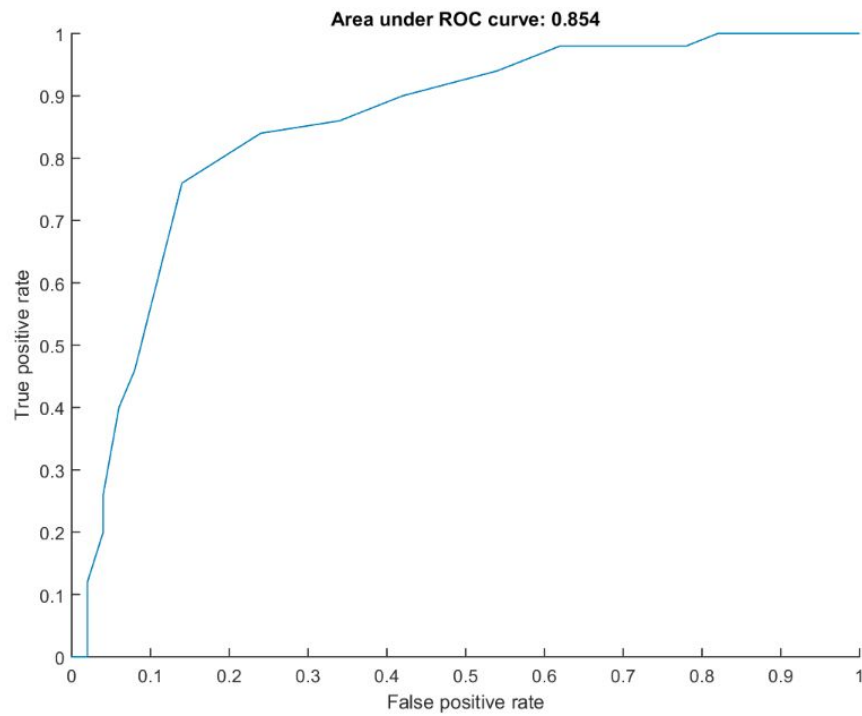
Abaixo, a figura 13 é um exemplo de fixações de um indivíduo do grupo grave sobreposto ao mapa de fixação de seu grupo.



**Figura 13** - Mapa de saliência com fixações sobrepostas. Em vermelho estão as fixações, sobrepostas ao mapa de saliência em cinza.

Para treinamento e teste do classificador, foi utilizado o método de *cross-validation* com 5 *folds*. O conjunto de treinamento selecionado continha 135 *frames* selecionados aleatoriamente. Para avaliar o desempenho de classificação, criou-se uma curva ROC, utilizando como parâmetro a quantidade de *frames* utilizados como limiar para classificação. Utilizando os dados brutos do rastreamento do olhar através da análise dos mapas de fixação foi possível

classificar os grupos Grave e não Grave. Na curva ROC, *area under curve* (AUC) média foi de 0,85, e a máxima 0,88 como mostra a figura 14.



**Figura 14** - Curvas ROC com os dados do classificador. A segunda curva é a de melhor resultado.

Através da curva AUC, foi possível calcular o limiar mínimo para classificação, que foi de 20. O limiar foi definido a partir do cálculo do ponto de valor máximo entre  $\min(\text{RTP}, 1-\text{RFP})$ , em que RTP é o *True Positive Rate*, proporção de positivos corretamente classificados, e o RFP é o *False Positive Rate*, proporção de positivos que não estão corretamente classificados.

Isso significa que se o indivíduo tiver 20 *frames*, dos 40 utilizados na rodada, semelhantes ao grupo Grave, esse indivíduo será classificado como Grave. Neste limiar, o valor de precisão obtido foi de, com sensibilidade de 87%, especificidade 60% e erro de 25%.

## 5-Discussão

O uso da técnica de rastreamento do olhar tem se mostrado promissor para o estudo em TEA, uma vez que as alterações do rastreio do ambiente estão presentes precocemente e são um marco do transtorno(Chawarska et al, 2013). Inúmeros trabalhos mostraram efetividade para separar os indivíduos com TEA da população controle(Chawarska et al, 2013; Dawson et al, 1998; Wang et al, 2015). Estudos mais recentes, associaram características do rastreamento do olhar com sintomas específicos, principalmente relacionados a comunicação(Moore et al, 2018; Pierce et al, 2016; Del Valle Rubido, 2018).

A amostra coletada tem intervalo de idade grande, de 03 a 16 anos, por isso, optamos por utilizar o paradigma de preferência visual por estímulos biológicos e geométricos. A preferência pelo estímulo biológico é primordial para a percepção adequada do ambiente, e está presente desde os primeiros meses de vida da criança, sendo mantida ao longo da vida(Di Giorgio et al, 2012). Além disso, no autismo, o déficit na percepção do movimento biológico, é um dos primeiros sinais do transtorno, aparecendo a partir dos seis meses de vida e mantendo-se ao longo do tempo(Chawarska et al, 2013). Os grupos são diferentes em relação ao sexo, sendo que o grupo leve/moderado tem maior número de meninas que o grupo Grave. A literatura mostra que a prevalência do TEA é maior no sexo masculino que no feminino, mas que essa diferença diminui quanto maior a gravidade do transtorno. Para confirmar que a diferença de idade e o sexo não iriam impactar no resultado do nosso trabalho, aplicamos o teste t, que não mostrou associação entre a idade e o resultado da CARS, independente do sexo.

O presente estudo foi desenvolvido a fim de investigar a relação entre a gravidade do TEA e os padrões do rastreamento do olhar. Em continuidade aos trabalhos sendo desenvolvidos pelo grupo de pesquisa, em colaboração com o grupo da EACH, optamos por utilizar técnicas de aprendizado de máquina no processamento dos dados a fim de classificar a população entre Grave e não Grave. Além da diferença do direcionamento do olhar para regiões geométrica ou biológica, utilizamos o modelo de mapas de saliência para identificar características relevantes para os dois grupos, possibilitando análise sobre o processamento da imagem pelos indivíduos com TEA e a diferença entre os diferentes níveis do transtorno.

A princípio, utilizamos as ROIs para relacionar os dados de rastreamento do olhar e a classificação da gravidade do TEA. O escore relativo a contagem das fixações na região geométrica foi o único que mostrou alguma relação com a CARS, porém, explicando pouco o modelo. O resultado da regressão, considerando sexo e idade, mostrou que a contagem das fixações explica 10% do modelo, sendo a única variável das avaliadas que teve resultado satisfatório. As outras, duração das fixações, contagem e duração das visitas, não mostrou relação com os valores da CARS.

Sabemos que os indivíduos com TEA fixam menos vezes na região biológica do que na geométrica, apresentando um padrão disperso do olhar. Como as ROIs utilizam apenas os dados referentes às regiões pré definidas, esse padrão disperso e inconstante pode tornar difícil a classificação em subtipos. Além disso, ROIs não levam em conta o modelo biológico de atenção visual, ou seja, não considera quais



as características importantes das imagens que produzem modelos de saliência visual. Por isso, optamos por utilizar técnicas de construção de mapas de saliência visual e reconhecimento de padrões para classificação. Como esperado a análise dos mapas de fixação permitiu que os dados referentes a toda tela fossem analisados, gerando melhores resultados.

Na fase de seleção de características, o grupo não Grave teve duas características, presença de cor de pele e face, que não apareceram no grupo Grave. Indivíduos com TEA têm dificuldade em reconhecer pessoas e rosto em cenas complexas, e fixam menos o olhar nessas regiões. Porém, não há dados objetivos sobre essa diferença em subgrupos de TEA, que relacionem esses dados com a gravidade.

Segundo Wang et al 2015, indivíduos com TEA utilizam menos o processamento *top-down*, que processa informações relacionadas a aprendizado e experiências pregressas. A preferência por características como cor de pele e face dos indivíduos do grupo não grave aponta para o processamento *top-down*, enquanto o grupo grave permanece com a atenção em características mais básicas da imagem, como cor, movimento, profundidade, processadas pelo sistema *bottom-up*.

O trabalho de Seymour et al, 2019 corrobora essa hipótese, mostrando que durante o processamento visual indivíduos com TEA têm menor feedback top-down, ativando regiões diferentes do sistema nervoso central em comparação com controles.

Apesar do resultado apresentado, nosso trabalho possui algumas limitações. Um número maior de pacientes poderia melhorar o resultado do classificador. Outro ponto a ser considerado é a ausência de dados sobre o quociente de inteligência e outras características de comportamento, como agressividade, impulsividade, agitação, entre outros. A análise desse conjunto de comportamentos, possibilitaria a identificação de perfis fenotípicos, difíceis de serem observados na clínica. Analisar a trajetória do olhar, ao invés das fixações, também é uma melhoria possível, uma vez que dados de trajetória podem trazer informações importantes sobre a dinâmica do processamento visual na população com TEA. Como as trajetórias são mais variadas que as fixações, é necessário um grupo maior de pacientes para que a análise seja realizada.

O uso do rastreamento do olhar para classificar subgrupos de TEA pode ser uma ferramenta importante. O refinamento da técnica pode auxiliar na tomada de decisão desde o diagnóstico até a definição do tratamento. Novos estudos, utilizando outros dados fenotípicos, como presença de comorbidades, perfil de comportamento, bem como diferentes paradigmas são necessários para que o uso da técnica de rastreamento do olhar seja possível na prática clínica, orientando a escolha do tratamento além de auxiliar o diagnóstico.

## **6-Conclusão:**

Este é o primeiro trabalho que mostra que o uso dos dados de rastreamento do olhar baseadas em modelos de atenção visual com o auxílio de técnicas computacionais de reconhecimento de padrões contribui para identificação de subgrupos de gravidade no TEA.

O reconhecimento de pessoas e rosto , fundamental para identificação sobre o estado emocional, intenção, imitação, são características importantes para diferenciar grupos, e pode sugerir que mecanismos *top-down* se encontram mais prejudicados nos pacientes mais graves. Além dessa diferença, nosso classificador foi capaz de identificar e separar a população em grupo Grave e não Grave com acurácia de 88%.

Os dados de eye tracking são eficazes para identificar sub grupos entre indivíduos com TEA, o que é importante para auxílio diagnóstico, orientação de prognóstico e tratamento, bem como para permitir grupos mais homogêneos em pesquisas futuras.

Além disso, auxiliam no entendimento da fisiopatologia do transtorno, corroborando com pesquisas que utilizaram outras técnicas, como ressonância magnética, eletroencefalograma.



## 7 - ANEXOS

Tabela suplementar 1 - Caracterização Clínica da Amostra

id	idade	CARS	sexo
1	10	33	M
2	13	32,5	M
3	7	38	F
4	15	33,5	M
5	15	39,5	M
6	4	31	M
7	3	42	F
8	3	36,5	M
9	15	40	M
10	8	34	F
11	9	32	F
12	4	35	F
13	12	39	F
14	4	36	M
15	6	34,5	M
16	4	30	F
17	12	31,5	M
18	11	31	F
19	13	35,5	M
20	10	30,5	F
21	11	30	M
22	13	34	M
23	6	32	F
24	11	30,5	M
25	7	35	M
26	7	36	F
27	10	30	M
28	14	37,5	F
29	9	30,5	F

30	3	36	M
31	8	30,5	M
32	4	32,5	F
33	11	35	M
34	10	34	M
35	6	38	F
36	10	33	M
37	9	36,5	M
38	10	37	M
39	10	34,5	M
40	7	45	F
41	11	41	F
42	11	31	M
43	5	37	M
44	12	42	M
45	7	30,5	F
46	4	37	M
47	8	35,5	M
48	11	32	M
49	9	33	M
50	9	31	M
51	8	33,5	F
52	12	32,5	F
53	11	40,5	M
54	11	37	M
55	10	44,5	M
56	8	30,5	M
57	7	33	M
58	10	38,5	F
59	3	32,5	F
60	11	42	F
61	5	39	F
62	15	36	F
63	4	38	F

64	5	35,5	F
65	13	38	F
66	10	46,5	F
67	8	34	F
68	9	36	F
69	6	37	F
70	7	35,5	F
71	4	36,5	F
72	4	37,5	F
73	5	39	F
74	3	43	F
75	11	39,5	M
76	12	41	F
77	15	37	F
78	5	39,5	F
79	11	41,5	F
80	4	43,5	F
81	6	36	F
82	8	35,5	M
83	11	38,5	F
84	16	37	F
85	10	38	F
86	5	39	F
87	16	32	F
88	12	33	F

Planilha com os dados clínicos dos pacientes, id = identificação, CARS = Childhood Autism Rating Scale, Sexo. Valores mínimo, máximo, da média e desvio padrão de cada característica; M: Masculino e F: Feminino.

Tabela suplementar 2

Tabela dos dados de Rastreamento do Olhar

id	Duração das Fixações Biológico	Duração das Fixações Geométrico	Contagem das Fixações Biológico	Contagem das Fixações Geométrico	Duração das Visitas Biológico	Duração das Visitas Geométrico	Contagem das Visitas Biológico	Contagem das Visitas Geométrico
1	12,33	33,88	74	114	14,97	36,35	22	19
2	28,69	18,1	81	60	30,7	19,85	27	24
3	17,79	33,71	55	78	18,09	34,52	25	26
4	18,52	32,56	38	36	18,81	33,26	14	14
5	19,99	18,32	71	87	19,99	18,51	14	17
6	25,61	20,9	72	33	26,25	20,9	18	19
7	13,22	34,14	23	33	13,93	35,36	10	11
8	10,79	18,01	88	80	10,92	18,01	23	13
9	9,4	43,16	23	77	9,4	43,33	14	18
10	30,36	19,1	104	49	30,78	19,1	22	18
11	29,16	18,74	86	45	29,16	18,74	24	23
12	25,88	27,65	65	48	25,88	27,65	18	19
13	25,22	15,62	74	28	30,12	15,77	11	12
14	12,26	36,31	39	72	12,5	36,7	19	22
15	19,9	27,68	49	34	20,35	27,99	16	15
16	20,99	32,07	32	23	20,99	32,07	11	11
17	23,14	25,14	87	85	24,84	26,05	36	35



18	30,77	21,71	65	39	31,3	22,08	21	20
19	32,34	19,96	82	36	33,21	20,1	23	21
20	25,64	22,56	223	192	25,77	23,19	49	43
21	24,55	16,68	92	68	24,74	16,78	27	28
22	7,54	40,73	24	81	7,54	42,14	15	16
23	12,43	38,18	40	73	12,43	39,85	17	17
24	23,74	29,63	51	23	23,74	29,63	11	12
25	41,14	11,75	125	22	41,14	11,75	19	16
26	38,55	13,68	79	21	38,79	13,68	16	14
27	26,04	26,84	72	60	26,04	26,84	22	21
28	23,2	5,54	93	24	24,62	6	13	13
29	12,77	21,6	203	402	16,99	24,05	21	36
30	14,91	34,13	47	70	15,59	35,3	17	18
31	28,44	24,55	93	66	28,44	24,55	26	25
32	13,78	35,3	45	41	13,78	35,3	18	18
33	37,45	12,72	71	25	37,45	12,72	16	16
34	28,37	23,04	52	43	28,55	23,21	24	28
35	7,14	22,88	48	104	7,14	27,24	13	25
36	31,82	20,6	55	18	32,03	20,6	15	13
37	71	1	1	14	15	1	1	15
38	16,26	27,2	40	47	16,26	27,34	18	23
39	4,3	42,01	17	53	4,91	43,81	5	9

40	6,54	38,5	29	160	6,54	40,77	13	16
41	15,79	13,19	159	125	16,48	13,69	60	54
42	19,29	32,34	45	49	19,42	32,74	18	18
43	17,19	25,65	63	82	17,29	27,7	25	28
44	28,91	20,7	41	40	29,09	21,07	12	12
45	34,33	19,26	79	40	34,47	19,26	25	25
46	17,9	30,93	82	55	18,39	31,71	17	18
47	17,47	23,72	274	279	17,68	24,05	138	125
48	12,92	35,8	45	58	13,04	37,92	16	17
49	22,07	29,07	54	53	22,38	29,82	25	24
50	17,22	31,85	53	62	17,22	31,97	21	21
51	26,68	26,17	24	14	26,68	26,17	13	11
52	28,43	15,42	96	61	29,03	17,36	27	23
53	19,55	24,01	52	38	19,88	24,56	11	13
54	17,99	31,06	57	25	18,32	31,37	13	13
55	6,58	12,57	28	44	7,7	26,13	9	10
56	22,63	30,24	67	64	22,63	30,24	26	23
57	21,43	30,36	59	39	21,43	30,36	18	18
58	16,62	32,38	60	119	17,44	34,14	19	19
59	12,07	35,35	34	38	12,07	37,26	16	18
60	24,25	24,81	43	27	24,42	24,81	19	19
61	14,39	37,94	40	57	14,52	38,53	21	23

62	22,53	30,15	37	33	22,82	30,15	11	11
63	25,21	22,53	72	46	25,21	22,53	28	25
64	36,58	16,75	66	28	36,58	16,85	15	14
65	26,16	26,62	81	31	26,16	26,74	19	17
66	16,02	30,2	22	32	16,57	35,03	11	12
67	22,99	28,04	50	38	23,19	28,04	20	18
68	27,85	21,38	128	45	28,33	21,77	30	23
69	26,05	9,66	312	115	35,27	9,66	58	29
70	13,17	20,61	182	185	13,29	20,94	66	65
71	29,41	21,11	79	55	29,65	21,32	25	26
72	9,54	37,25	27	41	10,62	37,5	12	16
73	19,65	20,45	45	44	19,65	20,55	22	22
74	10,24	15,94	31	42	10,38	16,15	22	25
75	10,9	15,19	46	61	11,01	15,19	16	17
76	22,71	30,75	56	39	22,71	30,75	15	14
77	11,59	29,25	37	60	12,23	31,03	13	18
78	7,71	34,5	28	66	8,66	36,81	8	8
79	10,82	27,2	91	208	10,82	31,48	25	33
80	10,66	10,55	138	119	22,86	12,7	43	40
81	21,67	16,67	83	52	25,99	18,09	16	16
82	27,37	25,42	23	23	27,37	25,42	11	11
83	24,98	9,29	63	28	27,21	9,53	15	9

84	24,42	28,79	86	35	24,42	28,79	12	11
85	16,26	16,38	55	47	16,26	16,38	14	12
86	31,11	21,66	64	31	31,11	21,66	14	12
87	38,98	12,91	33	13	39,26	13,08	8	8
88	8,87	20,86	31	66	8,87	21,28	20	23

Tabela dos dados do rastreamento do olhar em milissegundos.

## Bibliografia:

1. Abrams, D. A., Lynch, C. J., Cheng, K. M., Phillips, J., Supekar, K., Ryali, S., ... & Menon, V. (2013). Underconnectivity between voice-selective cortex and reward circuitry in children with autism. *Proceedings of the National Academy of Sciences*, 110(29), 12060-12065.
2. American Psychiatric Association. (2013). *Diagnostic and statistical manual of mental disorders (DSM-5®)*. American Psychiatric Pub.
3. Baron-Cohen, S. (1997). How to build a baby that can read minds: Cognitive mechanisms in mindreading. *The maladapted mind: Classic readings in evolutionary psychopathology*, 207-239.
4. Bayliss, A. P., & Tipper, S. P. (2006). Predictive gaze cues and personality judgments: Should eye trust you?. *Psychological Science*, 17(6), 514-520.
5. Bayliss, A. P., Frischen, A., Fenske, M. J., & Tipper, S. P. (2007). Affective evaluations of objects are influenced by observed gaze direction and emotional expression. *Cognition*, 104(3), 644-653.
6. Borji, A. (2012, June). Boosting bottom-up and top-down visual features for saliency estimation. In *2012 IEEE Conference on Computer Vision and Pattern Recognition* (pp. 438-445). IEEE.
7. Borji, A., & Itti, L. (2012). State-of-the-art in visual attention modeling. *IEEE transactions on pattern analysis and machine intelligence*, 35(1), 185-207.
8. Brentani, H., Paula, C. S. D., Bordini, D., Rolim, D., Sato, F., Portolese, J., ... & McCracken, J. T. (2013). Autism spectrum disorders: an overview on diagnosis and treatment. *Brazilian Journal of Psychiatry*, 35, S62-S72.
9. Casanova, M. F., & Casanova, E. L. (2019). The modular organization of the cerebral cortex: Evolutionary significance and possible links to neurodevelopmental conditions. *Journal of Comparative Neurology*, 527(10), 1720-1730.
10. Casanova, M. F., van Kooten, I. A., Switala, A. E., van Engeland, H., Heinsen, H., Steinbusch, H. W., ... & Schmitz, C. (2006). Minicolumnar abnormalities in autism. *Acta neuropathologica*, 112(3), 287.
11. Chauhan, V., & Chauhan, A. (2015). Contribution of oxidative stress to the pathophysiology of autism spectrum disorders: impact of genetic and environmental factors. In *Studies on Psychiatric Disorders* (pp. 89-120). Humana Press, New York, NY.
12. Chawarska, K., Macari, S., & Shic, F. (2013). Decreased spontaneous attention to social scenes in 6-month-old infants later diagnosed with autism spectrum disorders. *Biological psychiatry*, 74(3), 195-203.
13. Christensen, D. L., Braun, K. V. N., Baio, J., Bilder, D., Charles, J., Constantino, J. N., ... & Lee, L. C. (2018). Prevalence and characteristics of autism spectrum disorder among children aged 8 years—autism and developmental disabilities monitoring network, 11 sites, United States, 2012. *MMWR Surveillance Summaries*, 65(13), 1.
14. Constantino, J. N., Kennon-McGill, S., Weichselbaum, C., Marrus, N., Haider, A., Glowinski, A. L., ... & Jones, W. (2017). Infant viewing of social scenes is under genetic control and is atypical in autism. *Nature*, 547(7663), 340.
15. Dawson, G., Meltzoff, A. N., Osterling, J., Rinaldi, J., & Brown, E. (1998). Children with autism fail to orient to naturally occurring social stimuli. *Journal of autism and developmental disorders*, 28(6), 479-485.
16. Dawson, G., Toth, K., Abbott, R., Osterling, J., Munson, J., Estes, A., & Liaw, J. (2004). Early social attention impairments in autism: social orienting, joint attention, and attention to distress. *Developmental psychology*, 40(2), 271.
17. Del Valle Rubido, M., McCracken, J. T., Hollander, E., Shic, F., Noeideke, J., Boak, L., ... & Umbricht, D. (2018). In Search of Biomarkers for Autism Spectrum Disorder. *Autism Research*, 11(11), 1567-1579.

18. Di Giorgio, E., Turati, C., Altoè, G., & Simion, F. (2012). Face detection in complex visual displays: An eye-tracking study with 3-and 6-month-old infants and adults. *Journal of experimental child psychology*, 113(1), 66-77.
19. Duchowski, A. T. (2002). A breadth-first survey of eye-tracking applications. *Behavior Research Methods, Instruments, & Computers*, 34(4), 455-470.
20. Duchowski, A. T. (2007). Eye tracking methodology. *Theory and practice*, 328(614), 2-3.
21. Ecker, C., Bookheimer, S. Y., & Murphy, D. G. (2015). Neuroimaging in autism spectrum disorder: brain structure and function across the lifespan. *The Lancet Neurology*, 14(11), 1121-1134.
22. Elsabbagh M, Johnson MH. Autism and the social brain: the first-year puzzle. *Biological psychiatry*. 2016 Jul 15;80(2):94-9.
23. Emery, N. J. (2000). The eyes have it: the neuroethology, function and evolution of social gaze. *Neuroscience & Biobehavioral Reviews*, 24(6), 581-604.
24. Fan, R. E., Chang, K. W., Hsieh, C. J., Wang, X. R., & Lin, C. J. (2008). LIBLINEAR: A library for large linear classification. *Journal of machine learning research*, 9(Aug), 1871-1874.
25. Felleman, D. J., & Van, D. E. (1991). Distributed hierarchical processing in the primate cerebral cortex. *Cerebral cortex (New York, NY: 1991)*, 1(1), 1-47.
26. Frazier, T. W., Klingemier, E. W., Parikh, S., Speer, L., Strauss, M. S., Eng, C., ... & Youngstrom, E. A. (2018). Development and Validation of Objective and Quantitative Eye Tracking– Based Measures of Autism Risk and Symptom Levels. *Journal of the American Academy of Child & Adolescent Psychiatry*, 57(11), 858-866.
27. Freeth, M., Chapman, P., Ropar, D., & Mitchell, P. (2010). Do gaze cues in complex scenes capture and direct the attention of high functioning adolescents with ASD? Evidence from eye-tracking. *Journal of autism and developmental disorders*, 40(5), 534-547.
28. Geschwind, D. H., & Levitt, P. (2007). Autism spectrum disorders: developmental disconnection syndromes. *Current opinion in neurobiology*, 17(1), 103-111.
29. Ghazanfar, A. A., & Santos, L. R. (2004). Primate brains in the wild: the sensory bases for social interactions. *Nature Reviews Neuroscience*, 5(8), 603.
30. Gredebäck, G., Johnson, S., & von Hofsten, C. (2009). Eye tracking in infancy research. *Developmental neuropsychology*, 35(1), 1-19.
31. Haith, M. M., Bergman, T., & Moore, M. J. (1977). Eye contact and face scanning in early infancy. *Science*, 198(4319), 853-855.
32. Holmqvist, K., Nyström, M., Andersson, R., Dewhurst, R., Jarodzka, H., & Van de Weijer, J. (2011). *Eye tracking: A comprehensive guide to methods and measures*. OUP Oxford.
33. Howes, O. D., Rogdaki, M., Findon, J. L., Wichers, R. H., Charman, T., King, B. H., ... & Povey, C. (2018). Autism spectrum disorder: Consensus guidelines on assessment, treatment and research from the British Association for Psychopharmacology. *Journal of Psychopharmacology*, 32(1), 3-29.
34. Huber, E. (1931). Evolution of facial musculature and facial expression.
35. Insel, T. R., & Fernald, R. D. (2004). How the brain processes social information: searching for the social brain. *Annu. Rev. Neurosci.*, 27, 697-722.
36. Itti, L. (2005). Models of bottom-up attention and saliency. In *Neurobiology of attention* (pp. 576-582). Academic Press.
37. Itti, L., Koch, C., & Niebur, E. (1998). A model of saliency-based visual attention for rapid scene analysis. *IEEE Transactions on Pattern Analysis & Machine Intelligence*, (11), 1254-1259.
38. Judd, T., Ehinger, K., Durand, F., & Torralba, A. (2009, September). Learning to predict where humans look. In *2009 IEEE 12th international conference on computer vision* (pp. 2106-2113). IEEE.
39. Kaas, J. H. (1989). The evolution of complex sensory systems in mammals. *Journal of Experimental Biology*, 146(1), 165-176.

40. Kandel, E., Schwartz, J., Jessell, T., Siegelbaum, S., & Hudspeth, A. J. (2014). *Princípios de Neurociências-5*. AMGH Editora.
41. Kanner, L. (1943). Autistic disturbances of affective contact. *Nervous child*, 2(3), 217-250.
42. Kienzle, W., Wichmann, F. A., Franz, M. O., & Schölkopf, B. (2007). A nonparametric approach to bottom-up visual saliency. In *Advances in neural information processing systems* (pp. 689-696).
43. Klein, J. T., Shepherd, S. V., & Platt, M. L. (2009). Social attention and the brain. *Current Biology*, 19(20), R958-R962.
44. Kliemann, D., Dziobek, I., Hatri, A., Steimke, R., & Heekeren, H. R. (2010). Atypical reflexive gaze patterns on emotional faces in autism spectrum disorders. *Journal of Neuroscience*, 30(37), 12281-12287.
45. Kobayashi, H., & Kohshima, S. (2001). Unique morphology of the human eye and its adaptive meaning: comparative studies on external morphology of the primate eye. *Journal of human evolution*, 40(5), 419-435.
46. Koch, C., & Ullman, S. (1987). Shifts in selective visual attention: towards the underlying neural circuitry. In *Matters of intelligence* (pp. 115-141). Springer, Dordrecht.
47. Koch, K., McLean, J., Segev, R., Freed, M. A., Berry II, M. J., Balasubramanian, V., & Sterling, P. (2006). How much the eye tells the brain. *Current Biology*, 16(14), 1428-1434.
48. Kou, J., Le, J., Fu, M., Lan, C., Chen, Z., Li, Q., ... & Kendrick, K. (2019). Comparison of three different eye-tracking tasks for distinguishing autistic from typically developing children and autistic symptom severity. *BioRxiv*, 547505.
49. Lord, C., & Bishop, S. L. (2015). Recent advances in autism research as reflected in DSM-5 criteria for autism spectrum disorder. *Annual review of clinical psychology*, 11, 53-70.
50. Lord, C., Schopler, E., & Revicki, D. (1982). Sex differences in autism. *Journal of Autism and developmental disorders*, 12(4), 317-330.
51. Lorenz, K. (1970). Companions as factors in the bird's environment. *Studies in animal and human behavior*, 1, 101-258.
52. Loth, E., Murphy, D. G., & Spooren, W. (2016). Defining precision medicine approaches to autism spectrum disorders: concepts and challenges. *Frontiers in psychiatry*, 7, 188.
53. MATLAB, V. (2015). 8.5. 0.197613 (R2015a). The MathWorks Inc., Natick, Massachusetts.
54. Masi, A., DeMayo, M. M., Glozier, N., & Guastella, A. J. (2017). An overview of autism spectrum disorder, heterogeneity and treatment options. *Neuroscience bulletin*, 33(2), 183-193.
55. Mason, M. F., Tatlow, E. P., & Macrae, C. N. (2005). The look of love: Gaze shifts and person perception. *Psychological science*, 16(3), 236-239.
56. Moore, A., Wozniak, M., Yousef, A., Barnes, C. C., Cha, D., Courchesne, E., & Pierce, K. (2018). The geometric preference subtype in ASD: identifying a consistent, early-emerging phenomenon through eye tracking. *Molecular autism*, 9(1), 19.
57. Mundy, P. (2018). A review of joint attention and social-cognitive brain systems in typical development and autism spectrum disorder. *European Journal of Neuroscience*, 47(6), 497-514.
58. Munn, S. M., Stefano, L., & Pelz, J. B. (2008, August). Fixation-identification in dynamic scenes: Comparing an automated algorithm to manual coding. In *Proceedings of the 5th symposium on Applied perception in graphics and visualization* (pp. 33-42). ACM.
59. Naber, F. B., Bakermans-Kranenburg, M. J., Van IJzendoorn, M. H., Dietz, C., Van Daalen, E., Swinkels, S. H., ... & Van Engeland, H. (2008). Joint attention development in toddlers with autism. *European Child & Adolescent Psychiatry*, 17(3), 143-152.
60. Oliveira, J. D. S. D. *Classificador para auxílio ao diagnóstico de TEA baseado em um modelo computacional de atenção visual* (Doctoral dissertation, Universidade de São Paulo).

61. Paula, C. S., Ribeiro, S. H., Fombonne, E., & Mercadante, M. T. (2011). Brief report: prevalence of pervasive developmental disorder in Brazil: a pilot study. *Journal of autism and developmental disorders*, 41(12), 1738-1742.
62. Pellicano, E., & Burr, D. (2012). When the world becomes 'too real': a Bayesian explanation of autistic perception. *Trends in cognitive sciences*, 16(10), 504-510.
63. Pelphrey, K. A., Sasson, N. J., Reznick, J. S., Paul, G., Goldman, B. D., & Piven, J. (2002). Visual scanning of faces in autism. *Journal of autism and developmental disorders*, 32(4), 249-261.
64. Pereira, A., Riesgo, R. S., & Wagner, M. B. (2008). Childhood autism: translation and validation of the Childhood Autism Rating Scale for use in Brazil. *Jornal de Pediatria*, 84(6), 487-494.
65. Perrett, D. I., & Emery, N. J. (1994). Understanding the intentions of others from visual signals: neurophysiological evidence.
66. Peters, R. J., & Itti, L. (2007, June). Beyond bottom-up: Incorporating task-dependent influences into a computational model of spatial attention. In *2007 IEEE conference on computer vision and pattern recognition* (pp. 1-8). IEEE.
67. Pierce, K., Marinero, S., Hazin, R., McKenna, B., Barnes, C. C., & Malige, A. (2016). Eye-tracking reveals abnormal visual preference for geometric images as an early biomarker of an ASD subtype associated with increased symptom severity. *Biological psychiatry*, 79(8), 657.
68. Rayner, K. (1998). Eye movements in reading and information processing: 20 years of research. *Psychological bulletin*, 124(3), 372.
69. Rigo, G. V. (2010). *Modelando a atenção seletiva e a saliência visual através de redes complexas* (Doctoral dissertation, Universidade de São Paulo).
70. Rubenstein, J. L. R., & Merzenich, M. M. (2003). Model of autism: increased ratio of excitation/inhibition in key neural systems. *Genes, Brain and Behavior*, 2(5), 255-267.
71. Sacrey, L. A. R., Bryson, S. E., & Zwaigenbaum, L. (2013). Prospective examination of visual attention during play in infants at high-risk for autism spectrum disorder: A longitudinal study from 6 to 36 months of age. *Behavioural brain research*, 256, 441-450.
72. Saitovitch, A., Bargiacchi, A., Chabane, N., Brunelle, F., Samson, Y., Boddart, N., & Zilbovicius, M. (2012). Social cognition and the superior temporal sulcus: implications in autism. *Revue neurologique*, 168(10), 762-770.
73. Senju, A., & Johnson, M. H. (2009). The eye contact effect: mechanisms and development. *Trends in cognitive sciences*, 13(3), 127-134.
74. Seymour, R. A., Rippon, G., Gooding-Williams, G., Schoffelen, J. M., & Kessler, K. (2019). Dysregulated oscillatory connectivity in the visual system in autism spectrum disorder. *bioRxiv*, 440586.
75. Simoncelli, E. P., & Freeman, W. T. (1995, October). The steerable pyramid: A flexible architecture for multi-scale derivative computation. In *Proceedings., International Conference on Image Processing* (Vol. 3, pp. 444-447). IEEE.
76. Swanson, M. R., & Siller, M. (2013). Patterns of gaze behavior during an eye-tracking measure of joint attention in typically developing children and children with autism spectrum disorder. *Research in Autism Spectrum Disorders*, 7(9), 1087-1096.
77. Treisman, A. M., & Gelade, G. (1980). A feature-integration theory of attention. *Cognitive psychology*, 12(1), 97-136.
78. Viola, P., & Jones, M. (2001). Robust real-time object detection. *International journal of computer vision*, 4(34-47), 4.
79. Walther, D., & Koch, C. (2006). Modeling attention to salient proto-objects. *Neural networks*, 19(9), 1395-1407.



80. Wang, S., Jiang, M., Duchesne, X. M., Laugeson, E. A., Kennedy, D. P., Adolphs, R., & Zhao, Q. (2015). Atypical visual saliency in autism spectrum disorder quantified through model-based eye tracking. *Neuron*, 88(3), 604-616.
81. Wong, C., Odom, S. L., Hume, K. A., Cox, A. W., Fettig, A., Kucharczyk, S., ... & Schultz, T. R. (2015). Evidence-based practices for children, youth, and young adults with autism spectrum disorder: A comprehensive review. *Journal of autism and developmental disorders*, 45(7), 1951-1966.
82. Yamaguchi, K. (2013). mexopencv. Collection and a development kit of matlab mex functions for OpenCV library, available at <http://www.cs.stonybrook.edu/~kyamagu/mexopencv>.
83. Yirmiya, N., Gamliel, I., Pilowsky, T., Feldman, R., Baron-Cohen, S., & Sigman, M. (2006). The development of siblings of children with autism at 4 and 14 months: Social engagement, communication, and cognition. *Journal of Child Psychology and Psychiatry*, 47(5), 511-523.
84. Zikopoulos, B., & Barbas, H. (2013). Altered neural connectivity in excitatory and inhibitory cortical circuits in autism. *Frontiers in human neuroscience*, 7, 609.
85. Zilbovicius, M., Meresse, I., Chabane, N., Brunelle, F., Samson, Y., & Boddaert, N. (2006). Autism, the superior temporal sulcus and social perception. *Trends in neurosciences*, 29(7), 359-366.
86. Zwaigenbaum, L., Bauman, M. L., Choueiri, R., Kasari, C., Carter, A., Granpeesheh, D., ... & Pierce, K. (2015). Early intervention for children with autism spectrum disorder under 3 years of age: recommendations for practice and research. *Pediatrics*, 136(Supplement 1), S60-S81.