DATASETS PÚBLICOS PARA ANÁLISE COMPORTAMENTAL DE ROEDORES: CARACTERÍSTICAS E COMPATIBILIDADE COM FRAMEWORKS DE VISÃO COMPUTACIONAL

José Pablo Pinto da Silva[1], André do Amaral Penteado Bíscaro[2], Ivan Luiz Pedroso Pires[3], Tales Nereu Bogoni[4], Nádia Aléssio Velloso[5], Érico Fernando de Oliveira Martins[6]

[1,2]Universidade do Estado de Mato Grosso, Faculdade de Ciências Exatas e Tecnológicas, Engenharia Elétrica, Sinop, Brasil (jose.pablo@unemat.br)
[5]Universidade Federal de Mato Grosso, Programa de Pós-Graduação em Ciências da Saúde, Sinop, Brasil
[3,4,6]Universidade do Estado de Mato Grosso, Faculdade de Ciências Exatas e Tecnológicas, Sinop, Sistemas de Informação, Sinop, Brasil

Resumo: Este estudo apresenta uma análise comparativa de *datasets* aplicados ao rastreamento comportamental de roedores, considerando características de captura, tipos de anotação, organização dos rótulos e compatibilidade com *frameworks* computacionais. Foram avaliados os *datasets Rodent3D*, *RGB-D Rat* e *MARS Dataset*. Os resultados mostram que diferenças estruturais influenciam diretamente a reutilização dos dados e sua adequação a tarefas de detecção, estimativa de pose e classificação comportamental.

Palavras-chave: Comportamento animal; Visão computacional; Estruturas anotacionais.

INTRODUÇÃO

Os modelos animais desempenham um papel imprescindível na investigação de distúrbios neuropsiquiátricos, especialmente em estudos voltados à compreensão de mecanismos biológicos, respostas farmacológicas e alterações comportamentais. Entre esses modelos, os roedores ocupam posição de destaque devido à elevada similaridade genética com os seres humanos, à facilidade de manejo em ambiente laboratorial e à ampla utilização em protocolos experimentais. Ainda assim, conforme discutido por Nestler e Hyman (2010), a validação de modelos animais plenamente representativos de condições humanas permanece como um desafio, o que reforça a necessidade de métodos analíticos mais objetivos, reprodutíveis e tecnicamente consistentes.

Com o avanço da visão computacional e do aprendizado de máquina, a análise automatizada do comportamento animal tem adquirido crescente relevância em pesquisas experimentais. Técnicas de rastreamento, estimativa de pose e classificação comportamental vêm substituindo progressivamente abordagens baseadas exclusivamente em observação manual, tradicionalmente sujeitas à subjetividade do avaliador, à variabilidade entre observadores e à limitação de escala analítica. *Frameworks* como *DeepLabCut*, YOLO e MARS exemplificam esse avanço ao possibilitar desde a extração de coordenadas corporais até a classificação de comportamentos complexos em diferentes contextos experimentais (Kabra et al., 2013; Bello-Arroyo et al., 2018; Segalin et al., 2021; Lauer et al., 2022).

No contexto da visão computacional, *frameworks* podem ser entendidos como conjuntos estruturados de algoritmos, modelos, bibliotecas e procedimentos computacionais que oferecem suporte à execução de tarefas específicas, como detecção de objetos, estimativa de pose, rastreamento e classificação comportamental. Esses ambientes fornecem métodos padronizados para processamento, treinamento e interpretação dos dados, tornando a organização e a estrutura dos *datasets* fatores decisivos para sua utilização eficiente.

Os *datasets* públicos correspondem a conjuntos organizados de vídeos, imagens, anotações e metadados disponibilizados para utilização pela comunidade científica. Essas bases constituem recursos essenciais para treinamento, validação e comparação de modelos computacionais, favorecendo a reprodutibilidade dos experimentos e o desenvolvimento de novas metodologias de rastreamento, estimativa de pose e classificação comportamental.

Nesse sentido, essas bases assumem função estratégica, pois permitem comparar metodologias, reproduzir experimentos e avaliar a capacidade de generalização dos modelos computacionais. Conjuntos de dados como *Rodent3D*, *RGB*-D *Rat* e *MARS Dataset* ilustram a diversidade atualmente disponível, mas também evidenciam diferenças importantes quanto ao tipo de dado fornecido, ao nível de anotação e ao escopo experimental contemplado (Gerós et al., 2020; Patel et al., 2023; Segalin et al., 2021).

Diante desse cenário, este artigo propõe uma análise comparativa de *datasets* públicos aplicados ao

rastreamento e à análise automatizada do comportamento de roedores, considerando aspectos relacionados às estruturas de vídeo, modalidades de captura, tipos de anotação, organização dos rótulos e compatibilidade com frameworks amplamente utilizados em visão computacional. Busca-se, ainda, discutir as potencialidades e limitações dessas bases quanto à reutilização computacional e à sua aplicação em tarefas de detecção, estimativa de pose e classificação comportamental.

MATERIAL E MÉTODOS

Este trabalho caracteriza-se como uma pesquisa documental, comparativa e exploratória, desenvolvida a partir de um levantamento sistematizado de *datasets* públicos e *frameworks* computacionais aplicados ao rastreamento e à análise automatizada do comportamento de roedores.

A seleção dos *datasets* e *frameworks* considerou sua disponibilidade pública, ampla utilização na literatura especializada e representatividade em diferentes tarefas de visão computacional aplicadas ao comportamento animal. Buscou-se contemplar bases de dados com distintas modalidades de captura, estruturas anotacionais e contextos experimentais, bem como *frameworks* voltados à detecção de objetos, estimativa de pose e classificação comportamental, permitindo uma análise comparativa abrangente das relações de compatibilidade entre dados e métodos computacionais.

O corpus documental do estudo foi composto por artigos científicos, repositórios públicos de dados, documentações oficiais, manuais de instalação, tutoriais técnicos, materiais audiovisuais de apoio e documentos institucionais relacionados aos *datasets* e frameworks selecionados. Entre os materiais analisados, incluíram-se documentos técnicos e científicos associados aos *datasets* *Rodent3D*, *RGB*-D *Rat* e *MARS Dataset*, bem como aos *frameworks* *DeepLabCut*, *YOLO* e *MARS*.

Os *dataset*s foram analisados quanto à estrutura dos vídeos, resolução espacial e temporal, modalidade de captura, número de animais por cena, diversidade de cenários experimentais, presença de oclusões, disponibilidade de metadados, tipo de anotação e organização dos rótulos. Além disso, foram avaliadas características relacionadas à estrutura anotacional de cada base, considerando a forma como as informações são representadas e disponibilizadas para utilização em diferentes modelos computacionais.

Os *frameworks* foram analisados quanto às suas finalidades principais, tipos de entrada de dados suportados, estruturas anotacionais requeridas e tarefas computacionais predominantes, incluindo tarefas de detecção de objetos, estimativa de pose e classificação comportamental. Com base nessas características, foram identificadas relações de compatibilidade entre os requisitos operacionais de cada *framework* e as informações disponibilizadas pelos *datasets* analisados.

Realizou-se uma análise comparativa entre as características estruturais dos *datasets* e os requisitos operacionais dos frameworks selecionados, buscando identificar relações de compatibilidade metodológica e potencial de reutilização dos dados.

Complementarmente, foi conduzida uma avaliação exploratória das anotações presentes em cada *dataset*, buscando compreender diferenças relacionadas aos formatos de representação dos dados, à organização dos rótulos e aos tipos de informação disponibilizada. Essa etapa contribuiu para a identificação de potencialidades e limitações associadas à reutilização dessas bases em diferentes tarefas de visão computacional.

RESULTADOS E DISCUSSÃO

A análise comparativa dos *datasets* públicos voltados ao rastreamento e à análise comportamental de roedores evidenciou uma expressiva pluralidade entre as bases disponíveis, tanto em relação à estrutura dos vídeos quanto aos tipos de anotação, à organização dos rótulos e às condições experimentais. Essa diversidade não representa apenas uma diferença de formato ou organização interna, mas influencia diretamente o potencial de uso de cada *dataset* em tarefas de detecção, rastreamento, estimativa de pose e classificação comportamental.

Entre os principais aspectos observados, destacam-se a modalidade de captura, a resolução espacial e temporal, o número de animais por cena, a diversidade de cenários experimentais, a presença de metadados e o tipo de anotação associado a cada base. Tais fatores são cruciais para determinar os potenciais impactos ocasionados à capacidade de generalização em experimentos distintos, os quais não foram originalmente desenvolvidos. A Tabela 1 sintetiza os principais aspectos técnicos dos *dataset*s analisados, incluindo modalidades de captura, resolução e número de animais por cena.

Tabela 1. Aspectos gerais dos *dataset*s analisados.

Dataset	Captura	Resolução	Animais
Rodent3D	Multivista, RGB-D e térmico	1024×1024 (130 Hz) / 848×480 (60 Hz)	1-2
RGB-D Rat	RGB-D	1920×1080 / 512×424; até 30 fps	1
MARS Dataset	Infravermelho	1024×570 e 1280×500; 30 fps	2

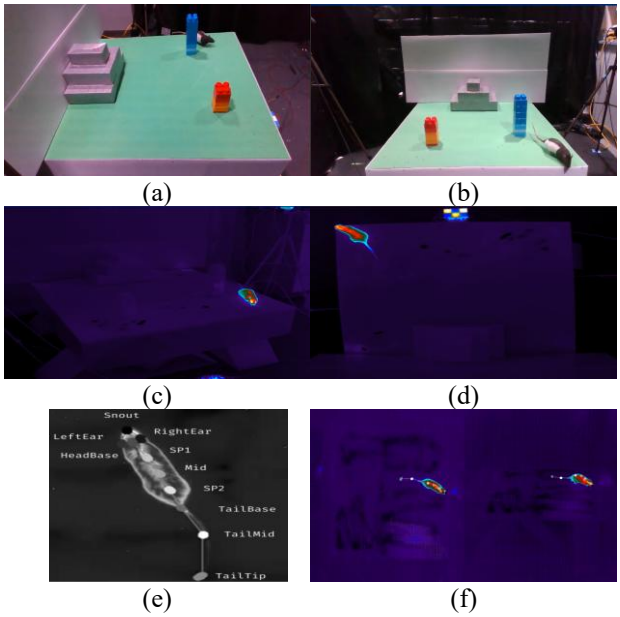
A partir das informações sintetizadas na Tabela 1, compreende-se que os três *dataset*s analisados não diferem apenas em termos de formato de captura, mas também quanto ao tipo de tarefa para a qual parecem mais naturalmente orientados.

O *Rodent3D* apresenta a estrutura técnica mais complexa entre as bases comparadas, ao combinar aquisição multivista (diferentes pontos de captura) com dados RGB-

De térmicos, além de contemplar um a dois animais por cena e três condições experimentais distintas (Com objetos, sem objetos e dispensador de comida). Essa configuração amplia significativamente o potencial analítico espacial e postural, sobretudo em contextos nos quais a reconstrução tridimensional do movimento e a observação de interações entre indivíduos sejam relevantes.

No entanto, essa mesma riqueza estrutural tende a elevar a dificuldade de processamento, anotação e adaptação para pipelines mais simplórios, especialmente aqueles voltados apenas à detecção bidimensional ou ao rastreamento convencional.

A Figura 1 apresenta exemplos de imagens e representações disponibilizadas pelo *Rodent3D*, evidenciando a diversidade de modalidades de aquisição presentes no *dataset*, incluindo imagens RGB, imagens térmicas e representações corporais baseadas em keypoints.



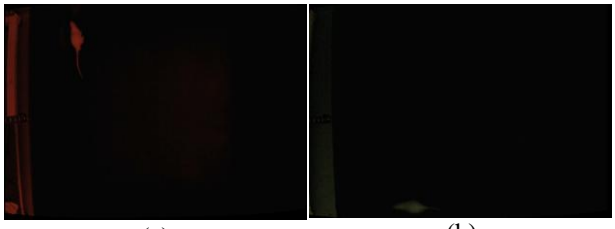
(a) (b) (c) (d) (e) (f)

Figura 1. Exemplos de imagens do *dataset Rodent3D*: (a) captura RGB lateral; (b) captura RGB frontal; (c) imagem térmica diagonal; (d) imagem térmica superior; (e) representação corporal por *keypoints*; e (f) estrutura multimodal do conjunto de dados. Fonte: adaptado de Patel et al. (2023).

O *RGB*-D *Rat Dataset*, por sua vez, apresenta uma configuração mais controlada, centrada em um único animal em arena de campo aberto, com resolução RGB elevada e profundidade. Essa característica o aproxima de protocolos voltados à observação locomotora individual em ambientes experimentais controlados. Além disso, a presença de anotação comportamental frame a frame favorece análises temporais mais diretas e tarefas de classificação supervisionada.

Ainda assim, a existência de apenas um cenário experimental principal reduz a diversidade contextual da base, o que pode limitar sua capacidade de generalização para situações mais variadas de captura, iluminação ou organização experimental. Ou seja, trata-se de um *dataset* potencialmente útil para aplicações específicas e mais controladas, mas com menor abrangência metodológica quando comparado a bases estruturalmente mais amplas.

A Figura 2 apresenta exemplos de imagens obtidas no *RGB-D Rat* sob diferentes condições de iluminação, ilustrando uma das características experimentais consideradas durante a construção da base de dados.



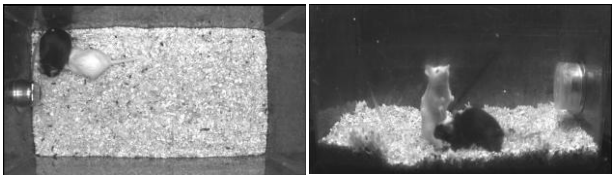
(a) (b)

Figura 2. Exemplos de imagens do dataset *RGB*-D *Rat* obtidas sob diferentes condições de iluminação: (a) luz vermelha e (b) luz branca. Fonte: adaptado de Gerós et al. (2019).

No caso do *MARS Dataset*, a combinação entre gravação em infravermelho, registro em duas vistas e anotação simultânea de pose e comportamento confere à base uma característica intermediária entre controle experimental e riqueza informacional.

Todavia, o fato de trabalhar com dois animais por cena e com foco em interação social amplia o nível de complexidade comportamental representada, o que é vantajoso para tarefas de classificação de ações e reconhecimento de padrões sociais. Por outro lado, justamente por estar fortemente ligado a um protocolo de interação entre pares, seu uso tende a ser mais adequado a estudos voltados à interação social e ao reconhecimento de padrões comportamentais entre múltiplos indivíduos.

A Figura 3 apresenta exemplos de imagens do *MARS Dataset* obtidas em diferentes perspectivas de captura, evidenciando o contexto experimental utilizado para análise de interações sociais entre roedores.



(a) (b)

Figura 3. Exemplos de imagens do *MARS Dataset* ilustrando diferentes perspectivas empregadas na aquisição de dados para rastreamento comportamental multi-indivíduo: (a) vista superior e (b) vista frontal. Fonte: adaptado de Segalin et al. (2021).

Outro aspecto importante refere-se à relação entre resolução espacial e resolução temporal. Em bases como o *Rodent3D*, a presença de modalidades de captura com alta taxa temporal e diferentes resoluções indica maior competência para rastreamento refinado de movimento e reconstrução de postura, o que pode beneficiar tarefas mais exigentes em termos cinemáticos. Já no *RGB-D Rat*

e no *MARS Dataset*, embora as taxas de quadros sejam mais modestas, elas ainda se mostram suficientes para análises comportamentais em nível de quadro, sobretudo quando associadas a anotações previamente estruturadas.

A presença de metadados em todos os *datasets* analisados também representa um fator metodologicamente relevante. Ainda que essa característica não elimine as limitações inerentes a cada base, ela favorece a reprodutibilidade, a interpretação técnica dos experimentos e a organização de pipelines de treinamento e validação.

Além das características relacionadas à captura dos dados, aspectos como diversidade de cenários experimentais, disponibilidade de metadados e tipos de anotação também contribuem para diferenciar os *datasets* analisados. A Tabela 2 sintetiza essas informações.

Tabela 2. Cenários experimentais, metadados e tipos de anotação dos *datasets* analisados.

Dataset	Cenários	Metadados	Anotação
Rodent3D	3	Sim	Pose 2D e 3D
RGB-D Rat	1	Sim	Comportamento por frame
MARS Dataset	1	Sim	Pose e comportamento

Em estudos envolvendo inteligência artificial aplicada ao comportamento animal, a existência de metadados bem descritos tende a reduzir ambiguidades quanto às condições de captura, ao protocolo adotado e ao significado das anotações, contribuindo para o reaproveitamento das bases em novos contextos analíticos.

Nesse sentido, a disponibilidade de metadados deve ser entendida não apenas como um recurso auxiliar, mas como um elemento que impacta diretamente a transparência e a robustez metodológica do *dataset*. Em relação ao tipo de anotação, as diferenças observadas entre as bases ajudam a explicar por que sua compatibilidade com diferentes frameworks não é homogênea.

O *Rodent3D*, ao enfatizar pose 2D e 3D, tende a dialogar melhor com pipelines voltados à estimação de pose e modelagem postural. O *RGB-D Rat*, com anotação comportamental frame a frame, mostra-se mais imediatamente orientado a tarefas de classificação temporal supervisionada.

Já o *MARS Dataset*, por reunir pose e comportamento, apresenta maior versatilidade analítica, embora dentro de um recorte experimental específico. Dessa forma, a natureza da anotação influencia não apenas o tipo de tarefa que pode ser executada com maior facilidade, mas também o grau de adaptação necessário para tornar cada base compatível com frameworks como *YOLO*, *DeepLabCut* e *MARS*.

Levando em consideração os aspectos anteriormente mencionados, torna-se fundamental examinar de maneira mais específica como cada *dataset* organiza e disponibiliza suas anotações, uma vez que esse fator influencia diretamente em sua utilização em diferentes pipelines computacionais.

Nesse sentido, a análise a seguir destaca, para cada base, os principais tipos de anotação empregados, incluindo pontos corporais e comportamentos registrados, bem como a organização dos rótulos e o grau de compatibilidade com frameworks amplamente utilizados, como *YOLO*, *DeepLabCut* e *MARS*.

No caso do *Rodent3D*, trata-se de um *dataset* claramente voltado à pose, com dez pontos corporais anotados: focinho, orelha direita, orelha esquerda, base da cabeça, Sp1, região média do dorso, Sp2, base da cauda, porção média da cauda e ponta da cauda. Esse formato favorece tarefas de rastreamento fino e modelagem postural, o que explica sua maior proximidade com frameworks como o DeepLabCut.

As Tabelas 3 e 4 sintetizam essas informações com o objetivo de oferecer uma visão comparativa, que servirá de base para a discussão detalhada das potencialidades, limitações e exigências de adaptação de cada *dataset*.

Tabela 3. Anotações.

Dataset	Pontos corporais	Comportamentos
Rodent3D	10	-
RGB-D Rat	-	7
MARS Dataset	9 topo / 13 frontal	3

A Tabela 3 mostra que os *datasets* analisados não se diferenciam apenas pelo tipo de informação anotada, mas também pela forma como essas anotações são organizadas. Isso influencia diretamente a maneira como cada base pode ser reutilizada. Por outro lado, como o *Rodent3d* não foi organizado originalmente em torno de rótulos comportamentais por quadro, seu uso em classificação direta de comportamento exige etapas adicionais de interpretação ou extração de eventos a partir da pose.

O *RGB-D Rat* segue uma lógica diferente. Sua organização está centrada em comportamentos anotados quadro a quadro, o que o aproxima mais pontualmente de tarefas de classificação supervisionada. As sete classes presentes na base são: parado, exploração em movimento, exploração local, caminhada, alimentação suportada, criação sem suporte, tosa escovadora.

Nesse caso, os rótulos já estão muito próximos da saída final esperada por modelos de reconhecimento comportamental, o que favorece esse tipo de aplicação.

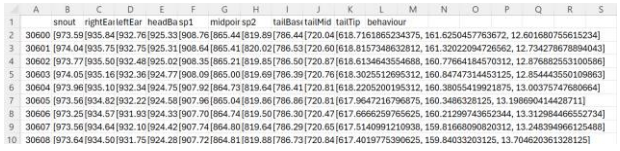
Em contrapartida, como a base não tem a pose como elemento central de anotação, seu aproveitamento em frameworks voltados à modelagem postural, como o

DeepLabCut, torna-se mais limitado, a menos que seja feito um novo processo de rotulagem corporal.

Assim, embora seja uma base forte para análise comportamental temporal, sua flexibilidade para outros usos é menor. O *MARS Dataset* ocupa uma posição versátil, porque combina pose, comportamento e *bounding boxes*. Na vista superior, os pontos corporais anotados incluem: nariz, orelhas, base do pescoço, quadril esquerdo, quadril direito, base da cauda, porção média da cauda e ponta da cauda, totalizando nove pontos corporais. Na perspectiva frontal, esses mesmos pontos aparecem acompanhados ainda pelas quatro patas, chegando a treze *keypoints*. No plano comportamental, a base está centrada principalmente em três ações sociais: ataque, montada e investigação próxima. Essa organização amplia as possibilidades de reutilização da base, permitindo desde tarefas de pose até detecção e classificação.

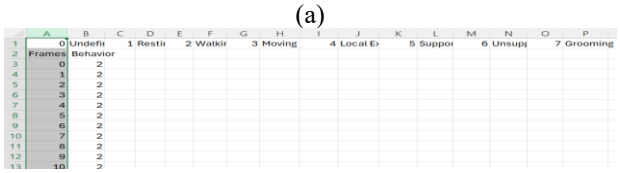
Por isso, sua compatibilidade com *MARS*, *DeepLabCut* e *YOLO* tende a ser mais direta do que nos outros casos, embora o grau de adaptação ainda dependa do objetivo analítico. Mesmo assim, essa versatilidade não elimina a necessidade de adaptações, especialmente quando se pretende empregar a base em contextos diferentes daqueles para os quais foi objetivamente criada.

A Figura 4 apresenta exemplos das estruturas anotacionais adotadas pelos *datasets* analisados, permitindo visualizar diferenças na organização dos rótulos e na representação das informações utilizadas pelos diferentes frameworks computacionais.



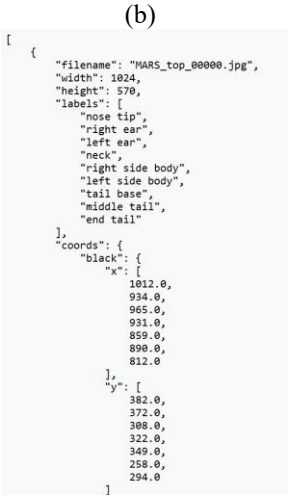
	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P	Q	R	S
1		snout	rightEar	leftEar	headBa	sp1	midpoir	sp2	tailBas	tailMid	tailTip	behaviour							
2	30600	[973.59	935.84	932.76	925.33	908.76	865.44	819.89	786.44	720.04	618.7161865234375, 161.6250457763672, 12.601680755615234]								
3	30601	[974.04	935.75	932.75	925.31	908.64	865.41	820.02	786.53	720.60	618.8157348632812, 161.3202209472656, 12.734278678894043]								
4	30602	[973.77	935.50	932.48	925.02	908.35	865.21	819.85	786.50	720.87	618.6134643554688, 160.7766418457031, 12.876882553100586]								
5	30603	[974.05	935.16	932.36	924.77	908.09	865.00	819.69	786.39	720.76	618.3025512695312, 160.8474731445312, 12.854443550109863]								
6	30604	[973.96	935.10	932.34	924.75	907.92	864.73	819.64	786.41	720.81	618.2205200195312, 160.38055419921875, 13.003757476806664]								
7	30605	[973.56	934.82	932.22	924.58	907.96	865.04	819.86	786.86	720.81	617.9647216796875, 160.3486328125, 13.198690414428711]								
8	30606	[973.25	934.57	931.93	924.33	907.70	864.74	819.50	786.30	720.47	617.6666259765625, 160.21299743652344, 13.312984466552734]								
9	30607	[973.56	934.64	932.10	924.42	907.74	864.80	819.64	786.29	720.65	617.5140991210938, 159.81668090820312, 13.248394966125488]								
10	30608	[973.64	934.50	931.75	924.28	907.72	864.81	819.88	786.73	720.84	617.4019775390625, 159.84033203125, 13.704620361328125]								

(a)



	A	B	C	D	E	F	G	H	I	J	K	L	M	N	O	P
1	0	Undefin		1 Restii		2 Walkir		3 Moving		4 Local E		5 Suppor		6 Unsupi		7 Grooming
2	Frames	Behavior														
3	0	2														
4	1	2														
5	2	2														
6	3	2														
7	4	2														
8	5	2														
9	6	2														
10	7	2														
11	8	2														
12	9	2														
13	10	2														

(b)




```
[
    {
        "filename": "MARS_top_00000.jpg",
        "width": 1024,
        "height": 570,
        "labels": [
            "nose tip",
            "right ear",
            "left ear",
            "neck",
            "right side body",
            "left side body",
            "tail base",
            "middle tail",
            "end tail"
        ],
        "coords": {
            "black": {
                "x": [
                    1012.0,
                    934.0,
                    965.0,
                    931.0,
                    859.0,
                    890.0,
                    812.0
                ],
                "y": [
                    382.0,
                    372.0,
                    308.0,
                    322.0,
                    349.0,
                    258.0,
                    294.0
                ]
```


(c)

Figura 4. Exemplos das estruturas anotacionais utilizadas nos datasets analisados: (a) coordenadas corporais e atributos comportamentais do *Rodent3D*; (b) rótulos comportamentais organizados por frame no *RGB-D Rat*; e (c) estrutura hierárquica de anotação em formato JSON do *MARS Dataset*. Fonte: elaborado pelo autor a partir de Patel et al. (2023), Gerós, Magalhães e Aguiar (2019) e Segalin et al. (2021).

A Figura 4 evidencia que os *datasets* analisados adotam estratégias anotacionais significativamente distintas, variando entre representações baseadas em keypoints corporais, rótulos comportamentais organizados por quadro e estruturas híbridas que combinam pose, comportamento e delimitação espacial por *bounding boxes*.

Essas diferenças não influenciam apenas a forma como as informações são armazenadas, mas também o nível de compatibilidade de cada base com diferentes frameworks computacionais.

Nesse contexto, a organização estrutural das anotações passa a desempenhar papel central na definição das tarefas que podem ser executadas com maior facilidade, bem como no grau de adaptação necessário para integração com pipelines voltados à detecção, estimativa de pose e classificação comportamental.

A Tabela 4 sintetiza essas relações considerando a organização predominante dos rótulos e sua compatibilidade principal com diferentes frameworks.

Tabela 4. Organização dos rótulos e compatibilidade principal.

Dataset	Organização dos rótulos	Compatibilidade
Rodent3D	Pose multimodal	DeepLabCut
RGB-D Rat	Comportamento por quadro	Classificação supervisionada
MARS Dataset	Pose, comportamento e boxes	MARS, DeepLabCut e YOLO

Outro fator relevante na análise dos *datasets* refere-se às diferentes perspectivas e modalidades de aquisição empregadas durante a captura dos dados. Essas características influenciam diretamente aspectos como visibilidade dos animais, ocorrência de oclusões, qualidade da estimativa de pose e desempenho dos algoritmos de rastreamento. Em frameworks de pose, tais aspectos tendem a melhorar a estimativa dos pontos corporais, principalmente em situações de rotação do corpo, mudanças bruscas de orientação e interações entre indivíduos. Ao mesmo tempo, essa riqueza estrutural aumenta a complexidade metodológica, uma vez que exige sincronização, calibração e integração entre diferentes modalidades. A Figura 5 apresenta exemplos das configurações de captura adotadas nos *datasets* analisados, evidenciando a diversidade de perspectivas, sensores e condições de aquisição utilizadas.

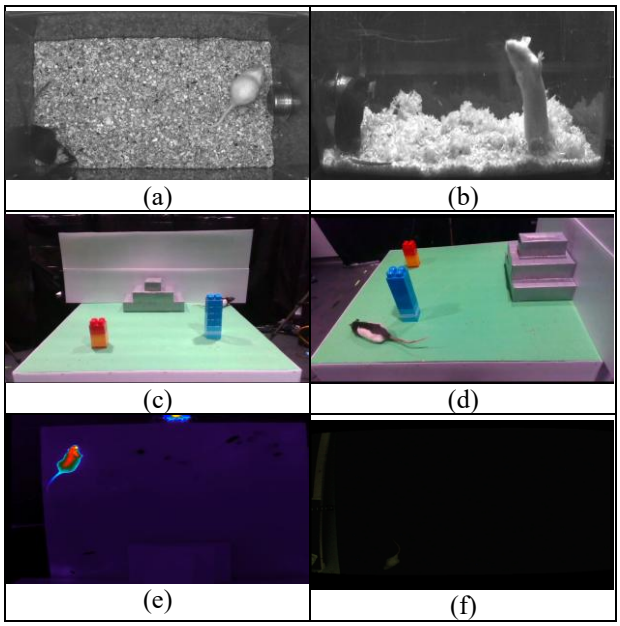


Figura 5. Diferentes perspectivas e modalidades de captura dos datasets analisados: (a) *MARS* (superior); (b) *MARS* (frontal); (c) *Rodent3D* (frontal); (d) Rodent3D (lateral); (e) *Rodent3D* (térmica); e (f) *RGB-D Rat* (luz branca). Fonte: elaborado pelo autor a partir de Patel et al. (2023), Gerós, Magalhães e Aguiar (2019) e Segalin et al. (2021).

No *Rodent3D*, o uso de captura multivista e multimodal reduz limitações comuns de uma única câmera, favorecendo uma reconstrução mais completa da postura e diminuindo ambiguidades de rastreio.

No *MARS Dataset*, a presença simultânea de vista superior e frontal também traz vantagens importantes. A vista superior favorece o rastreamento do deslocamento, da posição relativa entre os animais e da ocupação espacial da arena, sendo útil para métricas locomotoras e de proximidade.

A vista frontal, portanto, ajuda na observação de posturas específicas e de comportamentos sociais que poderiam ficar oclusos ou parcialmente ocultos se apenas a perspectiva superior fosse utilizada. Ainda assim, como há dois animais por cena, ambas as perspectivas estão sujeitas a oclusões, seja pela sobreposição dos corpos, seja pela ocultação parcial de segmentos anatômicos. Em frameworks de pose, isso pode dificultar a localização correta dos *keypoints*. Em frameworks de comportamento, pode gerar ambiguidade na segmentação temporal das ações ou na interpretação do evento observado.

No *RGB-D Rat*, como o cenário é mais controlado e envolve apenas um animal em campo aberto, a chance de oclusão por interação entre indivíduos é evidentemente menor. Isso tende a favorecer a estabilidade de modelos de classificação comportamental, principalmente porque os rótulos já estão organizados por quadro.

Ainda assim, a perspectiva única de gravação pode limitar a observação de detalhes posturais mais sutis, especialmente em comportamentos cuja distinção depende de mudanças corporais discretas. Por isso, embora a base seja adequada para reconhecimento comportamental em ambiente controlado, ela oferece menos riqueza estrutural para tarefas mais refinadas de modelagem postural ou reconstrução espacial.

De modo mais amplo, a comparação mostra que a organização dos rótulos e a perspectiva de captura não podem ser analisadas separadamente. Bases com pontos corporais dependem fortemente da boa visibilidade do corpo para redução de oclusões, enquanto bases centradas em comportamento por quadro dependem mais da consistência temporal das classes e da clareza visual do contexto observado. Assim, diferentes perspectivas de gravação tendem a favorecer tarefas distintas. Vistas múltiplas e multimodais costumam beneficiar frameworks de pose. Vistas mais controladas e centradas no contexto tendem a favorecer classificadores comportamentais. Já estruturas híbridas, como a do MARS, ampliam a flexibilidade de uso, embora exijam mais cuidado na escolha da tarefa e do pipeline.

A análise das Tabelas 3 e 4 reforça que a utilidade de cada dataset não depende apenas da quantidade de anotações disponíveis, mas da relação entre tipo de anotação, organização dos rótulos, perspectiva de gravação e objetivo analítico. Em consequência, a escolha da base mais adequada deve considerar não apenas o conteúdo dos dados, mas também o nível de adaptação necessário para a integrar em frameworks específicos e a robustez com que a perspectiva visual sustenta a tarefa que se pretende executar.

CONCLUSÃO

Este estudo apresentou uma análise comparativa de datasets públicos aplicados ao rastreamento e à análise comportamental de roedores, com ênfase em suas estruturas de anotação, organização dos dados e compatibilidade com diferentes frameworks computacionais. Os resultados evidenciaram que, embora exista uma ampla variedade de bases disponíveis, suas diferenças estruturais influenciam diretamente o potencial de utilização em tarefas de detecção, estimativa de pose e classificação comportamental.

Observou-se que os datasets analisados adotam diferentes formas de representação das informações, incluindo anotações baseadas em eventos temporais, rótulos por frame e pontos corporais associados à estimativa de pose. Essas diferenças impactam diretamente os níveis de compatibilidade com frameworks como *YOLO*, *DeepLabCut* e *MARS*, influenciando tanto a flexibilidade de reutilização das bases quanto sua adequação a diferentes tarefas computacionais.

De forma geral, observou-se que o Rodent3D apresenta maior potencial para aplicações voltadas à estimativa de pose e reconstrução espacial, enquanto o *RGB-D Rat* mostra-se particularmente adequado a tarefas de classificação comportamental supervisionada. Já o MARS *Dataset* destaca-se pela maior versatilidade estrutural, ao combinar informações de pose, comportamento e delimitação espacial, favorecendo sua integração com diferentes frameworks computacionais.

Além disso, verificou-se que fatores como modalidade de captura, perspectiva de gravação, quantidade de animais por cena e organização dos rótulos exercem influência significativa sobre a robustez metodológica e o potencial de generalização dos *datasets*. Nesse sentido, a escolha de uma base de dados adequada deve considerar não apenas a disponibilidade das informações, mas também sua compatibilidade estrutural com os objetivos analíticos e os modelos computacionais empregados.

Por fim, conclui-se que a diversidade estrutural dos *datasets* públicos representa simultaneamente um avanço e um desafio para pesquisas em visão computacional aplicada ao comportamento animal, reforçando a importância de análises comparativas que auxiliem na seleção mais criteriosa de bases de dados e no desenvolvimento de métodos mais robustos e reutilizáveis.

AGRADECIMENTOS

Os autores agradecem à Universidade do Estado de Mato Grosso Carlos Alberto Reyes Maldonado (UNEMAT), à Universidade Federal de Mato Grosso (UFMT), à Faculdade de Ciências Exatas e Tecnológicas (FACET) e ao Curso de Engenharia Elétrica pelo apoio institucional à realização desta pesquisa. Agradecem ainda ao Laboratório de Análise de Imagens Digitais (LAID) pelo suporte técnico e científico disponibilizado durante o desenvolvimento do estudo. Este trabalho contou com apoio do Programa de Bolsas de Iniciação Científica da UNEMAT (PROBIC), por meio do Edital nº 009/2025 – UNEMAT/PRPPG – Bolsas PROBIC – Seleção para Bolsas de Iniciação Científica.

REFERÊNCIAS

BELLO-ARROYO, E.; ROQUE, H.; MARCOS, A.; ORIHUEL, J.; HIGUERA-MATAS, A.; DESCO, M.; CAIOLFA, V. R.; AMBROSIO, E.; LARA-PEZZI, E.; GÓMEZ-GAVIRO, M. V. MouBeAT: A New and Open Toolbox for Guided Analysis of Behavioral Tests in Mice. Frontiers in Behavioral Neuroscience, [S. l.], v. 12, p. 201, 2018. Disponível em: https://doi.org/10.3389/fnbeh.2018.00201. Acesso em: 07 ago. 2025.

FRONTIERS IN BEHAVIORAL NEUROSCIENCE. Machine learning and behavior analysis: applications and perspectives in rodent studies. Frontiers in Behavioral Neuroscience, [S. l.], 2018. Disponível em: https://www.frontiersin.org/journals/behavioral-neuroscience/articles/10.3389/fnbeh.2018.00201/full. Acesso em: 07 ago. 2025.

GERÓS, A.; MAGALHÃES, A.; AGUIAR, P. Improved 3D tracking and automated classification of rodents' behavioral activity using depth-sensing cameras. Behavior Research Methods, [S. l.], 2020. Disponível em: https://link.springer.com/article/10.3758/s13428-020-01381-9. Acesso em: 13 ago. 2025.

KABRA, M.; ROBIE, A. A.; RIVERA-ALBA, M.; BRANSON, S.; BRANSON, K. JAABA: interactive machine learning for automatic annotation of animal behavior. Nature Methods, [S. l.], v. 10, p. 64–67, 2013. Disponível em: https://doi.org/10.1038/nmeth.2281. Acesso em: 09 ago. 2025.

LAUER, J.; ZHOU, M.; YE, S. et al. Multi-animal pose estimation, identification and tracking with DeepLabCut. Nature Methods, [S. l.], v. 19, p. 496–504, 2022. Disponível em: https://doi.org/10.1038/s41592-022-01443-0. Acesso em: 05 ago. 2025.

MATHIS, A.; HENRIQUEZ, J. et al. DeepLabCut: interface gráfica do gerenciador de projetos. Disponível em: https://deeplabcut.github.io/DeepLabCut/docs/gui/PROJECT_GUI.html. Acesso em: 10 ago. 2025.

NESTLER, E. J.; HYMAN, S. E. Animal models of neuropsychiatric disorders. Nature Neuroscience, [S. l.], v. 13, p. 1161–1169, 2010. Disponível em: https://pmc.ncbi.nlm.nih.gov/articles/PMC3750731/. Acesso em: 10 ago. 2025.

PATEL, M.; GU, Y.; CARSTENSEN, L. C.; HASSELMO, M. E.; BETKE, M. Animal pose tracking: 3D multimodal dataset and token-based pose optimization. International Journal of Computer Vision, [S. l.], 2022. Disponível em: https://link.springer.com/article/10.1007/s11263-022-01714-5. Acesso em: 12 ago. 2025.

SEGALIN, C.; WILLIAMS, J.; KARIGO, T.; HUI, M.; ZELIKOWSKY, M.; SUN, J. J.; PERONA, P.; ANDERSON, D. J.; KENNEDY, A. The Mouse Action Recognition System (MARS) software pipeline for automated analysis of social behaviors in mice. eLife, [S. l.], v. 10, e63720, 2021. Disponível em: https://doi.org/10.7554/eLife.63720. Acesso em: 08 ago. 2025.