

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM NEUROCIÊNCIAS

**DESENVOLVIMENTO DE UM MÉTODO PARA CLASSIFICAÇÃO DE  
COMPORTAMENTOS DE RATOS WISTAR UTILIZANDO O ALGORITMO DE  
APRENDIZADO SUPERVISIONADO FLORESTAS ALEATÓRIAS (*RANDOM  
FORESTS*)**

DISSERTAÇÃO DE MESTRADO EM NEUROCIÊNCIAS

Porto Alegre

Fevereiro de 2020

UNIVERSIDADE FEDERAL DO RIO GRANDE DO SUL  
INSTITUTO DE CIÊNCIAS BÁSICAS DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM NEUROCIÊNCIAS

**DESENVOLVIMENTO DE UM MÉTODO PARA CLASSIFICAÇÃO DE  
COMPORTAMENTOS DE RATOS WISTAR UTILIZANDO O ALGORITMO DE  
APRENDIZADO SUPERVISIONADO FLORESTAS ALEATÓRIAS (*RANDOM  
FORESTS*)**

Dissertação apresentada ao PPG Neurociências  
da Universidade Federal do Rio Grande do Sul  
como requisito parcial para obtenção do título de  
Mestre em Neurociências.

MATHEUS COSTA SCHIAFFINO

ORIENTADORA: Dra. Carmem Juracy Silveira Gottfried

Porto Alegre

Fevereiro de 2020

## **Agradecimentos**

Agradeço aos familiares e amigos que me acompanharam na vida pessoal e acadêmica.

Aos colegas de laboratório, especialmente Gui e Mell, pelo apoio e parceria.

À Professora Carmem pelo acolhimento no laboratório.

Aos professores Altamiro Amadeu Susin, Jorge Alberto Quillfeldt e Mirna Bainy Leal pela contribuição com este projeto.

## Resumo

A despeito da grande importância dos testes comportamentais na pesquisa pré-clínica sobre os transtornos neuropsiquiátricos e do neurodesenvolvimento, como o transtorno do espectro do autismo (TEA), poucas ferramentas computacionais existem para avaliar quantitativamente o comportamento animal. Além disso, o funcionamento da maior parte das poucas ferramentas existentes depende de softwares adicionais ou de hardware de alto custo, como câmeras de profundidade e gaiolas instrumentadas. Neste trabalho, é apresentado um método para avaliação do comportamento de ratos Wistar a partir de vídeos do teste de campo aberto, empregando técnicas de visão computacional e aprendizado de máquina. Todos os passos do método, do processamento e segmentação das imagens até a aplicação dos algoritmos de aprendizado de máquina estão descritos aqui. Foram rotulados manualmente 47.466 quadros de 52 vídeos do teste de campo aberto, atribuindo a cada um deles uma de sete classes: caminhando, cheirando, escalando, *rearing*, autolimpeza, imobilidade e outros/sem certeza. Foi realizada subtração de fundo e binarização quadro-a-quadro para segmentar as imagens, separando a região correspondente ao animal do aparato de teste. A partir das imagens segmentadas, foram extraídos parâmetros geométricos de pose, e o método das janelas deslizantes foi empregado para levar em consideração o aspecto dinâmico do comportamento. Por fim, os dados extraídos foram utilizados juntamente com os rótulos manuais para treinar e comparar o desempenho de diversos classificadores utilizando algoritmos de aprendizado de máquina. Nosso sistema foi treinado para classificar os comportamentos caminhando (F1-score = 91,85%), cheirando (F1-score = 85,39%), *rearing* (F1-score = 80,94%), autolimpeza dianteira (F1-score = 59,39%), autolimpeza traseira (F1-score = 72,48%) e imobilidade (F1-score = 75,04%). O método desenvolvido permite detecção rápida e precisa dos comportamentos treinados, consistindo, portanto, em uma ferramenta valiosa para a pesquisa pré-clínica.

## Abstract

Despite the large importance of behavioral tests in the preclinical research on neuropsychiatric and neurodevelopmental disorders, such as autism spectrum disorder (ASD), few computational tools exist to quantitatively evaluate animal behavior. In addition, most of the few existing tools rely on additional software or expensive hardware such as depth cameras or instrumented cages to run. In this paper, we present a method for albino rat behavior assessment from open field test videos utilizing computer vision and supervised machine learning techniques. All steps of the method, from image processing and segmentation to the application of machine learning techniques are here described. We manually labeled 46,521 frames from 52 videos of the open field test, assigning each of the frames to one of seven classes: walking, sniffing, rearing, fore body grooming, hind body grooming, immobility and other/unsure. We utilized background subtraction, thresholding and morphological opening on a frame-by-frame basis to segment the images, separating the region corresponding to the animal from the test apparatus. We then extracted geometrical pose estimators from the segmented region and applied the sliding windows technique to them to account for the time dynamics of behavior. Finally, we utilized the extracted data, along with the manually scored labels, to train binary classifiers using the Random Forest algorithm. Our system was trained for classifying walking (F1-score = 91.85%), sniffing (F1-score = 85.39%), rearing (F1-score = 80.94%), fore body grooming (F1-score = 59.39%), hind body grooming (F1-score = 72.48%) and immobility (F1-score = 75.04%). The method we developed allows for reliable and rapid detection of the trained behaviors, therefore consisting of a valuable tool for preclinical research.

## Lista de abreviaturas e siglas

ABA – *Applied Behavior Analysis*

APA – *American Psychiatric Association*

CDC – *Centers for Disease Control and Prevention*

DSM-5 – Manual Diagnóstico e Estatístico de Transtornos Mentais, do inglês *Diagnostic and Statistical Manual of Mental Disorders*

FN – *False Negatives*

FP – *False Positives*

JAABA - *Janelia Automatic Animal Behavior Annotator*

KNN – *K Nearest Neighbors*

MMWR – *Morbidity and Mortality Weekly Report*

OVA – *One Versus All*

OVO – *One Versus One*

RF – *Random Forests*

SVM – *Support Vector Machines*

TEA – Transtorno do Espectro do Autismo

TEACCH – *Treatment and Educational of Autistic and Related Communication-handicapped Children*

TOC – Transtorno Obsessivo Compulsivo

TN – *True Negatives*

TP – *True Positives*

VPA – *Valproic Acid*

## Lista de figuras

Figura 1: Teste de sociabilidade recíproca.....	14
Figura 2: Teste de três câmaras.....	15
Figura 3: Teste de campo aberto .....	16
Figura 4: Software ANY-maze.....	20
Figura 5: Visão geral do funcionamento do software JAABA .....	22
Figura 6: Subtração de fundo, binarização e abertura morfológica .....	23
Figura 7: Erosão de uma imagem binária .....	25
Figura 8: Dilatação de uma imagem binária.....	26
Figura 9: Abertura morfológica de uma imagem binária.....	27
Figura 10: Método das janelas deslizantes .....	31
Figura 11: kNN .....	33
Figura 12: SVM .....	34
Figura 13: Árvores de decisão.....	35
Figura 14: Random forests .....	36
Figura 15: A maldição da dimensionalidade.....	40

## SUMÁRIO

1. INTRODUÇÃO .....	9
1.1. O Transtorno do Espectro Autista (TEA).....	9
1.1.1. Sintomatologia e epidemiologia .....	9
1.1.2. Desencadeamento do TEA.....	12
1.1.3. O modelo animal de autismo induzido por exposição ao ácido valproico.....	12
1.1.4. Análise comportamental .....	13
1.2. Fundamentação teórica computacional.....	22
1.2.1. Segmentação das imagens .....	23
1.2.2. Extração e processamento de dados .....	27
1.2.3. Aprendizado de máquina.....	31
2. JUSTIFICATIVA .....	41
3. OBJETIVOS .....	42
3.1. Objetivo geral .....	42
3.2. Objetivos específicos .....	42
4. TRABALHO NA FORMA DE ARTIGO CIENTÍFICO .....	43
5. CONCLUSÕES .....	71
6. PERSPECTIVAS.....	73
REFERÊNCIAS.....	74

## 1. INTRODUÇÃO

### 1.1. O Transtorno do Espectro Autista (TEA)

O Transtorno do Espectro Autista (TEA) compreende um conjunto de desordens do neurodesenvolvimento caracterizadas por uma tríade comportamental, com a presença de déficits de habilidades sociais, como a comunicação e interação, e de comportamentos repetitivos e estereotipados (American Psychiatry Association (APA), 2013).

O termo “autismo” foi introduzido pelo psiquiatra suíço Paul Eugen Bleuler, em 1911. Bleuler fez uso do prefixo grego autós (αὐτός, traduzido como “auto”) para cunhar o termo, que utilizou inicialmente na descrição de pacientes com esquizofrenia que tinham a característica de se “retirar para o mundo de suas fantasias” (Bleuler, 1911). Embora as primeiras descrições clínicas de autismo sejam comumente creditadas a Hans Asperger e Leo Kanner, o artigo de 1926 intitulado “Die schizoiden Psychopathien im Kindesalter”, de autoria da psiquiatra soviética Grunya Efiminova Sukhareva, é o mais antigo registro conhecido do transtorno (Wolff, 1996). No artigo, Sukhareva descreve seis casos de meninos com “tendências autistas”: todos introspectivos, com preferências pelos seus próprios mundos interiores, mas cada um com suas peculiaridades e talentos (Sukhareva, 1926).

#### 1.1.1. Sintomatologia e epidemiologia

Atualmente, utiliza-se o termo Transtorno do Espectro Autista para categorizar um conjunto de alterações comportamentais específicas, visto que cada paciente apresenta sintomas únicos, com diferentes graus de severidade, gerando um espectro de manifestações (Gadia et al., 2004; Rapin and Tuchman, 2008). Segundo o DSM-5 (Diagnostic and Statistical Manual of Mental Disorders), a tríade de alterações comportamentais que caracteriza o TEA consiste em (American Psychiatry Association (APA), 2013):

- Prejuízos na comunicação e interação social, incluindo déficits na reciprocidade social, na comunicação verbal e não-verbal e no estabelecimento e manutenção de relacionamentos.
- Presença de comportamentos repetitivos e estereotipados, com atividades e interesses restritos.

As características supracitadas podem estar presentes em diferentes intensidades, mas o critério de diagnóstico estabelece que ambos devem estar presentes ao longo de toda a vida (Rapin e Tuchman, 2008). Como o TEA é caracterizado por qualidades comportamentais, não havendo biomarcadores do transtorno, seu diagnóstico é estritamente clínico (Mayes et al., 2014).

Embora o diagnóstico do TEA exija ambas as características da tríade comportamental referida acima, trata-se de um transtorno multifatorial, que tipicamente possui diversas comorbidades associadas (Gottfried et al., 2013). Sintomas que frequentemente acompanham o diagnóstico do transtorno incluem comportamento agressivo, ansiedade, depressão, déficits cognitivos, crises epiléticas, alterações gastrointestinais, distúrbios do sono, disfunções do sistema imune e hipersensibilidade a estímulos táteis, sonoros e luminosos (Geschwind, 2009). Os sintomas se manifestam em diferentes formas entre os indivíduos, de modo que pacientes nunca compartilham exatamente as mesmas características (Rapin e Tuchman, 2008).

Além dos sintomas comportamentais já mencionados, diversas alterações do sistema nervoso já foram documentadas em pacientes e modelos animais de TEA, em regiões como o córtex pré-frontal (Courchesne et al., 2001), cerebelo (Aldinger et al., 2012), hipocampo e amígdala (Groen et al., 2010). No entanto, estas alterações não são suficientes para explicar as manifestações comportamentais exibidas por pacientes.

A abordagem clínica de tratamento do TEA consiste em controlar aspectos específicos de sua sintomatologia, visto que não há tratamento eficaz contra o transtorno (Riesgo, 2013). O tratamento farmacológico atualmente aprovado pelo FDA (do inglês *Food and Drug Administration*) restringe-se à risperidona e ao aripiprazol, antipsicóticos atípicos que visam controlar o comportamento agressivo apresentado por alguns pacientes (Riesgo, 2013), além de fármacos para controlar comorbidades

como distúrbios de humor e do sono (Tsujino et al., 2007). A fim de promover uma melhora dos sintomas, algumas abordagens de terapia comportamental foram desenvolvidas, como a ABA (do inglês *Applied Behavior Analysis*) e o TEACCH (do inglês *Treatment and Educational of Autistic and Related Communication-handicapped Children*) (Riesgo, 2013).

De acordo com o último estudo populacional do CDC (do inglês *Center for Disease Control and Preventions*), publicado em 2018 e avaliando dados do ano de 2014, o autismo tem incidência de 1:59 crianças nos Estados Unidos da América, sendo aproximadamente 5 vezes mais comum entre meninos (1:42) do que em meninas (1:189) (*Morbidity and Mortality Weekly Report* (MMWR), 2014). Ao longo das últimas décadas, a incidência vem aumentando com uma taxa alarmante, tendo aproximadamente dobrado entre 2004 e 2014 (Summaries, 2007). Acredita-se que este aumento se deva, parcialmente, a mudanças nos critérios de diagnósticos, o que aumenta o número de indivíduos identificados com TEA (Fombonne et al., 2009). Além disso, há indícios que a exposição a fatores ambientais de risco possa estar contribuindo para este aumento (Bryson et al., 2003; Schneider e Przewlocki, 2005).

No Brasil, existe carência de dados epidemiológicos a respeito do TEA. Os poucos dados existentes apontam prevalência de 0,3%, em um estudo realizado em Atibaia – SP (Paula et al., 2011), consideravelmente abaixo dos valores encontrados nos EUA. Isto pode possivelmente ser explicado por problemas metodológicos relacionados aos pequenos tamanhos de amostra utilizadas (*Morbidity and Mortality Weekly Report* (MMWR), 2014).

Existe um custo econômico elevado que acompanha a alta prevalência do TEA, visto que indivíduos diagnosticados com o transtorno frequentemente requerem atenção diferenciada. Estima-se que exista um custo adicional de, pelo menos, US\$ 17.000 por ano no cuidado de cada criança com TEA (Lavelle et al., 2014). Ao longo da vida, o custo médio total por indivíduo com TEA é em torno de US\$ 3.600.000 nos EUA (Cakir et al., 2019).

### 1.1.2. Desencadeamento do TEA

Acredita-se que a etiologia do TEA consista em uma interação entre fatores genéticos e ambientais. Estudos realizados com gêmeos demonstram que existe concordância de 60% entre gêmeos monozigóticos no autismo clássico, contra 0% entre gêmeos dizigóticos (Muhle et al., 2004). Considerando-se características mais abrangentes, como déficits de comunicação e alterações de sociabilidade, a concordância pode ser ainda mais elevada, sendo 92% entre gêmeos monozigóticos e 10% entre gêmeos dizigóticos (Casanova, 2006). Estes estudos de associação gênica fornecem indícios que existe um componente genético forte na origem do autismo. No entanto, embora mais de 1000 alterações genéticas documentadas tenham sido associadas ao autismo, elas permitem explicar apenas em torno de 25% dos casos do transtorno (Miles, 2011).

Em virtude da impossibilidade de explicar a etiologia do autismo considerando apenas fatores genéticos, possibilidades vêm sendo levantadas acerca dos fatores ambientais que podem estar envolvidos na expressão e desencadeamento dos sintomas (Muhle et al., 2004; Sandin et al., 2014). Pesquisas indicam que fatores de risco ambiental durante o primeiro trimestre de gestação podem estar relacionados com o desencadeamento do TEA, destacando-se o uso do fármaco ácido valproico (VPA, do inglês *valproic acid*) (Roullet et al., 2013; Smith e Brown, 2014) e a talidomida (Imai et al., 2014) durante esse período. Estudos demonstram que cerca de 30% dos indivíduos expostos à talidomida durante o período gestacional, especialmente entre os dias 20 e 24, desenvolveram TEA (Miller e Stromland, 1999). Cabe ressaltar que a talidomida apresenta efeitos diferentes em roedores e primatas, nestes últimos podendo causar anomalias nos membros, enquanto o VPA provoca alterações similares em roedores e primatas (Gottfried et al., 2013).

### 1.1.3. O modelo animal de autismo induzido por exposição ao ácido valproico

A partir de observações de que a exposição pré-natal ao ácido valproico aumenta significativamente o risco de desenvolver TEA, foi desenvolvido, pela

embriologista Patricia Rodier, um modelo animal utilizando ratos para o estudo do autismo (Rodier et al., 1997; Harden, 2013). Além do gatilho etiológico, animais induzidos com o modelo apresentam uma série de alterações comportamentais e neuromorfológicas em comum com pacientes humanos. Pode-se listar a presença da tríade comportamental característica do TEA (comportamentos repetitivos e estereotipados e déficits de sociabilidade) (Roullet et al., 2013), alterações no cerebelo (Ingram et al., 2000), nos nervos cranianos (Rodier et al., 1996) e outras regiões encefálicas (van Kooten et al., 2005). Além disso, constata-se a presença de padrões eletroencefalográficos anormais em resposta a estímulos auditivos (Gandal et al., 2010) e alterações sensoriais somáticas (Fontes-Dutra et al., 2018). Analogamente ao caso humano, em que existe maior incidência de autismo em indivíduos do sexo masculino, algumas características do modelo estão presentes somente em machos (Kim et al., 2013; Schneider et al., 2008). O modelo estabelecido por Rodier é atualmente amplamente empregado na pesquisa pré-clínica, na avaliação de alterações causadas pelo teratígeno e na busca por alternativas de tratamento.

#### 1.1.4. Análise comportamental

Como o autismo é caracterizado por alterações comportamentais, e visto que atualmente não existem biomarcadores validados do transtorno, testes que permitem avaliar alterações em modelos animais são amplamente empregados na pesquisa sobre o TEA. Conforme mencionado anteriormente, o autismo é caracterizado por déficits de sociabilidade e presença de comportamentos repetitivos e estereotipados. Alguns testes que permitem avaliar a presença da tríade em modelos animais serão descritos a seguir.

#### 1.1.4.1. Testes comumente empregados

Com o intuito de verificar o comportamento social de ratos do modelo de autismo, foram desenvolvidos testes em que se monitora a interação destes com coespecíficos. Os dois testes mais empregados nesta categoria são o de sociabilidade recíproca e o teste de três câmaras (Chang et al., 2017).

No teste de sociabilidade recíproca (Figura 1), dois animais são colocados em um aparato, tipicamente quadrado, com paredes, durante 10 minutos, podendo interagir livremente. A interação entre os animais é quantificada em categorias, como investigação ano-genital, investigação da cabeça, investigação do corpo, comportamentos agressivos, perseguir e comportamentos de dominância.

Figura 1: Teste de sociabilidade recíproca



Fonte: autor

O teste de três câmaras (Figura 2) também é utilizado para medida de comportamento social em roedores, sendo realizado em um aparato com três câmaras conectadas por portas. O animal-teste pode circular por todo o aparato. As câmaras laterais possuem gaiolas, que podem ser mantidas vazias ou podem conter um coespecífico do animal testado. O teste é realizado em duas etapas. Na primeira, afere-se o tempo de interação de um animal com um coespecífico que se encontra dentro de uma gaiola e o tempo de interação com uma gaiola vazia. Na segunda,

chamada teste de preferência por novidade social, afere-se o tempo de interação do animal com um coespecífico com que já está familiarizado e com outro com que ainda não está familiarizado. Animais do grupo controle têm a tendência a interagir mais tempo com a gaiola que contém o animal na primeira etapa, e de interagir mais tempo com a gaiola que contém o animal não conhecido na segunda etapa. Animais do modelo de autismo não possuem esta preferência (Bambini-Junior, 2014).

Testes que visam avaliar a presença de estereotípias no comportamento dos animais também vêm sendo empregados na pesquisa sobre o TEA. Os mais comuns são o de enterramento de bolinhas, alternância no labirinto em T e campo aberto.

Figura 2: Teste de três câmaras



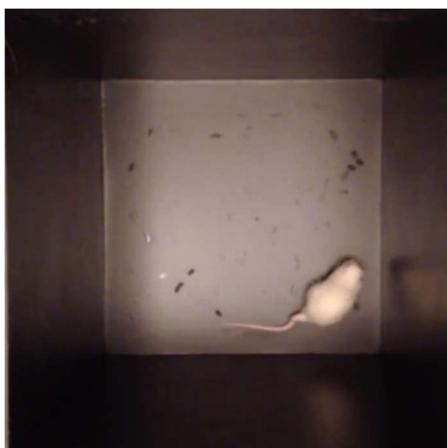
Fonte: autor

No teste de enterramento de bolinhas, determinada quantidade de pequenas bolas de vidro é colocada na maravalha do animal-teste. A quantidade de bolas que o animal enterra fornece um indicativo de ansiedade e rigidez comportamental, e vem sendo utilizado no estudo em modelos animais de Transtorno Obsessivo Compulsivo (TOC) e TEA (Njung'e e Handley, 1991; Chang et al., 2017).

No teste de alternância no labirinto em T, o animal-teste é colocado no centro de um aparato em forma de T. Animais do grupo controle tendem a alternar a exploração dos braços do aparato, enquanto animais de modelos que exibem rigidez comportamental tendem a explorar sempre o mesmo braço (Deacon e Rawlins, 2006).

O teste de campo aberto (Figura 3), que é objeto de estudo do presente trabalho, foi utilizado pela primeira vez em 1934 como forma de avaliar comportamento emocional em ratos através da observação de defecação e urinação em diferentes circunstâncias (Hall, 1934). No teste, um animal é colocado em um aparato, tipicamente quadrado, com paredes, onde pode explorar livremente durante 10 minutos. Atualmente, o teste de campo aberto é comumente utilizado para avaliar atividade locomotora, através do monitoramento da distância total percorrida ou do número de vezes que o animal passa por determinada região. O tempo que o animal passa no centro do aparato e próximo às paredes também é utilizado como indicativo de comportamento relativo à ansiedade, visto que os ratos tendem a preferir as regiões mais protegidas quando se sentem ameaçados. O tempo que o animal passa realizando comportamentos específicos como a autolimpeza, a imobilidade ou o *rearing* também podem ser monitorados (Gould et al., 2009). Alguns dos principais comportamentos avaliados no teste de campo aberto serão descritos a seguir.

Figura 3: Teste de campo aberto



Fonte: autor

O comportamento de caminhar é definido como atividade locomotora relativamente rápida e que não envolve exploração cheirando (Coleris et al., 2001). Este comportamento é frequentemente avaliado na pesquisa sobre transtornos como ansiedade e doença de Parkinson, que apresentam alteração na locomoção (Kang e Meredith, 2006; Ramos, 2008).

O comportamento de cheirar envolve atividade de exploração cheirando o chão do aparato de testes, suas paredes ou o ar (Coleris et al., 2001).

O comportamento de *rearing* ocorre quando o animal-teste levanta as patas dianteiras do chão, tocando ou não as paredes do aparato. O comportamento de *rearing* com apoio, também conhecido como escalar, ocorre quando o animal encosta nas paredes do aparato com suas patas dianteiras. Estes comportamentos são dependentes do contexto, possuindo ocorrência reduzida em condições aversivas. Além disso, fêmeas tendem a realizar menos dos dois comportamentos que machos (Sturman et al., 2018).

A autolimpeza é um comportamento comum em roedores, que ocorre quando o animal lambe ou esfrega seu próprio corpo. Tipicamente ocorre em sentido céfalo-caudal e possui uma microestrutura definida consistindo em cinco etapas, sendo altamente sensível a estressores e outros fatores. Em virtude da dificuldade de diferenciar precisamente as cinco etapas a partir da observação de vídeos, estudos comportamentais frequentemente empregam a separação entre autolimpeza dianteira e traseira. Modelos animais de ansiedade, autismo, depressão, TOC, síndrome de Tourette e distúrbios nos núcleos da base apresentam comportamento de autolimpeza alterado (McFarlane et al., 2007; Kalueff et al., 2007a). Tipicamente, a atividade de autolimpeza é avaliada no teste de campo aberto observando a latência, duração total e número de vezes que o animal realiza o comportamento (Kalueff et al., 2007a).

A imobilidade é caracterizada por ausência total de movimentos corporais em roedores, frequentemente ocorrendo como uma resposta normal a estímulos de medo (Brandão et al., 2008). Em ratos selvagens, a imobilidade consiste em um dos três mecanismos de resposta ao medo, ocorrendo, por exemplo, quando o animal constata a presença de um predador à distância (Blanchard et al., 1986). No teste de campo aberto, o excesso de comportamento de imobilidade é comumente relacionado com ansiedade, transtorno caracterizado por uma regulação anormal do medo (Díaz-Morán et al., 2014; Kandel et al., 2014).

#### 1.1.4.2. Métodos manuais de análise

Tradicionalmente, os testes comportamentais em roedores são filmados e avaliados posteriormente por pesquisadores treinados, que assistem aos vídeos, cronometram e marcam manualmente variáveis de interesse. No teste de campo aberto, por exemplo, caso se deseje monitorar o tempo que o animal passa realizando o comportamento de autolimpeza, é preciso que um investigador assista aos vídeos com atenção, detectando manualmente os instantes de tempo em que ocorreu o comportamento relevante.

Esta abordagem está sujeita a diversos erros e limitações. A avaliação realizada desta forma requer a atenção de investigadores treinados por longos períodos. Além disso, a análise, por ser subjetiva, frequentemente apresenta imprecisões e baixa reprodutibilidade. Diferentes pesquisadores avaliando o mesmo vídeo, ou mesmo um pesquisador avaliando o mesmo vídeo em duas ocasiões diferentes, podem fornecer resultados muito distintos. A análise manual acarreta também em um limite no número de variáveis que podem ser monitoradas com os recursos disponíveis, impossibilitando o monitoramento de padrões comportamentais complexos.

#### 1.1.4.3. Métodos automatizados de análise

Em virtude de todas as limitações citadas na subseção anterior, a busca por abordagens automatizadas e assistidas por computador para avaliação de comportamento animal vem surgindo no estudo do TEA e outros transtornos, constituindo um passo imprescindível para refinar os resultados obtidos na investigação sobre fatores de risco e fármacos para tratamento. Além disso, o domínio das neurociências está passando por uma revolução de novas tecnologias tratando do mapeamento, monitoramento e manipulação da atividade de populações específicas de neurônios, graças a técnicas como a optogenética e a

farmacogenética. Estas técnicas fornecem a possibilidade de estabelecer relações causais entre a atividade de redes de neurônios e comportamento. No entanto, a possibilidade de explorar e estabelecer estas relações depende diretamente da capacidade de aferir com precisão as consequências das manipulações (Anderson e Perona, 2014).

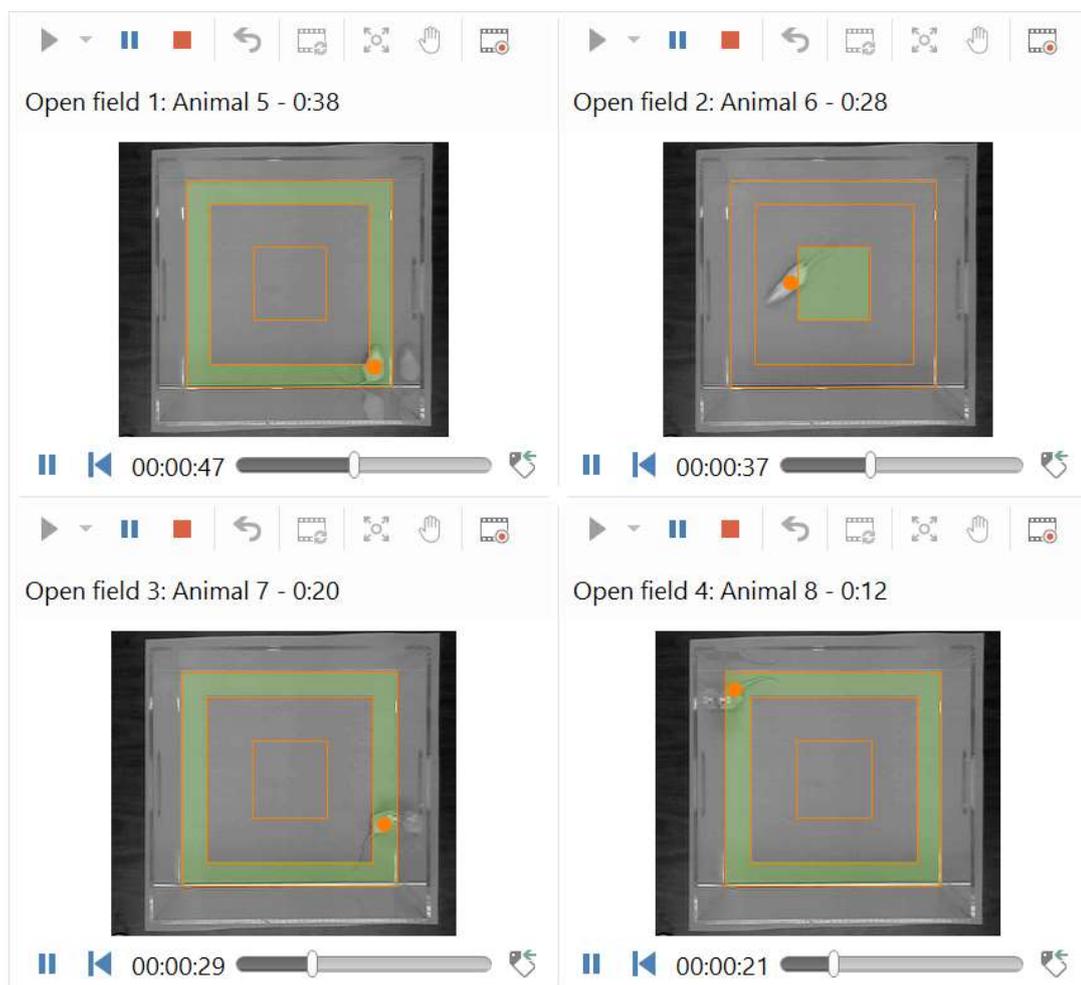
O campo da etologia computacional surgiu da união de áreas das ciências biológicas, da computação e engenharias, com o objetivo de fornecer métodos de alto desempenho para avaliação de comportamento animal. Atualmente, diversos métodos automatizados estão disponíveis, utilizando diferentes técnicas de rastreamento e classificação. As abordagens mais comumente empregadas serão descritas a seguir.

Os sistemas mais antigos empregam métodos de visão computacional com técnicas tradicionais de programação. Seu funcionamento consiste em:

- Localizar o animal-teste nos vídeos, o distinguindo do fundo.
- Extrair parâmetros da forma do animal (ex.: posição de seu centroide, limites de suas bordas, velocidade de deslocamento).
- Comparar os parâmetros extraídos com critérios previamente definidos pelo usuário para identificar comportamentos de interesse (ex.: os instantes de tempo em que a velocidade foi superior a determinado valor).

Esta abordagem possui como principal vantagem sua simplicidade e eficiência do ponto de vista computacional, não requerendo grande quantidade de recursos. Sua maior limitação advém do fato que, empregando técnicas tradicionais de programação, é preciso definir explicitamente os critérios que devem ser utilizados para identificar comportamentos de interesse. Desta forma, este paradigma tende a ser bem-sucedido e adequado para identificação de comportamentos simples. No teste de campo-aberto, por exemplo, esta abordagem pode ser empregada para avaliar o tempo que o animal passa em uma região do centro do aparato e em uma região próxima às paredes do aparato. No entanto, a identificação de comportamentos complexos é dificilmente praticável. Exemplos de sistemas que empregam as técnicas supracitadas são os softwares comerciais ANY-maze (Stoelting Co., 2019) (Figura 4) e EthoVision (Noldus, 2020), bem como algumas ferramentas de código aberto, como Argus (Shams et al., 2018) e ezTrack (Pennington et al., 2019).

Figura 4: Software ANY-maze



Fonte: Stoelting Co., 2019

Com o intuito de contornar as limitações da abordagem citada acima, algumas aplicações começaram a empregar aprendizado de máquina para substituir a etapa de definição explícita dos parâmetros para identificar os comportamentos relevantes. O funcionamento destes sistemas consiste em (Anderson e Perona, 2014):

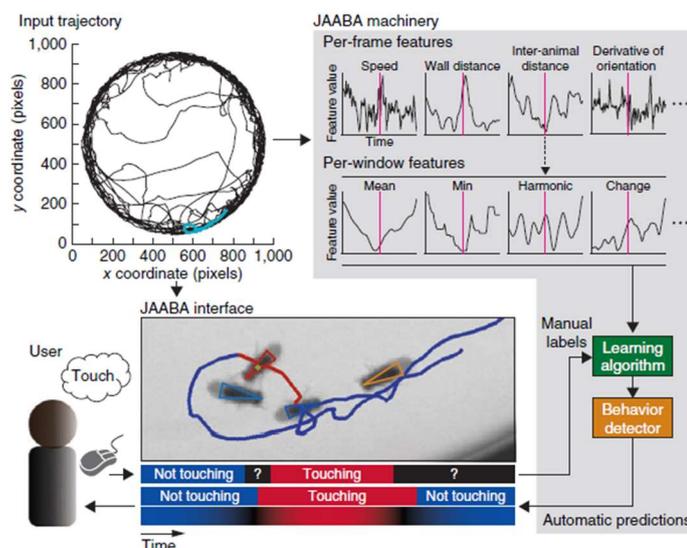
- Localizar o animal nos vídeos, o distinguindo do fundo.
- Extrair parâmetros da forma do animal (ex.: posição de seu centroide, limites de suas bordas, velocidade de deslocamento).

- Utilizar os parâmetros extraídos para treinamento de um sistema de classificação de comportamentos utilizando aprendizado de máquina supervisionado.

Os métodos de aprendizado de máquina (ver seção 1.2.3) permitem que critérios para classificação sejam aprendidos pelos algoritmos a partir de exemplos fornecidos pelo usuário, sem a necessidade de defini-los explicitamente. Isto permite que comportamentos mais complexos sejam avaliados. No entanto, o custo computacional de aplicações que empregam estas técnicas costuma ser elevado, o que fazia delas, até recentemente, pouco utilizadas. As primeiras publicações datam da última década, quando foram concebidos sistemas como JAABA (Figura 5), utilizado sobretudo para análise de comportamento de drosófilas e ratos (Kabra et al., 2013), os sistemas propostos por Hong e De Chaumont para análise de comportamento de camundongos (Hong et al., 2015; De Chaumont et al., 2019) e o sistema de Han e colaboradores para análise comportamental de cnidários da espécie hidra (Han et al., 2018). Outro inconveniente destas aplicações é o fato de que elas frequentemente requerem software ou hardware adicional para operar. O método de JAABA (Figura 5), por exemplo, funciona com base em parâmetros extraídos por software complementar de análise de vídeos, como Ctrax (Branson et al., 2009; Kabra et al., 2013). Os métodos descritos por Hong e De Chaumont requerem utilização de câmeras de profundidade (Hong et al., 2015; De Chaumont et al., 2019).

Há ainda métodos que consistem em aplicar técnicas de aprendizado de máquina, usualmente redes neurais artificiais, diretamente aos quadros de vídeos, sem qualquer processamento prévio das imagens. Exemplos incluem o método descrito por Stern e colaboradores para avaliação de comportamento de drosófila (Stern et al., 2015) e o método descrito por Ren e colaboradores para análise de memória em ratos (Ren et al., 2017). O treinamento e utilização de redes neurais artificiais para processamento de imagens tipicamente é um processo muito custoso do ponto de vista computacional, o que consiste em uma desvantagem destes métodos.

Figura 5: Visão geral do funcionamento do software JAABA



Fonte: KABRA et al., (2013)

## 1.2. Fundamentação teórica computacional

Nesta seção, serão abordados os fundamentos computacionais das técnicas empregadas na concepção do trabalho aqui apresentado, que possui os seguintes níveis de organização:

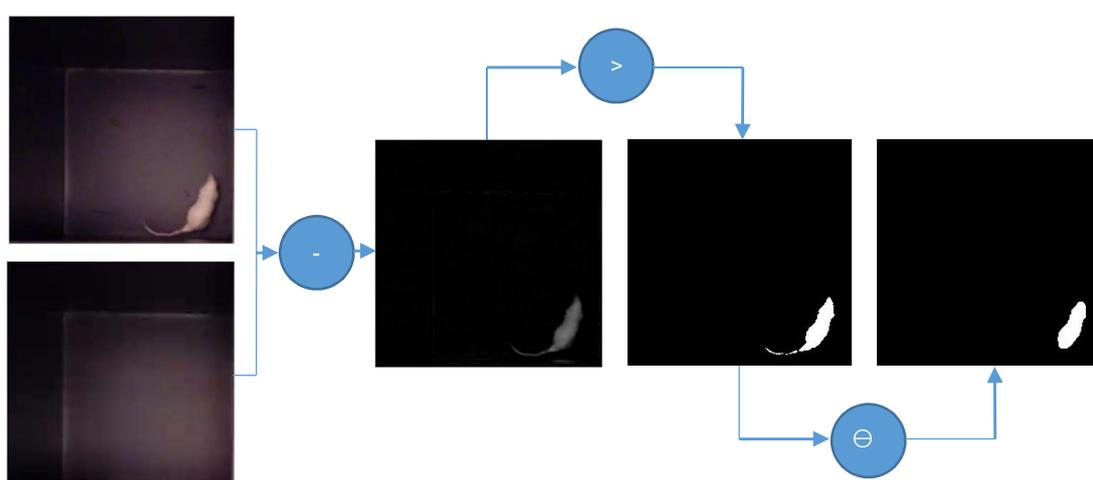
- a) Processamento e segmentação das imagens quadro-a-quadro a partir de vídeos, identificando qual região corresponde ao animal-teste e a separando do fundo.
- b) Extração de parâmetros das imagens segmentadas (ex.: área do animal, velocidade instantânea) que possam possuir correlação com o comportamento do animal.
- c) Criação de novos parâmetros a partir daqueles extraídos das imagens, a partir da aplicação do método das janelas deslizantes.
- d) Eliminação de parâmetros que possuem alta correlação com outros parâmetros.
- e) Utilização dos parâmetros restantes para treinar um algoritmo de aprendizado de máquina supervisionado para classificação dos comportamentos estudados.

Os níveis serão descritos a seguir.

### 1.2.1. Processamento e segmentação das imagens

O problema de detecção de objetos em vídeos é central em diversas aplicações de visão computacional, como detecção de intrusão, rastreamento de objetos e contagem de pessoas (Vacavant et al., 2013). No caso de câmeras estáticas, com pouca variação de luminosidade, o problema pode ser abordado utilizando técnicas elementares de subtração de fundo. Nesta operação, um quadro é definido como sendo a imagem de fundo. Todos os quadros subsequentes serão comparados com este primeiro, visando identificar as regiões que apresentam as maiores diferenças com o fundo, correspondendo a objetos que não se encontram na imagem de referência. A seguir, aplica-se uma operação de binarização na imagem resultante, obtendo-se em destaque o objeto que se deseja localizar. As imagens binárias comumente apresentam artefatos, que podem ser removidos com a aplicação de operações morfológicas como a abertura. A Figura 6 mostra um exemplo de imagem sendo processada através destas operações, que serão descritas em detalhe nas subseções a seguir.

Figura 6: Subtração de fundo, binarização e abertura morfológica



Fonte: autor

### 1.2.1.1. Subtração de fundo

A operação de subtração de fundo em imagens em escala de cinza, considerando o fundo estático, pode ser definida como (Tamersoy, 2009):

$$S(x, y, t) = |I(x, y, t) - I(x, y, t_0)| \quad (1)$$

Em que  $I(x, y, t)$  é a imagem em um instante qualquer  $t$  e  $I(x, y, t_0)$  é a imagem de fundo, definida em um instante de tempo  $t_0$ .  $S(x, y, t)$  é o resultado da operação. Em regiões em que  $I(x, y, t)$  for semelhante a  $I(x, y, t_0)$ , o resultado da subtração será próximo de 0. Em regiões onde houver grande diferença, o resultado será maior.

### 1.2.1.2. Método de Otsu

A subtração de fundo é tipicamente seguida da aplicação de uma operação de binarização na imagem, conforme mostra a Figura 6. Nesta operação, define-se um limiar  $L$ , e todos os pixels que possuem valor maior que esse limiar recebem valor 1; todos os pixels que possuem valor menor que o limiar recebem valor 0. Considerando  $B(x, y, t)$  o resultado da binarização, temos:

$$\begin{aligned} B(x, y, t) &= 1, \text{ onde } S(x, y, t) > L \\ &= 0, \text{ onde } S(x, y, t) < L \end{aligned} \quad (2)$$

O limiar para binarização pode ser definido manualmente ou calculado por métodos automáticos como o de Otsu (Otsu, 1979). O método de Otsu tem como objetivo definir automaticamente o limiar global ótimo para segmentação em imagens, analisando o histograma de uma imagem em escala de cinza e localizando o valor que permite melhor separação entre classes através da maximização da variância inter-classes das regiões de frente e fundo.

### 1.2.1.3. Abertura morfológica

Após a binarização, é comum que imagens apresentem artefatos, com o objeto detectado apresentando protruções em sua forma ou com objetos indesejados sendo detectados. Operações morfológicas como a abertura podem auxiliar na remoção destes artefatos, atuando na suavização da forma de objetos.

A erosão de uma imagem  $f$  por um elemento estruturante  $b$  em um ponto  $(x,y)$  é definida como o valor mínimo da imagem na região coincidente com  $b$  quando a origem de  $b$  está em  $(x,y)$  (Gonzalez e Woods, 2008). Em forma de equação, a erosão de  $f$  por  $b$  em  $(x,y)$  é dada por:

$$[f \ominus b](x, y) = \min_{(s,t) \in b} \{f(x + s, y + t)\} \quad (3)$$

De modo visual, após a erosão de uma imagem, ocorre uma contração dos limites de seus objetos, reduzindo seus tamanhos. Objetos menores que o elemento estruturante são eliminados (Murta, 2013). A Figura 7 apresenta um exemplo de erosão de uma imagem binária:

Figura 7: Erosão de uma imagem binária



Fonte: Mathworks

A dilatação é definida de maneira similar à erosão. A dilatação de uma imagem  $f$  por um elemento estruturante  $b$  em um ponto  $(x,y)$  é definida como o valor máximo da imagem na região coincidente com  $b$  quando a origem de  $b$  está em  $(x,y)$  (Gonzalez e Woods, 2008). Em forma de equação, temos que a dilatação de  $f$  por  $b$  em  $(x,y)$  é dada por:

$$[f \oplus b](x, y) = \max_{(s,t) \in b} \{f(x - s, y - t)\} \quad (4)$$

Após a aplicação de uma dilatação, ocorre a expansão dos contornos de objetos em uma imagem (Murta, 2013). A Figura 8 mostra um exemplo de dilatação em uma imagem binária:

Figura 8: Dilatação de uma imagem binária



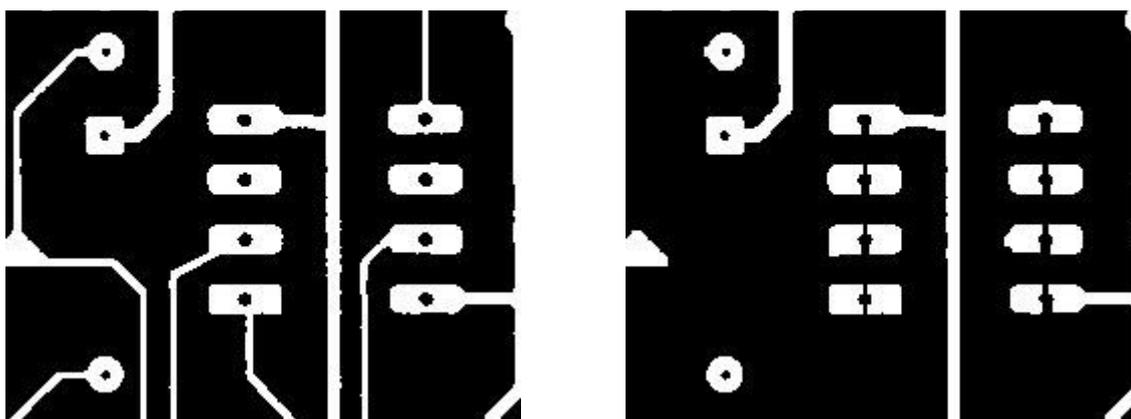
Fonte: Mathworks

A abertura morfológica é uma operação morfológica que consiste em realizar uma erosão seguida de uma dilatação (Gonzalez e Woods, 2008). Na forma de equação, temos:

$$A \circ B = (A \ominus B) \oplus B \quad (5)$$

Ela é geralmente utilizada com o objetivo de suavizar o contorno de objetos, eliminando protruções e pequenos objetos. A Figura 9 apresenta um exemplo em que a abertura é aplicada a uma imagem binária:

Figura 9: Abertura morfológica de uma imagem binária



Fonte: FEEC Unicamp

### 1.2.2. Extração e processamento de dados

A partir de objetos em imagens binárias, é possível extrair uma grande variedade de parâmetros numéricos representativos da sua geometria. Alguns dos parâmetros frequentemente utilizados em aplicações de visão computacional serão descritos a seguir.

#### **Área**

Em uma imagem binária, composta apenas por pixels de valor 0 ou 1, a área corresponde ao número de pixels com valor 1.

### **Área convexa**

Área do menor objeto convexo que contenha um conjunto de pontos. Um objeto  $O$  é dito convexo se qualquer segmento de reta ligando quaisquer dois pontos pertencentes a  $O$  estiver completamente dentro de  $O$ .

### **Solidez**

Razão entre a área de um objeto e sua área convexa:

$$\text{solidez} = \frac{\text{área}}{\text{área convexa}} \quad (6)$$

### **Comprimento do eixo maior**

Comprimento do eixo maior de uma elipse que possui o mesmo momento de inércia de área que dado objeto.

### **Comprimento do eixo menor**

Comprimento do eixo menor de uma elipse que possui o mesmo momento de inércia de área que dado objeto.

### **Excentricidade**

Excentricidade da elipse que possui o mesmo momento de inércia de área que dado objeto. Visualmente, pode ser entendida como uma medida do quão maior um eixo de um objeto é em relação ao outro.

## Orientação

Orientação do eixo maior da elipse que possui o mesmo momento de inércia de área que dado objeto. Pode ser definida como um ângulo com um eixo de referência.

## Distância de um objeto até uma região

Comprimento do menor segmento de reta que liga o centroide de um objeto a determinada região em uma imagem.

## Perímetro de um objeto

Somatório da distância entre pares pixels contíguos pertencentes à borda de um objeto. Os pixels definidos com valor 1 que possuem em suas adjacências pixels com valor 0 são considerados como pertencentes à borda.

## Posição do centroide

Ponto C de coordenadas (x,y) que corresponde à média aritmética da posição de todos os pontos de um objeto. O centroide C de um objeto formado por um número finito de pontos  $p_1, p_2, \dots, p_n$  em  $\mathbb{R}^n$  é dado por (Protter e Morrey, 1970):

$$C = \frac{p_1 + p_2 + \dots + p_n}{n} \quad (7)$$

## Intersecção de um objeto com uma região

Número de pixels que pertencem simultaneamente a dois objetos em duas imagens distintas. Pode ser definido como a área do resultado da multiplicação ponto-a-ponto de duas matrizes binárias.

### Velocidade instantânea

Comprimento do segmento de reta que liga a posição do centroide de um objeto entre dois quadros consecutivos  $i$  e  $i+1$ , definida como:

$$v_{instantânea} = \sqrt{(x_i - x_{i+1})^2 + (y_i - y_{i+1})^2} \quad (8)$$

### Velocidade n quadros

Comprimento do segmento de reta que liga a posição do centroide de um objeto entre dois quadros separados por  $n$  quadros,  $i$  e  $i+n$ , definida como:

$$v_{n \text{ quadros}} = \sqrt{(x_i - x_{i+n})^2 + (y_i - y_{i+n})^2} \quad (9)$$

### Ângulo da velocidade com o eixo maior

Ângulo entre o segmento de reta que liga a posição do centroide de um objeto entre dois quadros consecutivos  $i$  e  $i+1$  e sua orientação no quadro  $i$ .

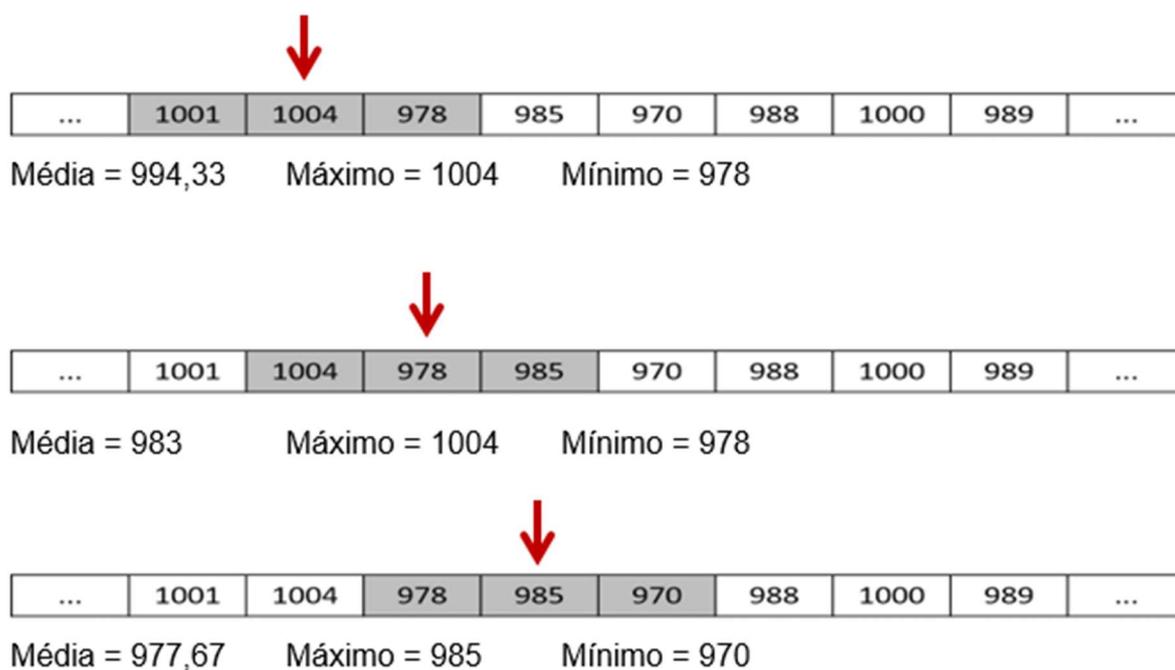
### Ângulo da velocidade com o eixo maior n frames

Ângulo entre o segmento de reta que liga a posição do centroide de um objeto entre dois quadros separados por  $n$  quadros,  $i$  e  $i+n$ , e sua orientação no quadro  $i$ .

Os parâmetros descritos acima descrevem extensivamente as características geométricas de um objeto a partir de uma imagem estática. No entanto, em algumas aplicações de visão computacional, como na análise do comportamento de animais, existe uma dinâmica temporal que deve ser levada em consideração. Um método frequentemente utilizado em reconhecimento de ações é o das janelas deslizantes

(Banos et al., 2014), que consiste em criar novos descritores para cada ponto de uma série temporal levando em consideração também os instantes de tempo anteriores e posteriores a ele. As janelas tipicamente são definidas com um tamanho ímpar (no exemplo da Figura 10, tamanho 3). Para criação dos novos descritores, a janela é colocada sobre cada ponto, e variáveis como média, valor máximo e valor mínimo são calculadas considerando todos os pontos pertencentes ao intervalo.

Figura 10: Método das janelas deslizantes



Fonte: autor

### 1.2.3. Aprendizado de máquina

O aprendizado de máquina é uma área de estudo da computação que consiste em fornecer a computadores a capacidade de aprender a realizar tarefas sem ser explicitamente programados para tanto, através da exposição a dados. Segundo Arthur Samuel, cientista da computação que cunhou o termo, a abordagem de aprendizado de máquina pode reduzir o esforço de programar detalhadamente computadores para resolução de alguns problemas (Samuel, 1959). O aprendizado pode ser classificado em três tipos, de acordo com a forma como ele ocorre:

aprendizado supervisionado, aprendizado não-supervisionado e aprendizado por reforço (Haykin, 2008).

No aprendizado supervisionado, são utilizados dados de entrada e os respectivos valores de saída para treinar sistemas. Durante o treinamento, os algoritmos buscam uma função que relacione as entradas com as saídas fornecidas, e que seja capaz de generalizar para novos casos. Cabe ressaltar que, no aprendizado supervisionado, é preciso conhecer os valores de saída no conjunto de dados de treinamento.

No aprendizado não-supervisionado, somente dados de entrada são utilizados. Neste caso, os algoritmos buscam agrupar os dados de acordo com suas características, visando a detecção de padrões nas informações.

No aprendizado por reforço, também são utilizados apenas dados de entrada. No entanto, existe um retorno sobre o sucesso do processo de aprendizagem, permitindo que o sistema treine com base na crítica sobre seu desempenho.

De modo geral, aplicações do aprendizado de máquina ao estudo de comportamento de animais empregam o paradigma de aprendizado supervisionado (Anderson e Perona, 2014). Nesta abordagem, um conjunto de vídeos é tipicamente dividido em uma parte para treino do algoritmo e uma parte para seu teste. No processo de aprendizagem, pesquisadores treinados rotulam manualmente vídeos de testes comportamentais (o conjunto de treino), indicando os instantes de tempo em que ocorre cada comportamento que se deseja estudar. Os algoritmos de aprendizado supervisionado utilizam estes rótulos, juntamente com parâmetros extraídos dos vídeos dos testes, para adquirir a habilidade de identificar os comportamentos.

Nas subseções a seguir, serão apresentados alguns dos algoritmos costumeiramente empregados em aplicações de análise comportamental em animais, bem como as métricas de avaliação usualmente utilizadas para aferir seu desempenho.

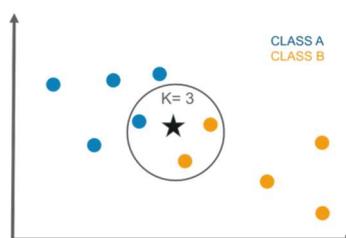
### 1.2.3.1. Algoritmos de classificação

Nesta seção serão apresentados alguns algoritmos de classificação em aprendizado de máquina, com seus princípios de funcionamento, vantagens e desvantagens.

O algoritmo *K Nearest Neighbors* (kNN) é um método baseado em vizinhança amplamente empregado em aprendizado de máquina devido à simplicidade de sua implementação e boa performance (Zhang et al., 2017). Algumas vantagens deste algoritmo incluem geralmente alta acurácia e insensibilidade a *outliers*. No entanto, trata-se de um método computacionalmente oneroso, que requer grandes quantidades de memória (Harrington, 2012).

Para realizar a classificação de uma nova instância, o algoritmo kNN compara os atributos deste ponto com os atributos de cada uma das instâncias do conjunto de dados de treino. A seguir, ele avalia os rótulos dos k pontos mais semelhantes a ele, sendo k um número inteiro definido pelo usuário, e realiza uma votação. À nova instância, será atribuída a classe mais comum entre os k vizinhos mais próximos. A Figura 11 exemplifica um caso de classificação binária utilizando k=3. Neste caso, 2 dos três vizinhos mais próximos do ponto a ser classificado (representado pela estrela) pertencem à classe B. O novo ponto será, portanto, atribuído a esta classe. No caso de comportamentos de animais, o algoritmo classificaria um novo ponto comparando-o com aqueles cujos parâmetros se assemelham a ele no conjunto de dados de teste. Instâncias em que o animal apresenta alta velocidade, por exemplo, tendem a estar mais próximas de pontos em que o animal está caminhando no conjunto de dados de teste do que de pontos em que o animal está imóvel.

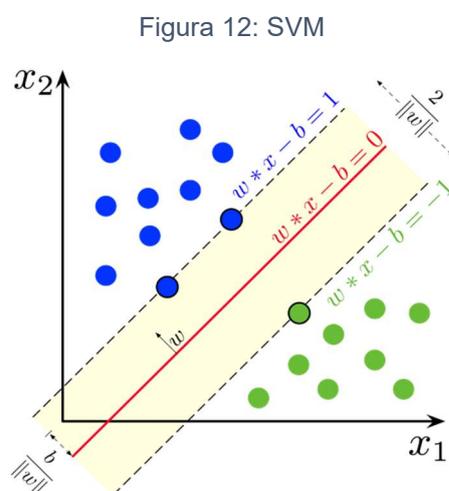
Figura 11: kNN



Fonte: Edureka, 2019

As máquinas de vetores de suporte (do inglês, *support vector machines*, SVM), operam buscando a linha de separação, ou hiperplano, que melhor divida duas classes em um conjunto de dados para treino. Para classificar novas instâncias, o algoritmo verifica de qual lado do hiperplano o ponto se encontra. Para criar a fronteira de decisão, as SVM trabalham mapeando um espaço vetorial de entrada em um espaço de dimensão mais elevada, onde as classes possam ser mais facilmente separadas. No novo espaço vetorial, é utilizado o princípio da minimização do risco estrutural (do inglês, *structural risk minimization*, SRM) para encontrar a fronteira, maximizando a separação entre as classes (Vapnik and Chervonekis, 1977).

As SVM tipicamente fornecem baixos erros de generalização, com custo computacional reduzido para classificação e resultados facilmente interpretáveis. No entanto, sua forma nativa prevê apenas classificação binária, e o processo de treinamento é custoso de um ponto de vista computacional (Harrington, 2012).



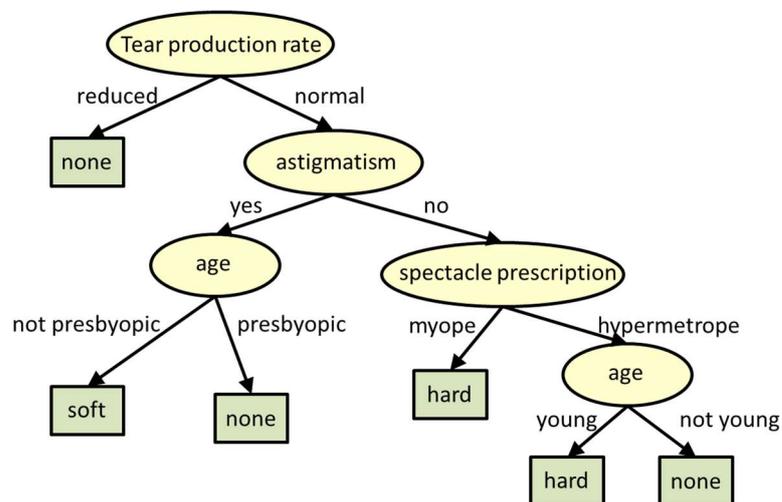
Fonte: Lahrmam, 2018

As árvores de decisão utilizam a entropia de Shannon, uma medida de organização de um conjunto de dados, para criar regras de classificação. Cada regra é criada de forma a dividir os dados da forma mais organizada possível. A Figura 13 mostra um exemplo de árvore. Neste caso, temos um classificador de diagnóstico de problemas de visão, indicando miopia, hipermetropia e presbiopia a partir de parâmetros como produção de lágrimas e idade. No caso de comportamento de

animais, as árvores utilizariam os parâmetros, como a velocidade, dos conjuntos de dados de teste para criar regras do tipo “se a velocidade for maior que determinado valor, então o animal está caminhando”.

As árvores de decisão possuem como principal vantagem sua alta interpretabilidade, visto que modelos criados assim podem ser facilmente representados graficamente, exibindo os critérios que foram utilizados no processo de classificação. Além disso, a classificação não tem alto custo computacional, e o método lida bem com atributos de baixa importância. No entanto, as árvores de decisão possuem alta variância, sendo muito sensíveis a variações nos dados de treino. Pequenas mudanças nos conjuntos de dados podem levar à criação de árvores muito diferentes. Além disso, este algoritmo é muito suscetível ao sobreajuste, se adaptando muito bem aos dados de treino mais possuindo pouca eficácia na classificação de novas instâncias (Harrington, 2012).

Figura 13: Árvores de decisão

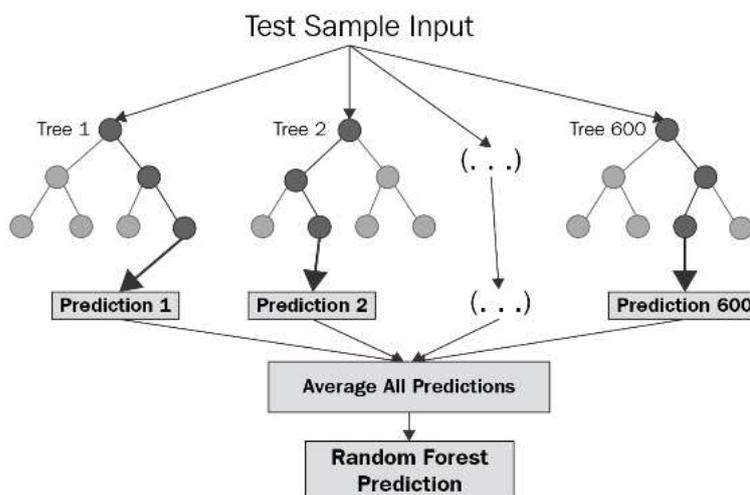


Fonte: Bhiksha, 2016

Visando usufruir das vantagens do método das árvores de decisão e contornar suas limitações, surgiu o método das florestas aleatórias (do inglês, *Random Forests*, RF) (Breiman, 2001). Trata-se de um método de *ensemble*, que combina o resultado de diversos modelos preditivos para realizar a classificação. Seu princípio de

funcionamento consiste em criar várias árvores de decisão onde o atributo de decisão é escolhido aleatoriamente. Em cada iteração, o algoritmo ignora parte aleatoriamente selecionada do conjunto de dados de treino, utilizando somente parte dele para criar uma árvore. O processo é repetido diversas vezes, de forma a construir um conjunto de árvores. Para classificar uma instância, ela é passada a todas as árvores do conjunto, que a analisam separadamente. O resultado final é obtido pela combinação dos resultados de cada árvore.

Figura 14: *Random forests*



Fonte: Chakure, 2019

Alguns algoritmos de classificação, como SVM, foram criados para solução de problemas de classificação binária, em que cada ocorrência pertence a uma de duas possíveis classes. Para aplicação a problemas de classificação multiclasse, em que uma instância pertence a uma de diversas classes, é preciso adotar estratégias de generalização dos métodos. Além disso, mesmo algoritmos que em sua forma nativa são adaptados para análise de problemas multiclasse podem apresentar melhor desempenho em problemas binários (Sánchez-Marño, 2010). Uma abordagem comumente empregada consiste em dividir problemas multiclasse em conjuntos de problemas de classificação binária, que são posteriormente combinados. Neste contexto, existem duas estratégias principais: *One versus All* (OvA) e *One versus One* (OvO).

Na abordagem OvA, um problema de classificação multiclasse contendo  $n$  classes é dividido em  $n$  classificadores binários. Cada classificador binário é treinado de forma a distinguir entre uma das classes e o restante. Na abordagem OvO, um problema de classificação com  $n$  classes é dividido em  $n(n-1)$  problemas de classificação binária, em que cada uma das classes é comparada com cada uma das outras individualmente.

### 1.2.3.2. Métricas de avaliação

Tendo em vista a grande variedade de aplicações das técnicas de aprendizado de máquina, existem também diversas formas de avaliar a eficácia de modelos. Cada aplicação, com suas peculiaridades, como o número de instâncias pertencentes em cada caso, bem como os resultados esperados em termos de falsos negativos ou falsos positivos definem a métrica mais adequada para utilização em uma aplicação. Nesta subseção, serão apresentadas as métricas mais comumente empregadas em problemas de classificação.

A matriz de confusão apresenta, em um problema de  $n$  classes, em forma de tabela de dimensão  $n \times n$ , os erros e acertos de um modelo de classificação, conforme mostrado no Quadro 1. Da matriz de confusão, podem ser derivados também os valores para cálculo de diversas outras métricas, como precisão e revocação.

Quadro 1: Matriz de confusão

		Classe prevista	
		A	B
Classe real	A	TP	FN
	B	FP	TN

No Quadro 1, é apresentada uma matriz de confusão de um problema de classificação binária, em que há uma classe positiva A e uma classe negativa B. Nas

linhas, temos as classes a que instâncias realmente pertencem, e, nas colunas, as classes previstas por um modelo. O elemento TP (do inglês *true positive*) corresponde ao número de instâncias da classe A previstas corretamente. O elemento FN (do inglês *false negative*) corresponde aos casos positivos que o modelo classifica como negativos. O elemento FP (do inglês *false positive*) corresponde aos casos negativos que o modelo classifica como positivos. O elemento TN (do inglês *true negative*) corresponde às instâncias da classe B previstas corretamente.

As matrizes de confusão podem ser empregadas também na avaliação de desempenho em problemas com mais de duas classes. Nestes casos, considera-se que uma classe é a positiva e as demais são negativas.

A acurácia computa o total de acertos realizados em uma predição, independentemente da classe. Ela é tomada dividindo o total de instâncias previstas corretamente pelo número total de instâncias em um conjunto de dados:

$$acurácia = \frac{TP + TN}{TP + TN + FP + FN} \quad (10)$$

Embora a acurácia seja uma métrica amplamente utilizada em aprendizado de máquina, ela possui eficácia limitada para descrever o desempenho de modelos utilizados em conjuntos de dados com classes não balanceadas, isto é, em que uma classe possui maior representatividade. Para ilustrar o problema, frequentemente utiliza-se o caso em que se deseja criar um modelo para detecção de transações fraudulentas em cartões de crédito. Nestes tipos de conjuntos de dados, apenas uma pequena parcela das transações, por exemplo, 1%, é fraudulenta. Um modelo que classifica todas as instâncias como sendo transações não fraudulentas possui acurácia de 99%, e, no entanto, não possui qualquer utilidade na detecção de fraude. Nestes casos, outras métricas são mais adequadas.

A precisão é definida como a razão entre o número de verdadeiros positivos e a soma dos verdadeiros positivos e os falsos positivos:

$$precisão = \frac{TP}{TP + FP} \quad (11)$$

Ela corresponde ao número de predições positivas que de fato pertencem à classe positiva.

A revocação é definida como a razão entre o número de instâncias positivas corretamente classificadas e a soma dos verdadeiros positivos com falsos negativos:

$$revocação = \frac{TP}{TP + FN} \quad (12)$$

Uma forma de combinar precisão e revocação é utilizando a medida F1 (ou F1-score), que corresponde à média harmônica entre as duas medidas:

$$F1 = \frac{2 * precisão * revocação}{precisão + revocação} \quad (13)$$

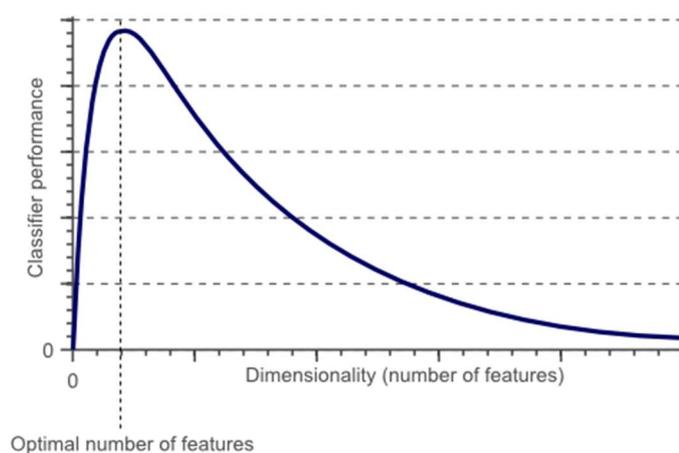
Para obter valores elevados de F1, é preciso que tanto a precisão quanto a revocação sejam altas. A medida F1 é adequada para avaliar o desempenho de modelos com classes não balanceadas.

### 1.2.3.3. Seleção de parâmetros

Problemas que possuem grande quantidade de parâmetros em relação ao número de exemplos costumam sofrer problemas de desempenho. A “maldição da dimensionalidade” é um problema comum em aprendizado de máquina, em que o

desempenho de classificadores tende a cair com o aumento do número de parâmetros utilizados acima de determinada quantidade ótima, conforme ilustrado na Figura 15. Sempre que possível, é preferível utilizar modelos simples, com número reduzido de descritores. Parâmetros redundantes ou irrelevantes devem, sempre que possível, ser removidos (Yu e Liu, 2004).

Figura 15: A maldição da dimensionalidade



Fonte: ODSC, 2019

Para evitar problemas relacionados com alta dimensionalidade em aprendizado de máquina, o treinamento de modelos é frequentemente antecedido por uma etapa de seleção de parâmetros, onde aqueles que não possuem validade preditiva para a variável desejada são eliminados.

Uma estratégia comumente empregada consiste em remover parâmetros que possuam alta correlação com outros. O coeficiente de correlação de Pearson entre duas variáveis pode ser usado como uma medida da dependência linear entre elas, sendo definido como (Mukaka, 2012):

$$p = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}} \quad (14)$$

Onde  $x_i$  e  $y_i$  são os  $i$ -ésimos valores de cada variável.

## 2. JUSTIFICATIVA

Em virtude das limitações dos métodos manuais de avaliação comportamental citados nas seções anteriores, bem como da relevância destes testes na pesquisa pré-clínica sobre o TEA e outros transtornos, torna-se essencial investigar métodos que permitam avaliar o comportamento dos animais com mais detalhes e precisão. Ademais, a possibilidade de usufruir de métodos como a optogenética e a farmacogenética para desvendar mecanismos neurais por trás da formação do comportamento dependem da capacidade de avaliar com precisão o comportamento animal.

### **3. OBJETIVOS**

#### **3.1. Objetivo geral**

Desenvolver um método computacional para identificação de comportamentos do etograma de ratos Wistar, de forma automatizada, a partir de vídeos do teste de campo aberto.

#### **3.2. Objetivos específicos**

- Definir um etograma relevante para o estudo de TEA e outros transtornos, com comportamentos de ratos Wistar no teste de campo aberto.
- Desenvolver um algoritmo que identifique de forma automatizada o contorno de roedores a partir de vídeos do teste de campo aberto.
- Extrair parâmetros geométricos do contorno dos roedores (área, centroide, eixos etc.) que possam fornecer indícios sobre o comportamento do animal.
- Investigar os melhores métodos de aprendizado de máquina supervisionado para realizar a classificação dos comportamentos relevantes a partir dos parâmetros extraídos.
- Rotular manualmente os estados comportamentais estudados em um conjunto de vídeos do teste de campo aberto.
- Utilizar os dados para treinar e validar o funcionamento de um modelo de classificação utilizando aprendizado de máquina.

#### 4. TRABALHO NA FORMA DE ARTIGO CIENTÍFICO

A ser submetido de acordo com as normas de publicação da revista *Behavior Research Methods*.

##### DEVELOPMENT OF A METHOD FOR WISTAR RAT BEHAVIOR CLASSIFICATION USING RANDOM FORESTS

Matheus Costa Schiaffino<sup>1,2,3</sup>, Guilherme Bauer Negrini<sup>1,2,3</sup>, Mellanie Fontes  
Dutra<sup>1,2,3</sup>, Carmem Gottfried<sup>1,2,3\*</sup>

<sup>1</sup>Laboratory of Neuroglial Plasticity, Institute of Health's Basics Science, Department of Biochemistry, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil.

<sup>2</sup>Department of Biochemistry, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil.

<sup>3</sup>Translational Research Group in Autism Spectrum Disorders (GETTEA), Institute of Health's Basics Science, Department of Biochemistry, Federal University of Rio Grande do Sul (UFRGS), Porto Alegre, Brazil.

**\*Corresponding author:**

Carmem Gottfried

cgottfried@ufrgs.br

## ABSTRACT

Despite the large importance of animal behavioral tests in the preclinical research on neuropsychiatric and neurodevelopmental disorders, such as autism spectrum disorder (ASD), few computational tools exist to quantitatively evaluate animal behavior. In addition, most of the few existing tools rely on additional software or expensive hardware such as depth cameras or instrumented cages to run. In this paper, we present a method for albino rat behavior assessment from open field test videos utilizing computer vision and supervised machine learning techniques. All steps of the method, from image processing and segmentation to the application of machine learning techniques are here described. We manually labeled 46,521 frames from 52 videos of the open field test, assigning each of the frames to one of seven classes: walking, sniffing, rearing, fore body grooming, hind body grooming, immobility and other/unsure. We utilized background subtraction, thresholding and morphological opening on a frame-by-frame basis to segment the images, separating the region corresponding to the animal from the test apparatus. We then extracted geometrical pose estimators from the segmented region and applied the sliding windows technique to them to account for the time dynamics of behavior. Finally, we utilized the extracted data, along with the manually scored labels, to train binary classifiers using the Random Forest algorithm. Our system was trained for classifying walking (F1-score = 91.85%), sniffing (F1-score = 85.39%), rearing (F1-score = 80.94%), fore body grooming (F1-score = 59.39%), hind body grooming (F1-score = 72.48%) and immobility (F1-score = 75.04%). The method we developed allows for reliable and rapid detection of the trained behaviors, therefore consisting of a valuable tool for preclinical research.

## KEYWORDS

Rat behavior, ethology, computer vision, supervised learning

## INTRODUCTION

Autism spectrum disorder (ASD) consists of a developmental disorder characterized by deficits in communication and social interaction and repetitive and stereotyped behaviors (American Psychiatry Association (APA), 2013). The criteria for diagnostics are exclusively based on behavioral assessment, and there are currently no validated biomarkers for this disorder. According to the last study performed by the Center for Disease Control and Prevention (CDC), ASD has an incidence of 1:59 children in the USA, being approximately 5 times more common in boys (1:42) than in girls (1:189) (Morbidity and Mortality Weekly Report (MMWR), 2014). During the last decade, the incidence of ASD has been rapidly growing, having approximately doubled between 2004 and 2014 (Summaries, 2007; Morbidity and Mortality Weekly Report (MMWR), 2014). In addition to the changes in diagnostic criteria, which are believed to be increasing the number of patients included in the spectrum (Fombonne et al., 2009), the exposition to environmental risk factors may be contributing to this increase in incidence (Bryson et al., 2003; Schneider e Przewłocki, 2004).

Although ASD etiology remains unknown, it is believed that it consists of an interaction between genetic and environmental factors (Sandin et al., 2014). Research indicates that the exposition to certain teratogens such as valproic acid (VPA) and thalidomide during the first trimester of pregnancy may be related to ASD triggering (Rouillet et al., 2013; Smith e Brown, 2014; Imai et al., 2014).

Because ASD is characterized exclusively by behavioral alterations, preclinical research on this disorder frequently involves the utilization of animal behavioral tests to evaluate social and stereotypical behavior. Tests which are commonly employed to evaluate social behavior include the reciprocal interaction test and the three chamber test (Chang et al., 2018). For stereotypical behavior assessment, some widely employed tests include the open field test (OFT) (Gould et al., 2009), the marble burying test (Njung'e and Handley, 1991) and the T maze spontaneous alternation test (Deacon e Rawlins, 2006).

The open field test was first employed in 1934 to evaluate emotionality in rats by measuring the animals' urination and defecation (Hall, 1934), and since then has been used to observe several other aspects of rat behavior. It is performed in a generally square apparatus with surrounding walls, preventing escape, where a rat is put to behave without restraints. It is one of the most employed tests in the research of ASD and other disorders, being commonly used to evaluate the rat exploratory behavior and general activity. The monitored variables usually include the total distance traveled, the time spent by the animal along the walls and in the center of the apparatus, and the number of crossings into the center. Other parameters that can be assessed by the OFT include the number, duration and latency to behaviors such as walking, sniffing, rearing, grooming and freezing (Gould et al., 2009). Walking behavior is defined as relatively fast locomotor activity without sniffing exploration (Choleris et al., 2001). It is commonly assessed in the research of disorders such as anxiety and Parkinson's, in which the animals' locomotion is usually affected (Kang e Meredith, 2006; Ramos, 2008). Sniffing is defined as exploratory activity where the animal sniffs the apparatus' floor, its walls or the air (Choleris et al., 2001). Rearing behavior occurs when the animal lifts its fore paws from the ground, touching the apparatus walls or

not. It is a context dependent behavior, occurring less frequently in adverse situations and being less common in females than in males (Sturman et al., 2018). Grooming occurs when the animal performs self-cleaning licking or rubbing itself. This behavior has been shown to be very sensitive to stressors and several experimental factors (Kalueff et al., 2007), being usually altered in animal models of disorders such as anxiety, ASD, depression, OCD, and Tourette's (McFarlane et al., 2007). Freezing is characterized by complete lack of body movements, often occurring as a response to fear stimuli (Brandão et al., 2008). In the OFT, excessive freezing activity is usually related to anxiety (Díaz-Morán et al., 2014).

The analysis of behaviors such as the aforementioned is usually performed manually by trained researchers, who score variables such as number, duration and latency of relevant behaviors by hand. In addition to being slow and very labor-intensive, this method tends to provide subjective and hardly reproducible results. The interdisciplinary field of computational ethology emerged from the pursuit to create methods to accurately measure, quantify and describe animal behavior, involving a cooperation between biological sciences, mathematics, engineering and computer sciences (Anderson and Perone, 2014).

The evaluation of some variables in the OFT, such as total distance moved, tend to be relatively simple from a computational point of view, and several software solutions to perform this kind of analysis currently exist. Examples include commercial software such as Anymaze (Stoelting Co., 2019) and EthoVision (Noldus, 2020), as well as some open source tools such as ezTrack (Pennington et al., 2019).

The evaluation of complex behaviors, however, is much more complicated, and there are currently few computational tools available for identifying them from open field test

videos. This is due to the difficulty to explicitly define criteria for behavior identification using traditional programming techniques. This limitation can nevertheless be overcome with the application of machine learning, which is defined as a technique that provides computers the ability to automatically learn and improve from experience through exposition to data, without being explicitly programmed (Samuel, 1959).

Some successful computational ethology systems employing machine learning techniques have been published during the past decade. Some examples include JAABA, which analyses drosophila and rat behavior based on the output of tracking software such as Ctrax (Kabra et al., 2013), the methods proposed by Hong and de Chaumont, which analyze depth sensing imaging to evaluate rat behavior (Hong et al., 2015; De Chaumont, 2019), and the method proposed by Han to evaluate hydra behavior (Han et a., 2018). These systems, however, require additional software or expensive hardware such as depth cameras to run.

Our aim was to develop a reliable, cost-effective, automated method for complex albino rat behavior detection from videos of the open field test. In this paper, we present an end-to-end implementation which includes frame-by-frame segmentation of the videos, animal segmentation from the background in each image, feature extraction from the segmented images, feature creation by application of sliding windows and training of machine learning classifiers based on the Random Forest (RF) algorithm for walking, sniffing, rearing, fore body grooming, hind body grooming and immobility detection.

## **MATERIALS AND METHODS**

### **Behavioral tests setup and recording**

We used recordings of projects that were approved by the Ethics Committee of both Federal University of Rio Grande do Sul (CEUA-UFRGS 31921) and Clinical Hospital of Porto Alegre (HCPA/FIPE 2014-0367). The Wistar rats utilized were handled according to the National Council for the Control of Animal Experimentation (CONCEA) of Brazil. The open field tests were performed in a square 50 cm by 50 cm apparatus, and the videos were recorded on 640x480 resolution with a constant frame rate of 30 FPS. The duration of each test was 10 minutes.

### **System overview**

The method we propose consists of five basic steps, which will be detailed in the next sections:

- (a) Image segmentation to localize the animal in the videos
- (b) Feature extraction from the segmented images
- (c) Application of sliding windows to the extracted features
- (d) Removal of highly correlated features
- (e) Training of classifiers with supervised learning

### **Animal detection**

The first step of our method consists of defining what part of the image corresponds to the animal and what part of the image corresponds to the background in each frame

of the analyzed videos. For this, we utilized Matlab Computer Vision Toolbox. We segmented the images by performing static background subtraction, followed by thresholding using Otsu's method (Otsu, 1979) and morphological opening with a disk-shaped structuring element with a radius of 5 pixels. We only kept objects with area between 200 and 700 pixels. The background for each video was a frame recorded before the animal was introduced in the apparatus. The results were validated through visual inspection by a trained researcher.

### **Feature extraction**

We extracted 16 parameters from each segmented image. Some parameters were extracted directly from the animals' shape in the images, while others were extracted from fitted ellipses which had the same second moments as the animals' shapes. The extracted parameters were 1) area, 2) convex area, 3) solidity, 4) major axis length, 5) minor axis length, 6) eccentricity, 7) orientation, 8) distance to the closest apparatus wall, 9) perimeter, centroid 10) x and 11) y coordinates, 12) area beyond floor borders, 13) linear velocity calculated as the centroid's position variation between two consecutive frames, 14) linear velocity calculated as the centroid's position variation between three consecutive frames, 15) angular velocity calculated as the orientation's variation between two consecutive frames, and 16) angular velocity calculated as the orientation's variation between three consecutive frames.

### **Feature processing**

To account for the time dynamics of behavior, we calculated new features by using the sliding windows technique, commonly used in the analysis of time series and action

recognition (Banos et al., 2014). We used windows of sizes 3, 7 and 11. The features of each frame were computed by placing the window centered in the frame and calculating the maximal, minimal, mean, standard deviation and absolute difference between the first and last values in each window.

After applying the sliding windows technique, we calculated Pearson's correlation between each pair of features, defined as

$$p = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}$$

where  $x_i$  and  $y_i$  are the  $i$ -eth value of each feature. We removed one of the features of each pair that had  $p > 0.9$ . Features that were created with larger window sizes were removed first.

### **Manual annotation of behaviors**

To assist with manual annotation of behaviors, we developed a Matlab application. Each frame was manually assigned to one of seven classes: walking, sniffing, rearing, fore body grooming, hind body grooming, immobility and other/unsure. Walking was defined as locomotion activity without sniffing exploration; sniffing was defined as the animal sniffing the air, the apparatus' walls or floor; rearing was defined as the animal raising its front paws while keeping both hind paws on the floor; fore body grooming was defined as grooming of the head and front paws; hind body grooming was defined as grooming of the body, hind paws, tail and genitals; immobility was defined as the animal maintaining a fixed posture. All frames where the observer was not sure of the behavior were scored as other/unsure.

## Supervised machine learning

The data sets were split into training sets, with approximately 80% of the videos, and test sets, with approximately 20% of the videos. We only used frames where the researcher who performed the manual scoring was sure of the behavior. Since the behaviors we studied occur with different frequencies and durations, the datasets of all the classifiers were unbalanced. To avoid bias toward the most represented classes, we performed random undersampling of the majority class in each of the binary training sets. The test sets were left unbalanced.

The manually annotated frames were used to test several different classifiers for performance comparison. The Random Forests algorithm gave us the best results and was chosen for creating the models we present in this article. We trained six binary classifiers with this algorithm, one for each of the aforementioned behaviors, using Matlab Statistics and Machine Learning toolbox.

## Evaluation metrics

We measured each classifier's accuracy, precision, recall and F1-score in order to evaluate the models' performance.

Accuracy was calculated as

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

Precision was calculated as

$$precision = \frac{TP}{TP + FP}$$

Recall was calculated as

$$recall = \frac{TP}{TP + FN}$$

F1-score was calculated as the harmonic mean of precision and recall

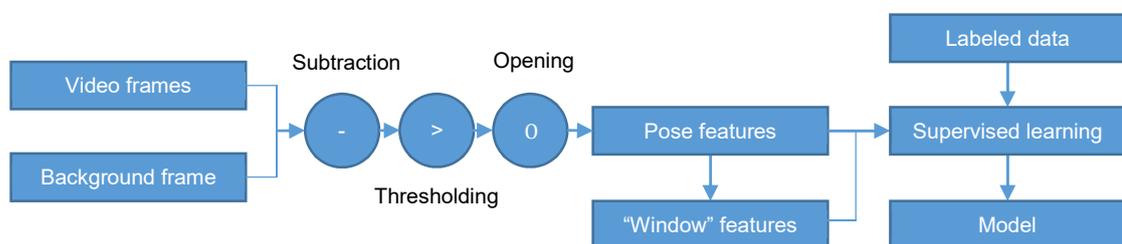
$$F1 = \frac{2 * precision * recall}{precision + recall}$$

Where TP is the number of true positives, TN is the number of true negatives, FP is the number of false positives and FN is the number of false negatives of each binary classifier. For unbalanced testing sets such as ours, F1-score gives the most reliable performance indication, as it considers both precision and recall.

## RESULTS

### System overview

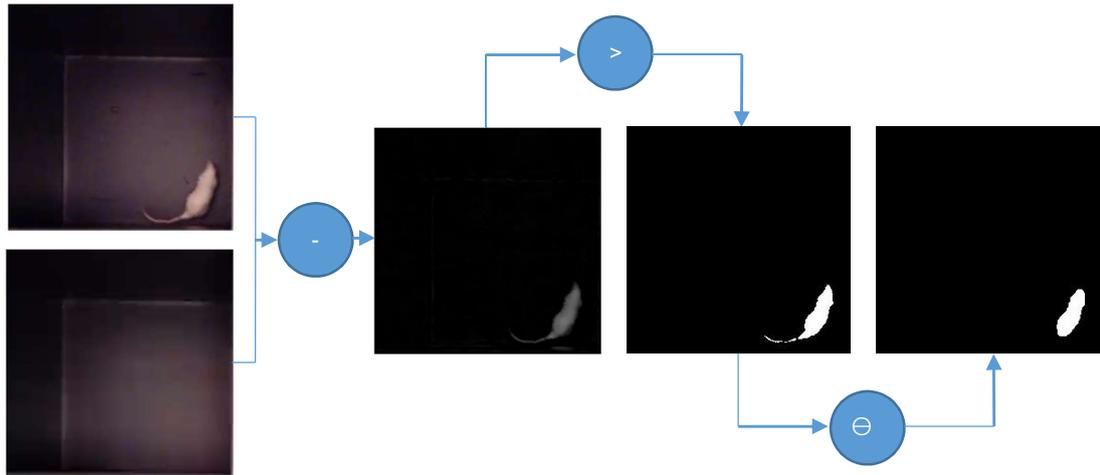
The system we developed works according to the workflow presented in Figure 1.



**Fig. 1** Workflow representing the major steps of image processing, feature extraction and model training of our system.

## Animal detection

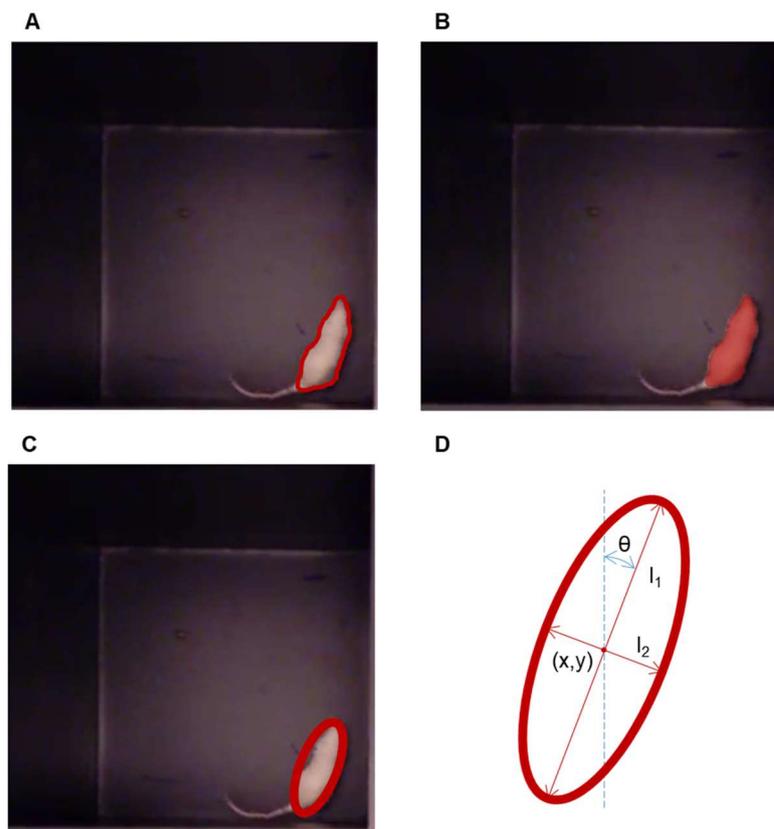
Figure 2 shows in detail the steps of background subtraction, thresholding with Otsu's method and morphological opening.



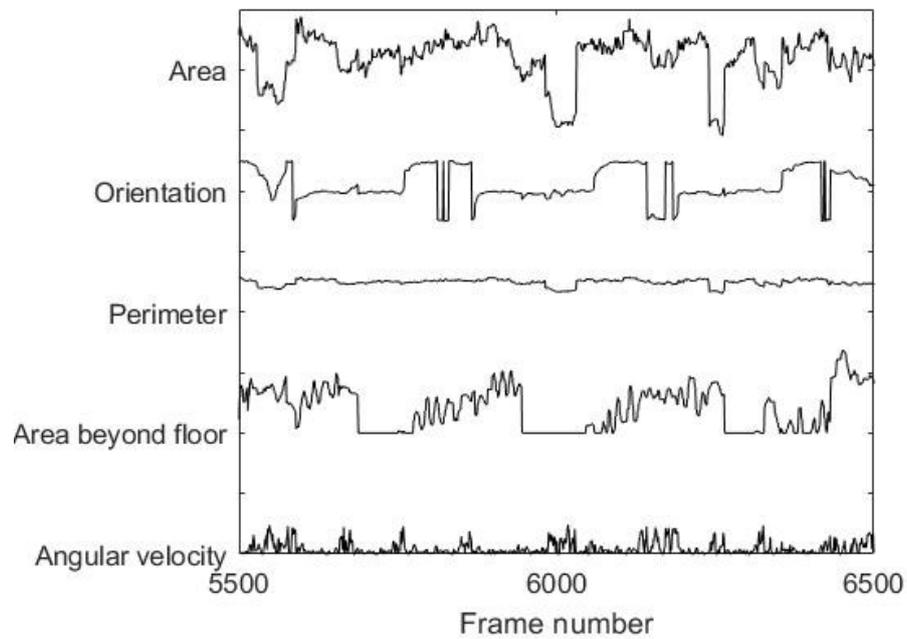
**Fig. 2** Summary of processing steps to segment the images. We performed background subtraction followed by thresholding using Otsu's method and morphological opening.

## Feature extraction

Our algorithm was capable of processing the images and extracting the parameters with an average rate between 100 and 150 frames per second. Figure 3 shows examples of parameters extracted from an animal's shape (A and B, respectively perimeter and area) and from a fitted ellipse (C). Figure 3-D shows parameters which can be extracted from a fitted ellipse, such as inclination, major and minor axis lengths. Figure 4 shows some of the extracted features in function of time.



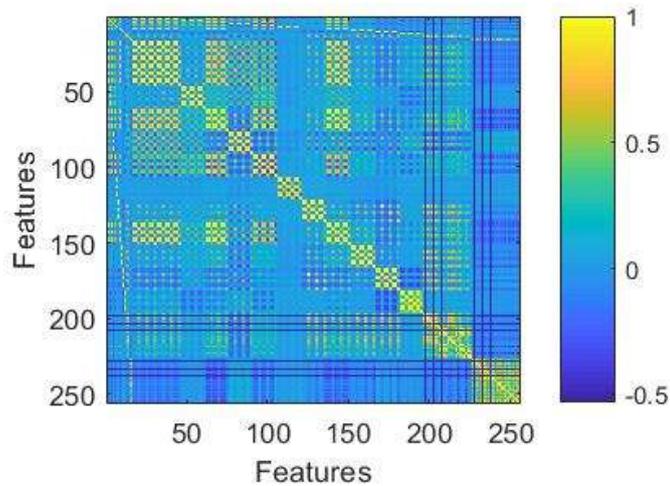
**Fig. 3** Examples of pose estimators extracted from the segmented images. Some features were extracted directly from the animal's contour, such as perimeter (A) and area (B). Others were extracted from a fitted ellipse (C), such as inclination, major and minor axis length (D).



**Fig. 4** Examples of extracted parameters in function of the frame number.

### Feature processing

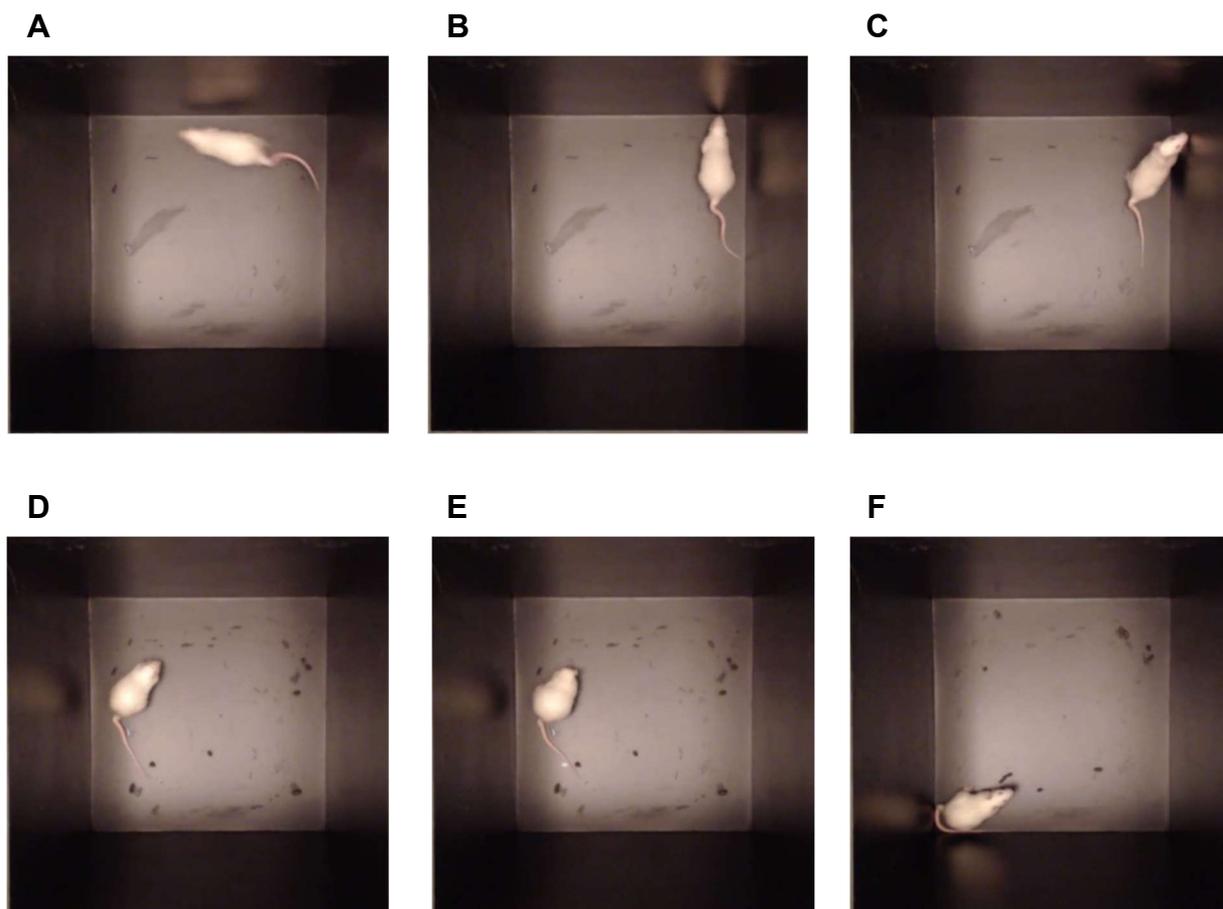
With the application of the sliding windows technique, we created 240 features in addition to the 16 extracted directly from the images. We calculated Pearson's correlation between each pair of variables, obtaining the results presented in the correlation matrix of Figure 5. By excluding one feature of each pair that had  $p > 0.9$ , our dataset was left with 94 features.



**Fig. 5** Correlation plot of the 256 originally extracted features.

### Manual annotation of behaviors

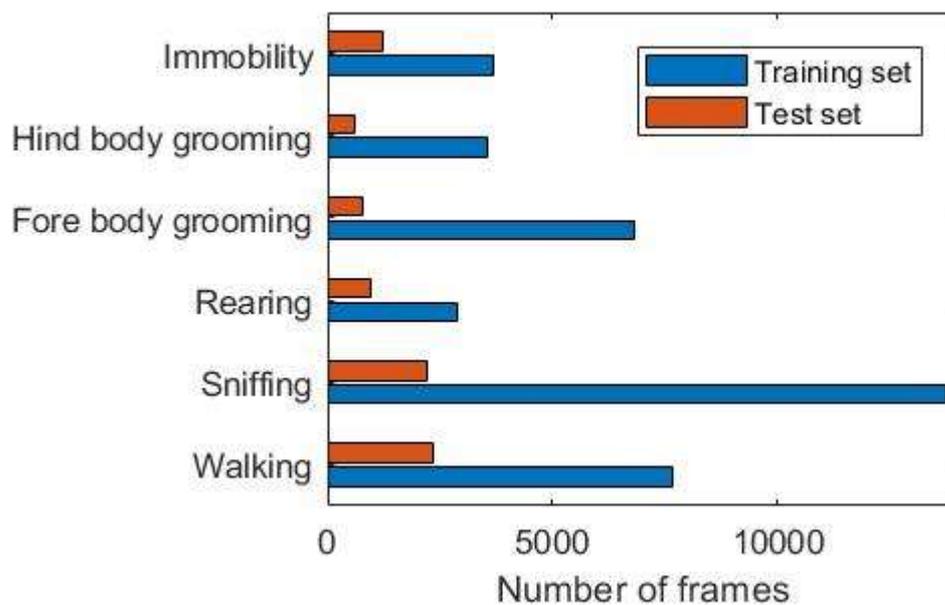
A total of 46,521 frames with clearly identifiable behaviors, belonging to 52 different videos, were manually scored on a frame-by-frame basis by a trained observer. The observer was able to label approximately 500 frames with clearly identifiable behaviors per hour of work. Walking was labeled in 36 videos; sniffing was labeled in 36 videos; rearing was labeled in 32 videos; fore body grooming was labeled in 31 videos; hind body grooming was labeled in 17 videos; immobility was labeled in 15 videos. The number of videos and frames we used in the training and test sets are shown in Table 1. Figure 6 presents examples of an animal performing each of the six studied behaviors. The number of frames where the animal performed each behavior is shown in Figure 7.

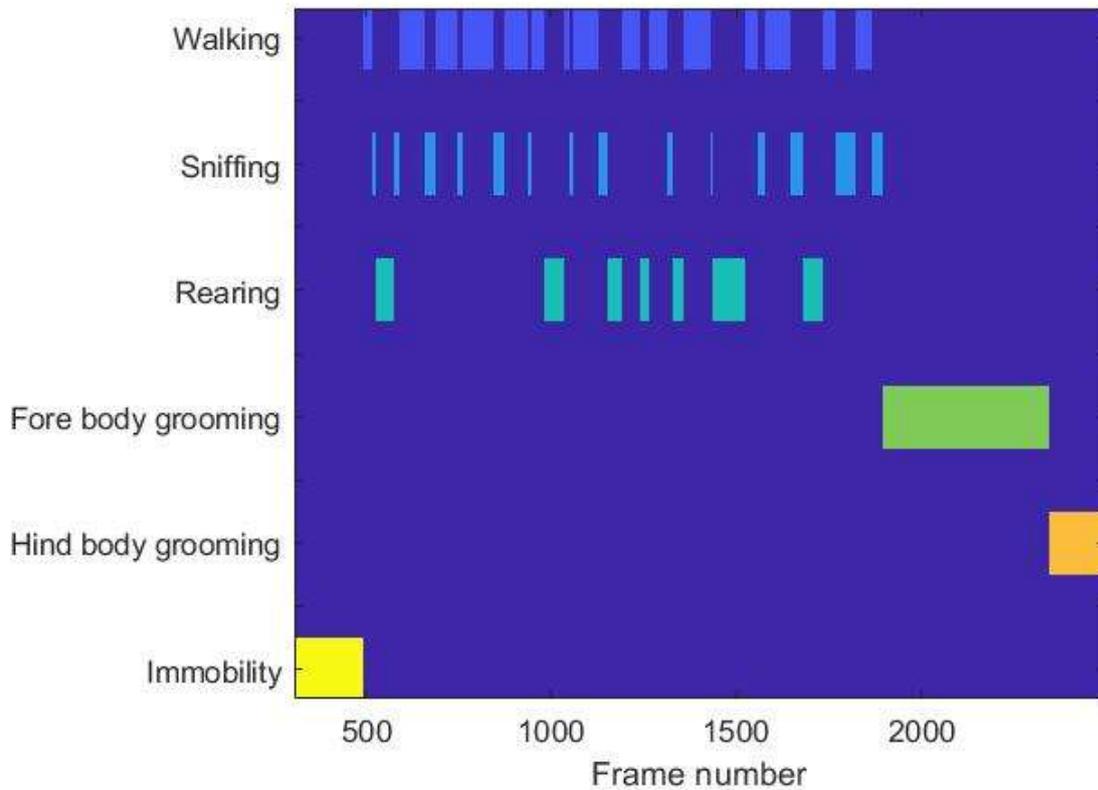


**Fig. 6** Six behaviors we used to train our system. (A) Walking. (B) Sniffing. (C) Rearing. (D) Fore body grooming. (E) Hind body grooming. (F) Immobility.

**Table 1** Training and test set size for each behavior.

Behavior	Number of training videos	Number of training frames	Number of test videos	Number of test frames
Walking	29	15,238	7	11,037
Sniffing	29	23,296	7	4,577
Rearing	26	5,742	6	7,390
Fore body grooming	25	11,128	6	6,046
Hind body grooming	13	6,670	4	4,577
Immobility	12	6,382	3	5,964

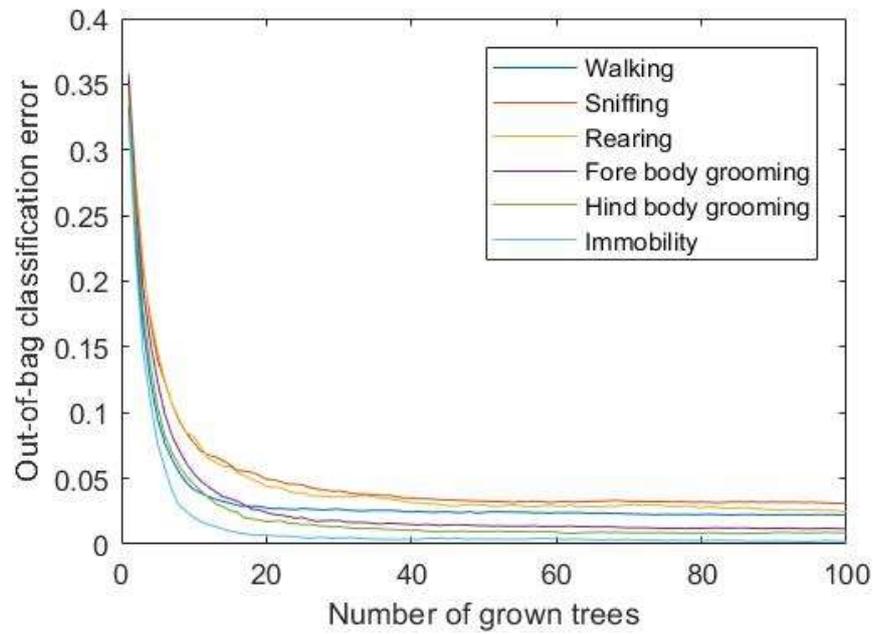
**Fig. 7** Instances of each behavior in the training and test sets.



**Fig. 8** Raster plot showing behaviors.

### Supervised machine learning

We grew 100 classification trees in each of our classifiers, which was beyond the number where error rate plateaued, as indicated by out-of-bag classification error (Figure 9). The results obtained utilizing the model on the test set are presented in Table 2. The time it took for each classifier to run on the test sets is presented in Table 3.



**Fig. 9** Out-of-bag classification error in function of the number of grown trees.

**Table 2** Accuracy, precision, recall and F1-score for each behavior

Behavior	Accuracy (%)	Precision (%)	Recall (%)	F1-score
Walking	96.37	87.83	96.25	91.85
Sniffing	88.14	80.76	90.58	85.39
Rearing	94.26	70.15	95.64	80.94
Fore body grooming	90.03	62.38	56.68	59.39
Hind body grooming	93.38	80.28	66.06	72.48
Immobility	91.40	94.49	62.23	75.04

**Table 3** Time for each analysis

Behavior	Number of frames	Time (s)	Time/1000 frames (s)
Walking	11,037	1.070933	0.097031
Sniffing	4,577	0.777661	0.169906
Rearing	7,390	0.786200	1.063870
Fore body grooming	6,046	0.676188	0.111840
Hind body grooming	4,577	0.525114	0.114728
Immobility	5,964	0.580113	0.097269

## DISCUSSION

The emerging field of computational ethology presents the perspective of great progress in neuroscience, by allowing for more robust, time and cost-efficient analysis of animal behavior. This could bring benefits to preclinical research on disorders such as ASD, which is characterized mainly by behavioral alterations. Screening for drugs to treat and searching for potential risk factors for these disorders, for instance, is directly dependent on behavioral studies with animal models. Furthermore, it's important to consider that a great deal of progress has been made during the past years allowing for researchers to map, monitor and manipulate the activity of specific neuron populations. These advances afford the ability to establish causal relations between the activity of specific neural circuits and behavior. However, the exploration of these new possibilities depends on the ability to quantitatively assess behavior (Anderson and Perone, 2014).

In this work, we present a method for rat behavior assessment based on the application of the Random Forest machine learning algorithm to features extracted from open field test videos and manually scored labels. Contrary to most state-of-the-art applications utilizing machine learning for animal behavior evaluation, our method doesn't require additional software or any hardware other than a standard video camera to operate. This makes it a cost-effective alternative to manual behavior analysis.

Our method was trained for classifying walking (F1-score = 91.85%), sniffing (F1-score = 85.39%), rearing (F1-score = 80.94%), fore body grooming (F1-score = 59.39%), hind body grooming (F1-score = 72.48%) and immobility (F1-score = 75.04%). Walking, sniffing and rearing were classified with high recall (96.25%, 90.58% and 95.64%, respectively) and slightly lower precision (87.83%, 80.76% and 70.15%), indicating more false positives than false negatives in these cases. Hind body grooming and immobility had high precision (80.28% and 94.49%) and lower recall (66.06% and 62.23%), indicating higher numbers of false negatives. Fore body grooming had similar values of precision (62.38%) and recall (56.68%).

In addition to the six classes we worked with in this project, we believe the method here proposed could be used to classify many other behaviors from the wide repertoire exhibited by Wistar rats, provided we extract representative features of these behaviors from the videos. The method could also be adapted for use in the evaluation of memory, social behavior, and many others. In memory tests such as object-location memory (OLM) and object recognition test (ORT), as well as in the three chamber sociability test, where researchers aim to evaluate the time that the animal passes interacting with a given object. We believe using supervised machine learning with parameters such as distance to the object, orientation and velocity could provide much more reliable results than simply computing the time the animal spends in proximity to

the object, as most systems currently do. In the reciprocal interaction test, parameters such as distance between two animals, its velocities and intersection area could be used to train a classifier for identifying different kinds of social interaction, although the task of identifying each individual in this kind of test tends to be a complicating factor for computational applications.

There are a few different strategies we believe could be employed in order to improve our model's performance. The first one is to increase the size of our data sets, since some of the behaviors we worked with were very underrepresented in the data. Fore body grooming, for instance, occurred in 6,670 frames in the training set, corresponding to less than 10% of the total number of frames we used. The random undersampling of the majority class we performed to balance our training set could be a source of imprecision, since the samples used to train the models could possibly not be describing the entire sets accurately. Although augmentation of the data set could be carried out through oversampling processes such as Synthetic Minority Oversampling Technique (SMOTE), we believe the best way of improving our models' performances would be to increase the amount of human-labeled data for training. This could be achieved by encouraging researchers of other institutions to contribute to our method's development by feeding it more training data. We also believe that other machine learning techniques, such as direct frame-by-frame analysis of images by artificial neural networks, could yield good results, especially in the classification of behaviors that are identifiable mostly through pose evaluation, such as fore and hind body grooming.

In terms of time consumption, our algorithm took around 10 seconds to extract the parameters from 1000 frames, while our classifiers took between 0.097 and 1.064 seconds to analyze them, by far outperforming trained human annotators realizing

manual analysis, who take on average two hours to evaluate the same amount of frames. Furthermore, computational analysis can operate nonstop for days at a time, not being prone to errors due to distraction and tiredness as humans are. This could allow for high throughput analysis of complex behavior, which is currently not possible with available methods. Higher rate of behavior classification could permit other analysis, such as T-patterns, to be realized. These analysis, as well as the application of other machine learning techniques, particularly unsupervised learning, could help unravel new characteristics of behavior which human observers are not capable of perceiving. In the long term, advances in computational ethology could provide the ability to measure and describe animal behavior in its totality.

#### **AUTHOR NOTE**

This project was supported by the National Council of Technologic and Scientific Development of Brazil (CNPq), INCT-NIM, AWARD Institute and HCPA/FIPE.

#### **REFERENCES**

- American Psychiatry Association (APA), 2013. Diagnostic and Statistical Manual of Mental Disorders, 5th edition, American Psychiatric Publishing, Inc, Washington, DC.
- Anderson, D. J., & Perona, P. (2014). Toward a Science of Computational Ethology. *Neuron*, 84(1), 18–31. doi:10.1016/j.neuron.2014.09.005

- Banos, O., Galvez, J.-M., Damas, M., Pomares, H., & Rojas, I. (2014). Window Size Impact in Human Activity Recognition. *Sensors*, 14(4), 6474–6499. doi:10.3390/s140406474
- Brandão, M. L., Zanoveli, J. M., Ruiz-Martinez, R. C., Oliveira, L. C., & Landeira-Fernandez, J. (2008). Different patterns of freezing behavior organized in the periaqueductal gray of rats: Association with different types of anxiety. *Behavioural Brain Research*, 188(1), 1–13. doi:10.1016/j.bbr.2007.10.018
- Bryson, S.E., Rogers, S.J., Fombonne, E., 2003. Autism spectrum disorders: early detection, intervention, education and psychopharmacological management. *Can J Psychiatry* 48, 506-516. doi: 10.1177/070674370304800802
- Chang, Y.-C., Cole, T. B., & Costa, L. G. (2017). Behavioral Phenotyping for Autism Spectrum Disorders in Mice. *Current Protocols in Toxicology*, 11.22.1–11.22.21. doi:10.1002/cptx.19
- Choleris, E., Thomas, A.W., Kavaliers, M., Prato, F.S. (2001). A detailed ethological analysis of the mouse open field test: effects of diazepam, chlordiazepoxide and an extremely low frequency pulsed magnetic field. doi: 10.1016/s0149-7634(01)00011-2
- Deacon, R. M. J., & Rawlins, J. N. P. (2006). T-maze alternation in the rodent. *Nature Protocols*, 1(1), 7–12. doi: 10.1038/nprot.2006.2
- De Chaumont, F., Ey, E., Torquet, N., Lagache, T., Dallongeville, S., Imbert, A., ... Olivo-Marin, J.-C. (2019). Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nature Biomedical Engineering*. doi:10.1038/s41551-019-0396-1

- Díaz-Morán, S., Estanislau, C., Cañete, T., Blázquez, G., Ráez, A., Tobeña, A., Fernández-Teruel, A. (2014). Relationships of open-field behaviour with anxiety in the elevated zero-maze test: Focus on freezing and grooming. doi: 10.4236/wjns.2014.41001
- Fombonne, E., Quirke, S., Hagen, A., 2009. Prevalence and interpretation of recent trends in rates of pervasive developmental disorders. *Mcgill J Med* 12, 73.
- Gould, T. D., Dao, D. T., & Kovacsics, C. E. (2009). The Open Field Test. *Neuromethods*, 1–20. doi:10.1007/978-1-60761-303-9\_1.
- Hall, C.S. (1934). Emotional behavior in the rat. doi: 10.1037/h0071444
- Han, S., Taralova, E., Dupre, C., Yuste, R. (2018). Comprehensive machine learning analysis of Hydra behavior reveals a stable basal behavioral repertoire. doi: 10.7554/eLife.32605
- Hong, W., Kennedy, A., Burgos-Artizzub, X.P., Zelikowsky, M., Navonneb, S.G., Perona, P., Anderson, D.J. (2015). Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning. doi: 10.1073/pnas.1515982112
- Imai, K., Iida, T., Yamamoto, M., Komatsu, K., Nukui, Y., Yoshizawa, A., 2014. Psychological and mental health problems in patients with thalidomide embryopathy in Japan. *Psychiatry Clin. Neurosci.* 68, 479-486. doi:10.1111/pcn.12152
- Kabra, M., Robie, A.A., Rivera-Alba, M., Branson, S., Branson, K. (2013). JAABA: interactive machine learning for automatic annotation of animal behavior. doi: 10.1038/nmeth.2281

- Kalueff, A.V., Aldridge, J.W., Laporte, J.L., Murphy, D.L., Tuohimaa, P. (2007). Analyzing grooming microstructure in neurobehavioral experiments. doi: 10.1038/nprot.2007.367
- Kang, U.J., Meredith, G.E. (2006). Behavioral models of Parkinson's disease in rodents: A new look at an old problem. doi: 10.1002/mds.21010
- McFarlane, H. G., Kusek, G. K., Yang, M., Phoenix, J. L., Bolivar, V. J., Crawley J.N. (2007). Autism-like behavioral phenotypes in BTBR T1tf/J mice. doi: 10.1111/j.1601-183X.2007.00330.x
- Morbidity and Mortality Weekly Report (MMWR), 2014. Prevalence of Autism Spectrum Disorder among Children Aged 8 Years – Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010, Vol. 63(2). ed. Surveillance Summaries, Washington, DC
- Njung'e, K., & Handley, S. L. (1991). Evaluation of marble-burying behavior as a model of anxiety. *Pharmacology Biochemistry and Behavior*, 38(1), 63–67. doi:10.1016/0091-3057(91)90590-x
- Noldus, January 20, 2020. EthoVisionXT. Retrieved from <https://www.noldus.com/ethovision-xt>
- Otsu, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. doi:10.1109/tsmc.1979.4310076
- Pennington, Z.T., Dong, Z., Feng, Y., Vetere, L.M., Page-Harley, L., Shuman, T., Denise J.C. (2019). eztrack: An open-source video analysis pipeline for the investigation of animal behavior. doi: 10.1038/s41598-019-56408-9

- Ramos, A. (2008). Animal models of anxiety: do I need multiple tests? doi: 10.1016/j.tips.2008.07.005
- Samuel, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229. doi:10.1147/rd.33.0210
- Sandin, S., Lichtenstein, P., Kujaj-Halkola, R., Larsson, H., Hultman, C.M., Reichenberg, A., 2014. The familial risk of autism. *JAMA* 311, 1770-7. doi: 10.1001/jama.2014.4144
- Schneider, T., & Przewłocki, R. (2004). Behavioral Alterations in Rats Prenatally Exposed to Valproic Acid: Animal Model of Autism. *Neuropsychopharmacology*, 30(1), 80–89. doi:10.1038/sj.npp.1300518.
- Smith, V., Brown, N., 2014. Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism. *Arch. Dis. Child. Educ. Pract. Ed.* 2013-2015. doi:10.1136/archdischild-2013-305636
- Stoelting Co., 2019. ANY-maze: Behavioral tracking software. Retrieved from [www.anymaze.co.uk](http://www.anymaze.co.uk)
- Sturman, O., Germainband, P.L., Bohacek, J. (2018). Exploratory rearing: a context- and stress-sensitive behavior recorded in the open-field test. doi: 10.1080/10253890.2018.1438405
- Summaries, S., 2007. Prevalence of autism spectrum disorders – autism and developmental disabilities monitoring network, six sites, United States, 2000. *MMWR. Sueveill. Summ. Morb. Mortal. Wkly. report. Surveill. Summ./ CDC* 56, 1-11

Roullet, F.I., Lai, J.K., Foster, J.A., 2013. In utero exposure to valproic acid and autism – a current review of clinical cases and animal studies. *Neurotoxicol Teratol* 36, 47-56. doi: S0892-0362(13)00005-6 [pii] 10.1016/j.ntt.2013.01.004

## 5. CONCLUSÕES

O presente trabalho se soma a alguns estudos que vêm sendo desenvolvidos nos últimos anos com o intuito de refinar a análise comportamental em animais, apresentando um método rápido e de baixo custo para classificação de comportamentos de ratos Wistar no teste de campo aberto. O sistema proposto foi treinado para identificar os comportamentos de caminhar (F1 = 91.85%), cheirar (F1 = 85.39%), *rearing* (F1 = 80.94%), autolimpeza dianteira (F1 = 59.39%), autolimpeza traseira (F1 = 72.48%) e imobilidade (F1 = 75.04%) através da análise de testes filmados com uma câmera filmadora padrão, em resolução 640x480.

O método desenvolvido neste projeto constitui um avanço considerável em relação à forma como são conduzidos testes tradicionalmente, visto que permite análises com maior reprodutibilidade, com necessidade significativamente reduzida de recursos humanos e em menor tempo. Ademais, a utilização deste método permite o monitoramento de uma quantidade maior de estados comportamentais: enquanto a análise realizada manualmente tipicamente pode ser empregada para avaliar um comportamento por vez, como a autolimpeza, a análise automatizada permite que se analise todo um etograma. Além dos comportamentos estudados neste projeto piloto, muitos outros poderiam ser avaliados empregando técnicas semelhantes. A implementação aqui apresentada poderia facilmente ser expandida para analisar, por exemplo, testes de memória, como memória de localização de objeto (do inglês *object-location-memory*, OLM) e teste de reconhecimento de objeto (do inglês *object recognition test*, ORT). Com refinamento da análise, a implementação poderia ser eficaz mesmo na identificação de comportamentos mais sutis, como as etapas da autolimpeza, e também na avaliação de interação social no teste de sociabilidade recíproca.

Para refinamento do desempenho do sistema proposto, seria imprescindível aumentar o conjunto de dados utilizados em seu treino. Como a tarefa de rotulagem manual de comportamentos é lenta e laboriosa, seria importante contar com a colaboração de pesquisadores que se disponham a contribuir com dados para desenvolvimento de um banco para treinamento do modelo. Além do aumento do volume de dados, acreditamos que outras estratégias dentro do aprendizado de

máquina podem se mostrar eficazes, como a utilização de redes neurais artificiais para avaliar imagens antes de qualquer tratamento.

Métodos como o aqui proposto podem ser empregados na pesquisa pré-clínica sobre causas e tratamentos de transtornos como o TEA, bem como no estudo das bases neurais do comportamento. Especialmente se empregados em conjunto com outras técnicas de análise, utilizando aprendizado de máquina não supervisionado, é possível que futuros projetos na área de etologia computacional forneçam instrumentos para novos patamares em termos de medida, descrição e compreensão do comportamento.

## 6. PERSPECTIVAS

A conclusão do presente estudo fornece uma série de perspectivas para desenvolvimento futuro, as mais relevantes sendo:

- Expansão do método para outros estados comportamentais de ratos Wistar e outros animais.
- Possibilidade de detectar sutilezas ainda maiores no comportamento, como etapas da autolimpeza.
- Utilização do método na análise comportamental de animais de grupos experimentais comumente utilizados na pesquisa pré-clínica sobre autismo, como controle, modelo de autismo induzido por exposição pré-natal ao ácido valproico, grupo tratado com resveratrol (RSV) e grupo VPA + RSV.
- Utilização do método em conjunto com técnicas como a análise de padrões T, que já se mostraram eficazes no estudo comportamental de modelos animais de autismo, mas possuem aplicabilidade limitada devido à dificuldade em realizar classificação manual de comportamento.
- Utilização do método em conjunto com outras formas de análise não supervisionada, visando obter novos indícios sobre o comportamento dos animais.
- Implementação junto a uma interface de fácil utilização, permitindo que o usuário treine classificadores para detecção do etograma mais conveniente em seu estudo.
- Hospedagem do sistema em nuvem, permitindo que pesquisadores de outras instituições utilizem e contribuam com o desenvolvimento futuro do método.

## REFERÊNCIAS

ALDINGER, K.A., KOGAN, J., KIMONIS, V., FERNANDEZ, B., HORN, D., KLOPOCKI, E., CHUNG, B., TOUTAIN., A., WEKSBERG, R., MILLEN, K.J., BARKOVICH, A.J., DOBYNS, W.B., (2012). Cerebellar and posterior fossa malformations in patients with autism-associated chromossome 22q13 terminal deletion. *Am J Med Genet A* 161A, 131-136. doi: 10.1002/ajmg.a.35700

AMERICAN PSYCHIATRY ASSOCIATION (APA), 2013. *Diagnostic and Statistical Manual of Mental Disorders*, 5<sup>th</sup> edition, American Psychiatric Publishing, Inc, Washington, DC.

ANDERSON, D. J., & PERONA, P. (2014). Toward a Science of Computational Ethology. *Neuron*, 84(1), 18–31. doi:10.1016/j.neuron.2014.09.005

BAMBINI-JUNIOR, V., ZANATTA, G., DELLA FLORA NUNES, G., MUELLER DE MELO, G., MICHELS, M., FONTES-DUTRA, M., NOGUEIRA FREIRE, V., RIESGO, R., GOTTFRIED, C. (2014). Resveratrol prevents social deficits in animal model of autism induced by valproic acid.

BANOS, O., GALVEZ, J.-M., DAMAS, M., POMARES, H., & ROJAS, I. (2014). Window Size Impact in Human Activity Recognition. *Sensors*, 14(4), 6474–6499. doi:10.3390/s140406474

BHIKSHA, R. (2016). *Decision Trees*. Disponível em <https://www.cs.cmu.edu/~bhiksha/courses/10-601/decisiontrees/>. Acesso em 29/01/20.

BLANCHARD RJ, FLANNELLY KJ, BLANCHARD DC. (1986). Defensive reactions of laboratory and wild *Rattus norvegicus*. *J Comp Physiol Psychol*.

BLEULER, P.E., 1911. *Dementia Praecox Oder Gruppe Der Schizophrenien*. *Handb. Der Psychiatr.*

BRANDÃO, M. L., ZANOVELI, J. M., RUIZ-MARTINEZ, R. C., OLIVEIRA, L. C., & LANDEIRA-FERNANDEZ, J. (2008). Different patterns of freezing behavior organized in the periaqueductal gray of rats: Association with different types of anxiety. *Behavioural Brain Research*, 188(1), 1–13. doi:10.1016/j.bbr.2007.10.018

- BRANSON, K., ROBIE, A.A., BENDER, J., PERONA, P., DICKINSON, M.H. (2009). High-throughput ethomics in large groups of drosophila. doi:10.1038/nmeth.1328.
- BROWN, A. E. X., & DE BIVORT, B. (2018). Ethology as a physical science. *Nature Physics*, 14(7), 653–657.
- BRYSON, S.E., ROGERS, S.J., FOMBONNE, E. (2003). Autism spectrum disorders: early detection, intervention, education, and psychopharmacological management.
- CAKIR, J., FRYE, R.E., WALKER, S.J. (2020). The lifetime social cost of autism: 1990-2029.
- CASANOVA, M.F. (2006). Neuropathological and genetic findings in autism: the significance of a putative minicolumnopathy.
- CHOLERIS, E., THOMAS, A.W., KAVALIERS, M., PRATO, F.S. (2001). A detailed ethological analysis of the mouse open field test: effects of diazepam, chlordiazepoxide and an extremely low frequency pulsed magnetic field.
- CHANG, Y.-C., COLE, T. B., & COSTA, L. G. (2017). Behavioral Phenotyping for Autism Spectrum Disorders in Mice. *Current Protocols in Toxicology*, 11.22.1–11.22.21. doi:10.1002/cptx.19.
- CHAKURE, A. (2019). Random Forest Regression. Disponível em <https://towardsdatascience.com/random-forest-and-its-implementation-71824ced454f>. Acesso em 24/01/20.
- COURCHESNE, E., MOUTON, P.R., CALHOUN, M.E., SEMENDEFERI, K., AHRENS-BARBEAU, C., HALLET, M.J., BARNES, C.C., PIERCE, K., 2001. Neuron number and size in prefrontal cortex of children with autism. *JAMA* 306, 2001-2010. doi: 306/18/2001 [pii] 10.1001/jama.2011.1638
- DEACON, R. M. J., & RAWLINS, J. N. P. (2006). T-maze alternation in the rodent. *Nature Protocols*, 1(1), 7–12.
- DE CHAUMONT, F., EY, E., TORQUET, N., LAGACHE, T., DALLONGEVILLE, S., IMBERT, A., OLIVO-MARIN, J.-C. (2019). Real-time analysis of the behaviour of groups of mice via a depth-sensing camera and machine learning. *Nature Biomedical Engineering*. doi:10.1038/s41551-019-0396-1

DÍAZ-MORÁN, S., ESTANISLAU, C., CAÑETE, T., BLÁZQUEZ, G., RÁEZ, A., TOBEÑA, A., FERNÁNDEZ-TERUEL, A. (2014). Relationships of open-field behaviour with anxiety in the elevated zero-maze test: Focus on freezing and grooming.

EDUREKA, (2019). K-Nearest Neighbors Algorithm Using Python. Acesso em 28/01/20. Disponível em <https://www.edureka.co/blog/k-nearest-neighbors-algorithm>.

FEEC Unicamp (2002). Morphological opening. Disponível em <http://www.dca.fee.unicamp.br/~lotufo/cursos/ia-870-2002/html/morph/mmopen.html>. Acesso em 27/01/20.

FOMBONNE, E. QUIRKE, S., HAGEN, A. (2009). Prevalence and interpretation of recent trends in rates of pervasive developmental disorders.

FONTES-DUTRA, M., SANTOS-TERRA, J., DECKMANN, I., BRUM SCHWINGEL, G., DELLA-FLORA NUNES, G., HIRSCH, M.M., BAUER-NEGRINI, G., RIESGO, R.S., BAMBINI-JÚNIOR, V., HEDIN-PEREIRA, C., GOTTFRIED, C. (2018). Resveratrol Prevents Cellular and Behavioral Sensory Alterations in the Animal Model of Autism Induced by Valproic Acid.

GADIA, Carlos A.; TUCHMAN, Roberto; ROTTA, Newra T. [Autism and pervasive developmental disorders]. *Jornal de pediatria*, [s. l.], v. 80, n. 2 Suppl, p. S83-94, 2004. Disponível em: <<http://www.ncbi.nlm.nih.gov/pubmed/15154076>>. Acesso em: 26 abr. 2017.

GANDAL, M.J., EDGAR, J.C., EHRLICHMAN, R.S., MEHTA, M., ROBERTS, T.P.L., SIEGEL, S.J. (2010). Validating  $\gamma$  oscillations and delayed auditory responses as translational biomarkers of autism.

GESCHWIND, D.H. (2009). Advances in autism. *Annu Rev Med* 60, 367-380.

GONZALEZ, R.C., WOODS, R.E. (2008). *Digital Image Processing*. Ed. Pearson Educational International.

GOTTFRIED, C., BAMBINI-JUNIOR, V., BARONIO, D., ZANATTA, G., SILVESTRIN, R.B., VACCARO, T. (2013). Valproic Acid in Autism Spectrum Disorder: From and Environmental Risk Factor to a Reliable Animal Model.

GOULD, T. D., DAO, D. T., & KOVACSICS, C. E. (2009). The Open Field Test. *Neuromethods*, 1–20. doi:10.1007/978-1-60761-303-9\_1.

GROEN, W., TELUIJ, M., BUITELAAR, J., TENDOLKAR, I. (2010). Amygdala and hippocampus enlargement during adolescence in autism. *J Am Acad Child Adolesc Psychiatry* 49, 552-560. doi:S0890-8567(10)00200-5 [pii] 10.1016/j.jaac.2009.12.023

HALL, C.S. (1934). Emotional behavior in the rat.

HAN, S., TARALOVA, E., DUPRE, C., YUSTE, R. (2018). Comprehensive machine learning analysis of Hydra behavior reveals a stable basal behavioral repertoire.

HARDEN, C.L. (2013). In utero valproate exposure and autism: long suspected, finally proven.

HARRINGTON, P. (2012). *Machine learning in action*. Shelter Island, N.Y.: Manning Publications Co.

HONG, W., KENNEDYA, A., BURGOS-ARTIZZUB, X.P., ZELIKOWSKYA, M., NAVONNEB, S.G., PERONAB, P., ANDERSONA, D.J. (2015). Automated measurement of mouse social behaviors using depth sensing, video tracking, and machine learning.

IMAI, K., IIDA, T., YAMAMOTO, M., KOMATSU, K., NUKUI, Y., YOSHIZAKA, A. (2014). Psychological and mental health problems in patients with thalidomide embryopathy in Japan.

KABRA, M., ROBIE, A.A., RIVERA-ALBA, M., BRANSON, S., BRANSON, K. (2013). JAABA: interactive machine learning for automatic annotation of animal behavior.

KALUEFF, A.V., ALDRIDGE, J.W., LAPORTE, J.L., MURPHY, D.L., TUOHIMAA, P. (2007). Analyzing grooming microstructure in neurobehavioral experiments.

KALUEFF, A.V., FOX, M.A., GALLAGHER, P.S., MURPHY, D.L. (2007). Hypolocomotion, anxiety and serotonin syndrome-like behavior contribute to the complex phenotype of serotonin transporter knockout mice.

KANDEL, E.R., SCHWARTZ, J.H., JESSELL, T.M., SIEGELBAUM, S.A., HUDSPETH, A.J. (2014). *Princípios de neurociências*.

KANG, U.J., MEREDITH, G.E. (2006). Behavioral models of Parkinson's disease in rodents: A new look at an old problem.

KIM, P., PARK, J.H., KWON, K.J., KIM, K.C., KIM, H.J., LEE, J.M., KIM, H.Y., HAN, S.H., SHIN, C.Y. (2013). Effects of Korean red ginseng extracts on neural tube defects and impairment of social interaction induced by prenatal exposure to valproic acid.

LARHMAM (2018). SVM margin. Disponível em [https://commons.wikimedia.org/wiki/File:SVM\\_margin.png](https://commons.wikimedia.org/wiki/File:SVM_margin.png). Acesso em 30/01/20.

LAVELLE, T.A., WEINSTEIN, M.C., NEWHOUSE, J.P., MUNIR, K., KUHLTHAU, KA., PROSSER, L.A. (2014). Economic burden of childhood autism spectrum disorders.

MATHWORKS. MATLAB documentation. Disponível em <https://www.mathworks.com/help/matlab/index.html>. Acesso em 27/01/20.

MAYES, S.D., CALHOUN, S.L., MURRAY, M.J., PEARL, A., BLACK, A., TIERNEY, C.D. (2014). Final DSM-5 under-identifies mild Autism Spectrum Disorder: Agreement between the DSM-5, CARS, CASD, and clinical diagnoses.

MCFARLANE, H. G., KUSEK, G. K., YANG, M., PHOENIX, J. L., BOLIVAR, V. J., CRAWLEY J. N. (2007). Autism-like behavioral phenotypes in BTBR T1tf/J mice.

MILES, J.H. (2011). Autism spectrum disorders – a genetics review.

MILLER, M.T., STROMLAND, K. (1999). Teratogen update: thalidomide: a review, with a focus on ocular findings and new potential uses.

MORBIDITY AND MORTALITY WEEKLY REPORT (MMWR), 2014. Prevalence of Autism Spectrum Disorder among Children Aged 8 Years – Autism and Developmental Disabilities Monitoring Network, 11 Sites, United States, 2010, Vol. 63(2). ed. Surveillance Summaries, Washington, DC.

MUHLE, R., TRENTACOSTE, S.V., RAPIN, I. (2004). The genetics of autism.

MURTA, L.O. Processamento e análise de imagens médicas. Disponível em <http://docplayer.com.br/35913779-Processamento-e-analise-de-imagens-medicas.html>. Acesso em 29/01/20.

NJUNG'E, K., & HANDLEY, S. L. (1991). Evaluation of marble-burying behavior as a model of anxiety. *Pharmacology Biochemistry and Behavior*, 38(1), 63–67. doi:10.1016/0091-3057(91)90590-x

Noldus, January 20, 2020. EthoVisionXT. Retrieved from <https://www.noldus.com/ethovision-xt>

OTSU, N. (1979). A Threshold Selection Method from Gray-Level Histograms. *IEEE Transactions on Systems, Man, and Cybernetics*, 9(1), 62–66. doi:10.1109/tsmc.1979.4310076

ODSC, (2019). Confronting the curse of dimensionality. Disponível em <https://medium.com/@ODSC/confronting-the-curse-of-dimensionality-5bcf2998b30d>. Acesso em 20/01/20.

RAMOS, A. (2008). Animal models of anxiety: do I need multiple tests?

RAPIN, I.; TUCHMAN, R. F. What is new in autism? *Curr Opin Neurol*, [s. l.], v. 21, n. 2, p. 143–149, 2008.

REN, Z., ANNIE, A. N., CIERNIA, V., & LEE, Y. J. (2017). Who Moved My Cheese? Automatic Annotation of Rodent Behaviors with Convolutional Neural Networks. 2017 IEEE Winter Conference on Applications of Computer Vision (WACV). doi:10.1109/wacv.2017.147

RIESGO, R., GOTTFRIED, C., BRISTOT, R. (2013). *Recent Advances in Autism Spectrum Disorders – Volume I*.

RODIER, P.M., INGRAM, J.L., TISDALE, B., CROOG, V.J. (1997). Linking etiologies in humans and animal models: studies of autism.

RODIER, P.M., INGRAM, J.L., TISDALE, B., NELSON, S., ROMANO, J. (1996). Embryological origin for autism: Developmental anomalies of the cranial nerve motor nuclei.

ROULLET, F.I., LAI, J.K., FOSTER, J.A. (2013). In utero exposure to valproic acid and autism – a current review of clinical and animal studies.

PAULA, C.S., RIBEIRO, S.H., FOMBONNE, E., MERCADANTE, M.T. (2011). Brief report: Prevalence of pervasive developmental disorder in Brazil: a pilot study.

PENNINGTON, Z.T., DONG, Z., FENG, Y., VETERE, L.M., PAGE-HARLEY, L., SHUMAN, T., DENISE J.C. (2019). eztrack: An open-source video analysis pipeline for the investigation of animal behavior.

- SAMUEL, A. L. (1959). Some Studies in Machine Learning Using the Game of Checkers. *IBM Journal of Research and Development*, 3(3), 210–229. doi:10.1147/rd.33.0210
- SANCHEZ-MARONO, N., ALONSO-BETANZOS, A., GARCIA-GONZALEZ, P., & BOLON-CANEDO, V. (2010). Multiclass classifiers vs multiple binary classifiers using filters for feature selection. *The 2010 International Joint Conference on Neural Networks (IJCNN)*. doi:10.1109/ijcnn.2010.5596567
- SANDIN, S., LICHTENSTEIN, P., KUJA-HALKOLA, R., LARSSON, H., HULTMAN, C.M., REICHENBERG, A. (2014). The familial risk of autism.
- SCHNEIDER, T., PRZEWLOCKI, R. (2005). Behavioral alterations in rat prenatally exposed to valproic acid: animal model of autism.
- SCHNEIDER, T., ROMAN, A., BASTA-KAIM, A., KUBERA, M., BUDZISZEWSKA, B., SCHNEIDER, K., PRZEWLOCKI, R. (2008). Gender-specific behavioral and immunological alterations in an animal model of autism induced by prenatal exposure to valproic acid.
- SHAMS S., AMLANI, S., SCICLUNA, M., GERLAI, R. (2018). Argus: An open-source and flexible software application for automated quantification of behavior during social interaction in adult zebrafish.
- SMITH, V., BROWN, N. (2014). Prenatal valproate exposure and risk of autism spectrum disorders and childhood autism.
- STERN, U., HE, R., YANG, C.H. (2015). Analyzing animal behavior via classifying each video frame using convolutional neural networks. *Scientific Reports*, 5(1). doi:10.1038/srep14351
- STOELTING CO., (2019). ANY-maze: Behavioral tracking software. Disponível em [www.anymaze.co.uk](http://www.anymaze.co.uk)
- STURMAN, O., GERMAINBAND, P.L., BOHACEK, J. (2018). Exploratory rearing: a context- and stress-sensitive behavior recorded in the open-field test.
- SUKHAREVA, G. E. (1926). Die schizoiden Psychopathien im Kindesalter.

SUMMARIES, S. (2007). Prevalence of autism spectrum disorders – autism and developmental disabilities monitoring network, six sites, United States, 2000.

TAMERSOY, B. (2009). Background subtraction. Disponível em [http://www.cs.utexas.edu/~grauman/courses/fall2009/slides/lecture9\\_background.pdf](http://www.cs.utexas.edu/~grauman/courses/fall2009/slides/lecture9_background.pdf). Acesso em 20/01/20.

TSUJINO, N., NAKATANI, Y., SEKI, Y., NAKASATO, A., NAKAMURA, M., SUGAWARA, M., ARITA, H. (2007). Abnormality of circadian rhythm accompanied by an increase in frontal cortex serotonin in animal model of autism.

VACAVANT, A., CHATEAU, T., WILHELM, A., & LEQUIEVRE, L. (2013). A Benchmark Dataset for Outdoor Foreground/Background Extraction. Lecture Notes in Computer Science, 291–300.

VAN KOOTEN, I.A., PALMEN, S.J., VON CAPPELN, P., STEINBUSCH, H.W., KORR, H., HEINSEN, H., HOF, P.R., VAN ENGELAND, H., SCHMITZ, C. (2008). Neurons in the fusiform gyrus are fewer and smaller in autism.

VAPNIK, V. and A. STERIN, (1977). On structural risk minimization or overall risk in a problem of pattern recognition.

WOLFF, S. (1996). The first account of the syndrome Asperger described?

ZHANG, S., LI, X., ZONG, M., ZHU, X., & CHENG, D. (2017). Learning k for kNN Classification. ACM Transactions on Intelligent Systems and Technology, 8(3).