

# Reptilerecon: Um Arcabouço para Extração e Análise de Sinais de Lagartos

João Gabriel Fernandes Zenóbio<sup>1</sup>, Pedro Henrique Lopes Silva<sup>1</sup>,  
Eduardo José da Silva Luz<sup>1</sup>, Gladston Juliano Prates Moreira<sup>1</sup>,  
Conrado Aleksander Barbosa Galdino<sup>2</sup>, Jadson Castro Gertrudes<sup>1</sup>

<sup>1</sup>Departamento de Ciência da Computação – Universidade Federal de Ouro Preto  
Campus Morro do Cruzeiro – Ouro Preto – MG – Brazil.

<sup>2</sup>Laboratório de Ecologia e Ecologia Comportamental de Vertebrados do  
Programa de Pós-graduação em Biologia de Vertebrados da  
Pontifícia Universidade Católica de Minas Gerais  
Av. Dom José Gaspar, 500 - Coração Eucarístico,  
30535-610, Belo Horizonte-MG, Brasil

joao.zenobio@aluno.ufop.edu.br

**Abstract.** *Animal communication is essential for survival, and researchers have dedicated themselves to studying patterns of signals emitted by lizards, particularly visual communication through head-bobbing movements. In this context, this work proposes using machine learning techniques to identify and analyze communication patterns in *Tropidurus* lizards. The methodology involves extracting signals from videos using deep learning algorithms and applying unsupervised algorithms to identify patterns in the extracted signals. The results demonstrate clear patterns in the analyzed signals.*

**Resumo.** *A comunicação animal é essencial para a sobrevivência, e os pesquisadores têm se dedicado ao estudo dos padrões de sinais emitidos por lagartos, com destaque para a comunicação visual por meio dos movimentos de cabeceios (headbobs). Nesse contexto, este trabalho propõe o uso de técnicas de aprendizado de máquina para identificar e analisar padrões de comunicação em lagartos do gênero *Tropidurus*. A metodologia consiste em extrair os sinais dos vídeos por meio de algoritmos de aprendizado profundo e, em seguida, aplicar algoritmos não supervisionados para identificar padrões nos sinais extraídos. Os resultados obtidos demonstram a presença de padrões claros nos sinais analisados.*

## 1. Introdução

A comunicação animal é um importante componente do campo de estudo do comportamento animal. Ela constitui um processo de transferência de informação entre um indivíduo emissor de um sinal e um outro indivíduo que ao receber o sinal o processa em informação [Alcock 2009]. Entre os animais a comunicação pode ter importantes funções como passar informações sobre a habilidade e força em combates empregados na defesa de recursos [Kaplan 2014, Passos 2016], a condição reprodutiva dos indivíduos e o alerta da presença ou o afugentamento de predadores. Lagartos constituem bons modelos para o estudo da comunicação por sinais visuais [Passos 2016]. Em algumas espécies o processo

de comunicação é realizado por meio do fluxo de informações através da repetição de gestos (sinais estereotipados). Dentre eles destacam-se as flexões corporais (*push-ups*) e os cabeceios (*headbobs*) [Carpenter and Ferguson 1977]. Estes sinais carregam informações sobre a identidade do indivíduo, idade, sexo e podem ser exibidos em diversos contextos sociais [Martins 1991, Martins 1993].

O campo da comunicação em lagartos é um campo ainda em crescimento (280 artigos na PubMed - `pubmed.ncbi.nlm.nih.gov`). Entretanto, os estudos têm encontrado interessantes resultados. A exemplo, o estudo de [Martins 1994] encontrou regras de utilização de sinais em lagartos do gênero *Sceloporus*, incluindo diferentes tipos de movimentos de cabeça em um único sinal. O estudo de [Passos 2016] encontrou padrões de comunicação em três espécies de lagartos do gênero *Tropidurus*, mas não explorou os componentes dos sinais emitidos.

Assim como os estudos destacados no parágrafo anterior, geralmente utiliza-se de uma abordagem supervisionada, que corresponde a anotação manual das movimentações pelo especialista, para a descrição dos tipos de sinais visuais destes animais, desde a captura de vídeo até a determinação dos cabeceios. Estas anotações manuais implicam em grande dispêndio de tempo por parte dos pesquisadores uma vez que grande quantidade de horas de vídeos devem ser analisadas. A automatização deste processo pode potencializar a produção no campo uma vez que implica em aumento da capacidade analítica por parte dos pesquisadores. Adicionalmente, o uso de ferramentas computacionais não-supervisionadas pode implicar em ganhos para o campo, como a apresentação de novos padrões não identificados pelos especialistas no domínio. [Silva 2018] propôs um *framework* para extração dos sinais emitidos por lagartos em vídeos. Contudo, a técnica utilizada no *framework* possui algumas desvantagens - não é possível detectar se o padrão de referência existe na imagem, pois o resultado da técnica é sempre uma das regiões da imagem original sem que haja a extração de informações derivada desta, e sensibilidade à variações em relação ao tamanho e orientação do objeto - que podem dificultar o processamento e a interpretação das informações extraídas dos vídeos. Além disso, o trabalho de [Silva 2018] é limitado por não realizar a segmentação e análise dos sinais emitidos pelos lagartos.

O presente trabalho busca utilizar técnicas de aprendizado de máquina para desenvolver um *framework* para detecção, segmentação e análise dos padrões de comunicação animal utilizando lagartos do gênero *Tropidurus* como objeto de estudo. A construção deste *framework* deverá contribuir para o desenvolvimento do campo de estudo da comunicação visual entre os animais.

A proposta é dividida da seguinte forma: a primeira etapa consiste na utilização de modelos de aprendizado profundo, em especial, algoritmos de detecção de objetos para identificação da posição da cabeça do lagarto e a posterior extração dos movimentos identificados ao longo do vídeo. Esta extração é apresentada por meio de um sinal, mais especificamente, uma série temporal. São empregados algoritmos de aprendizado não supervisionado com a finalidade de identificar padrões latentes presentes no conjunto de sinais obtidos a partir dos vídeos. Os resultados obtidos a partir das extrações realizadas por meio dos algoritmos de aprendizado não supervisionados apresentam bons índices de validação, além de padrões interessantes existentes no conjunto de sinais segmentados.

## 2. Fundamentação teórica

Nesta seção, são apresentados os principais conceitos para o pleno entendimento da proposta apresentada no presente trabalho.

### 2.1. Detecção de objetos

A detecção de objetos em aprendizado de máquina é uma técnica que permite identificar a presença e localização de objetos específicos em imagens ou vídeos [Goodfellow et al. 2016]. Para tal, são utilizados algoritmos de aprendizado de máquina, como redes neurais convolucionais (*Convolutional Neural Network* - CNN), que precisam ser treinadas com um grande conjunto de imagens rotuladas para aprender a detectar padrões específicos nas imagens.

A arquitetura YOLO (*You Only Look Once*) [Redmon et al. 2015], adotada no presente trabalho, é uma CNN para detecção de objetos em imagens em tempo real. Ao contrário de outras abordagens de detecção de objetos que usam várias etapas e modelos para detectar objetos em uma imagem, a YOLO usa uma única CNN para prever diretamente as caixas delimitadoras e as classes dos objetos em uma única etapa.

### 2.2. Segmentação e transformação de séries temporais

Séries temporais são conjuntos de observações feitas em intervalos regulares e ordem cronológica [Montgomery et al. 2015, Fu 2011]. Caracterizam-se por serem numéricas, contínuas e de alta dimensionalidade. A análise de séries temporais identifica padrões e tendências ao longo do tempo, sendo aplicada em áreas como mineração de dados e análise preditiva. Neste estudo, o objetivo é detectar padrões nos sinais visuais emitidos por lagartos.

Conforme [Fu 2011], a mineração de dados de séries temporais envolve representação, indexação, medida de similaridade, segmentação e visualização. A segmentação pode ser uma etapa de pré-processamento ou técnica de análise de tendências. A segmentação de tamanho fixo proposta em [Das et al. 1998] utiliza uma janela deslizante para dividir a série temporal em segmentos, podendo haver sobreposição entre fragmentos. Neste trabalho, será adotada a janela deslizante com sobreposição

Após a segmentação de uma série temporal, é possível transformá-la em um espaço latente  $n$ -dimensional, utilizando redes neurais artificiais, especialmente os *autoencoders* [Goodfellow et al. 2016]. Os *autoencoders* são uma técnica de aprendizado não supervisionado amplamente utilizada para modelar séries temporais, compostos por um codificador e um decodificador [Chiarot and Silvestri 2023, Goodfellow et al. 2016]. A representação latente obtida por meio do *autoencoder* permite reconstruir a série original ou gerar novas amostras, facilitando comparações eficientes entre sinais na comunicação dos lagartos.

### 2.3. Agrupamento de dados

A análise de *cluster* (ou agrupamento de dados) consiste em dividir um conjunto de dados em grupos/*clusters* significativos, em que um par de objetos no mesmo *cluster* compartilha características semelhantes quando comparado com um par de objetos em outros.

Os algoritmos de agrupamento tendem a encontrar *clusters* usando abordagens diferentes. Entre eles, podemos citar o agrupamento particional e hierárquico. Os algoritmos de agrupamento particionais geralmente agrupam os dados de acordo com um parâmetro especificado pelo usuário  $k$ , em que  $k$  representa o número de grupos que devem ser encontrados no conjunto de dados [Jain and Dubes 1988].

Nos algoritmos de agrupamento hierárquicos, não há uma partição explícita do conjunto de dados, mas sim uma representação hierárquica do mesmo. Essa hierarquia pode ser construída de forma divisiva ou aglomerativa. No primeiro caso, há divisões sucessivas do conjunto de dados até restar apenas um objeto em cada *cluster* (*singleton*). Os algoritmos aglomerativos, mais comuns para tarefas de agrupamento, realizam o oposto, começando pelos *singletons* e mesclando *clusters* semelhantes dois a dois até criar um único *cluster*. Este é o caso da ligação simples (*Single linkage*), ligação completa (*Complete linkage*), ligação média (*Average linkage*), ligação por média ponderada (*Weighted linkage*), método de Ward, entre outros [Jain and Dubes 1988].

A hierarquia produzida por um algoritmo de agrupamento pode ser visualizada usando um dendrograma. Para especificar uma partição nesse cenário, comumente é feito um corte horizontal no dendrograma que leva ao mesmo número de grupos  $k$ , especificado pelo usuário. No entanto, essa metodologia tem limitações, como a impossibilidade de selecionar *clusters* em diferentes níveis hierárquicos. Algumas implementações de algoritmos de agrupamento hierárquico permitem que o usuário especifique o parâmetro  $m_{ClSize}$ , que é o tamanho mínimo que os *clusters* relevantes devem ter dentro de uma hierarquia. Nesse caso, qualquer *cluster* menor que  $m_{ClSize}$  é automaticamente descartado como um ruído [Campello et al. 2013]. O uso do parâmetro  $m_{ClSize}$  produz uma hierarquia simplificada que pode ser usada posteriormente para extração de *clusters* ótimos.

[Campello et al. 2013] propuseram um *framework* para o pós-processamento ideal de uma hierarquia de agrupamento simplificada, chamado de FOSC, que pode extrair uma partição plana realizando cortes locais na hierarquia, para selecionar uma coleção de *clusters* não sobrepostos e com possibilidade de identificar objetos de ruído durante o processo. Isso é possível graças a um critério não supervisionado ou semi-supervisionado fornecido. Em nossos experimentos, foi adotada a medida de qualidade não supervisionada *Stability* para extração ótima de *clusters*. De maneira geral, a medida em questão tem como objetivo avaliar a duração do ciclo de vida de um *cluster*, desde sua criação até sua eventual divisão em *clusters* válidos ou até que todos os objetos a ele associados sejam classificados como ruído. Para maiores detalhes, sugerimos consultar a fonte [Campello et al. 2013].

### 3. Trabalhos relacionados

Esta seção apresenta os principais trabalhos relacionados ao presente estudo. Vale ressaltar que não foram encontrados muitos trabalhos que envolvessem a aplicação de algoritmos de aprendizado de máquina aos estudos em dados de sinais emitidos por lagartos.

Os principais trabalhos, tomados com base neste documento, são os trabalhos de [Martins 1994] e [Passos 2016], em que ambos estudam padrões de sinais emitidos por lagartos. Especificamente, [Martins 1994] realiza o estudo do *Sceloporus graciosus* considerando três componentes da comunicação dos lagartos: (i) o número e o tipo de cabeceio, (ii) o número de pernas flexionadas e estendidas para produzir movimentos de flexão

e (iii) o uso de posturas do corpo expositor. Em todos os aspectos, ela obteve padrões que, combinados, podem representar uma estrutura de linguagem do lagarto *S. graciosus*. Considerando o aspecto (i), [Martins 1994] também observou que os lagartos eram capazes de produzir, dentro de um sinal de comunicação, diferentes tipos de movimentação de cabeça (segmentos), tais como movimentos simples (*single-headbob*) e movimentos duplos (*double head-bobs*), o que poderiam indicar a existência de um padrão gramatical no movimento da cabeça.

[Ord and Martins 2006] discutem como a seleção natural e sexual trabalham juntas para moldar os sistemas comunicativos dos animais dentro de uma mesma espécie, produzindo um fenótipo complexo que funciona bem em vários contextos. Os autores examinam como um regime seletivo intrincado pode moldar a diversificação de um comportamento complexo entre espécies. O trabalho usa métodos comparativos filogenéticos e um conjunto de dados de mais de 50 fontes para testar a importância relativa de diferentes variáveis ecológicas na evolução de exibições visuais de lagartos *Anolis*. Os resultados mostram que diferentes forças seletivas estão associadas a mudanças em diferentes componentes da exibição. O estudo também destaca a influência das distinções ecomórficas na evolução da estrutura da exibição.

Em [Passos 2016], foram avaliadas as exibições visuais do cabeceio (*headbob*) de três espécies do gênero *Tropidurus*: *T. helenae*, *T. jaguaribanus* e *T. pinima*. Neste trabalho também foram identificados padrões de comunicação emitidos pelas diferentes espécies. Entretanto, diferente do trabalho de [Martins 1994], não foi realizada uma identificação dos diferentes tipos de movimento (*tokens*) que poderiam existir dentro de um sinal emitido pelos lagartos.

A respeito a tarefa de extração de informações de padrão de comunicação de lagartos, destacamos aqui os estudos realizados por [Silva 2018]. No trabalho foi desenvolvido um sistema para armazenamento de vídeos, rotulação do objeto a ser detectado (cabeça do lagarto) e extração das informações de movimento, denominado *Reptilerecon*. Na implementação do sistema foi utilizada a técnica *Template Matching* com o método de casamento por coeficiente de correlação normalizado, desenvolvido por [Briechle and Hanebeck 2001], para detecção da cabeça do lagarto e para extração das informações de movimentos verticais detectados nos vídeos.

Um dos problemas identificados no trabalho de [Silva 2018] é que a técnica utilizada para detecção da cabeça do lagarto é passível de uma alta taxa de erro, o que pode dificultar o processamento e a interpretação das informações extraídas dos vídeos. Além disso, a cada utilização da técnica é necessário a intervenção humana para geração do *template* a ser encontrado na imagem. Dessa forma, o estudo de [Silva 2018] não soluciona a questão da automação na análise dos cabeceios dos lagartos. Portanto, é necessário aprimorar esse sistema com novas técnicas de aprendizado para detectar e extrair informações mais confiáveis.

O trabalho de [Nelson and Ord 2022] apresenta um estudo sobre como os sinais de comunicação de lagartos machos podem ser usados para identificar a espécie e a população à qual eles pertencem. Foram analisadas, por meio de árvores aleatórias (*Random Forests*), 11 espécies de lagartos *Anolis*, das quais muitas coexistem no mesmo ambiente. Os resultados mostraram que é possível identificar corretamente a maioria dos indivíduos

com base em apenas alguns sinais de comunicação. No entanto, alguns indivíduos foram mal identificados devido à semelhança nos sinais de comunicação entre espécies relacionadas ou adaptadas a ambientes semelhantes. O presente estudo assemelha-se ao trabalho em questão por empregar técnicas de aprendizado de máquina para a classificação de sinais. No entanto, os autores não se propuseram a efetuar a extração automática nem a segmentação de sinais, com o intuito de identificar distintos tipos de movimento presentes nos sinais emitidos pelos lagartos.

## 4. Metodologia experimental e resultados

Esta seção descreve os experimentos e resultados obtidos na análise dos sinais dos lagartos. A subseção 4.1 aborda a construção do banco de dados; 4.2 detalha o pré-processamento, envolvendo extração e tratamento dos sinais, e segmentação; 4.3 apresenta o agrupamento e validação dos dados; e 4.4 discute os resultados e análise dos grupos. Os experimentos foram executados em servidores com processador Intel 8 núcleos, 64GB RAM DDR4 e GPU NVIDIA 12GB RAM DDR4 com mais de 3 mil cuda cores.

### 4.1. Dados

O projeto utiliza uma base de dados de vídeos de comunicação visual de répteis do gênero *Tropidurus*, fornecidos por Daniel Passos, pesquisador da Universidade Federal Rural do Semi-Árido em Mossoró, Rio Grande do Norte. As filmagens em alta definição, a 30 fps, com câmeras e tripés, garantiram distância adequada e estabilidade na captação. Foram disponibilizados 46 vídeos, sendo 44 utilizados, cada um com um único lagarto, seguindo procedimentos de [Passos 2016]. Os demais vídeos, com comunicação de múltiplos lagartos, não foram tratados neste trabalho.

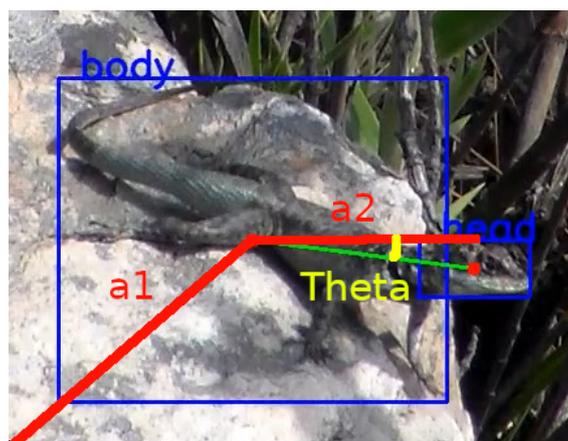
### 4.2. Pre-processamento

Após a coleta e curadoria dos vídeos, a fase de pré-processamento foi executada, abrangendo três etapas: (i) extração dos sinais emitidos pelos répteis; (ii) segmentação dos sinais; e (iii) compressão otimizada dos sinais.

Na etapa (i), a rede neural YOLOv4 [Bochkovskiy 2021] foi treinada para detectar a cabeça e o corpo do animal. Nesse processo, 600 quadros dos vídeos fornecidos foram rotulados manualmente com o software LabelImg (<https://github.com/tzutalin/labelImg>), demarcando as caixas delimitadoras correspondentes à localização da cabeça e do tronco do lagarto. Os dados foram divididos em 90% para treino e 10% para teste.

O treinamento da rede YOLO utilizou transferência de aprendizado e *fine tuning*, ajustando os pesos de uma rede neural pré-treinada para detectar objetos customizados. Os pesos pré-treinados e o algoritmo para treino e teste foram fornecidos pelos autores do YOLOv4 [Bochkovskiy 2021]. A transferência de aprendizado e *fine tuning* foi aplicada à partição de treinamento. É importante ressaltar que durante o processo de *fine tuning*, optamos por manter os hiperparâmetros padrão da arquitetura.

Com a rede treinada, foram extraídos os centros de posição do corpo e da cabeça de todos os vídeos disponíveis na partição de teste, e calculada a angulação relativa ( $\theta$ ) entre os centros em cada quadro, conforme ilustrado na Figura 1.



**Figura 1. Cálculo da angulação relativa ( $\theta$ ) entre a cabeça e o corpo do lagarto.**

Com base nas variáveis apresentadas na Figura 1, a angulação foi calculada conforme descrito por [Spong et al. 2005]:

$$\theta = \arctan\left(\frac{\sqrt{1 - D^2}}{D}\right), \quad (1)$$

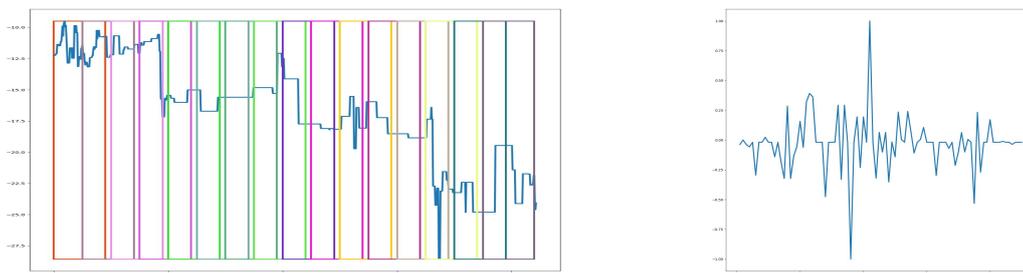
em que  $D$  é dado por

$$D = \frac{x^2 + y^2 - a_1^2 - a_2^2}{2 \times a_1 \times a_2}. \quad (2)$$

Na Equação 2,  $a_1$  e  $a_2$  representam, respectivamente, as distâncias em pixels entre o canto inferior esquerdo do vídeo e o centro do corpo do lagarto e entre o centro do corpo e o centro da cabeça do lagarto. Já  $x$  e  $y$  denotam a posição atual em pixels da cabeça do lagarto. O resultado da equação,  $\theta$ , expressa o ângulo formado entre o centro da cabeça e o centro do corpo do lagarto. Esse método garante invariância às movimentações do operador da câmera, concentrando-se na movimentação do animal, independentemente de quaisquer mudanças na câmera ou na orientação do lagarto durante a filmagem.

Ao executar o programa, os valores de angulação são registrados somente quando duas condições empíricas são satisfeitas: (i) ocorre movimentação do centro da cabeça; e (ii) essa movimentação é suficientemente significativa, de modo que o novo centro esteja a mais de 1 pixel de distância do anterior. Essa abordagem, que filtra ruídos indesejáveis, constitui uma contribuição relevante deste trabalho e um possível avanço para o campo.

Na etapa (ii) do pré-processamento, a segmentação dos sinais foi executada a fim de investigar os padrões sutis de comunicação visual dos lagartos do gênero *Tropidurus*. Para tal, empregamos a técnica de janela deslizante, ilustrada na Figura 2(a). Nos experimentos, a janela consistiu em 90 *frames*, correspondendo a 3 segundos de vídeo. Esse parâmetro foi estabelecido com base na análise empírica dos sinais por um especialista no domínio. O passo de deslizamento da janela foi definido como 10 *frames* para assegurar que nenhum padrão fosse desconsiderado e para reduzir a quantidade de dados redundantes. Após a segmentação, 8.827 novos sinais foram gerados.



(a) Técnica da janela deslizante (passos de tamanho 50 *frames*, apenas para demonstração).

(b) Partição do sinal após retirar o módulo e a primeira derivada dos valores.

**Figura 2. Apresentação de transformações realizadas na etapa (ii) do pré-processamento dos sinais.**

Posteriormente, o módulo e a primeira derivada dos valores dos novos sinais foram extraídos, permitindo a análise das movimentações da cabeça independentemente da orientação e pose do lagarto no vídeo, conforme ilustrado na Figura 2(b). Por fim, os dados foram normalizados no intervalo entre -1 e 1, preservando os sinais relativos ao afastamento (positivo) e à aproximação (negativo) e possibilitando a comparação entre diferentes conjuntos de dados do banco.

A etapa (iii) do pré-processamento consistiu na compressão inteligente dos sinais segmentados, para posterior utilização pelos algoritmos de agrupamento de dados. Para essa compressão foi utilizada uma rede do tipo *autoencoder* com camadas *Long Short-term Memory* [Hochreiter and Schmidhuber 1997]. A compressão é realizada pela parte da rede chamada codificador e a descompressão é realizada pela parte chamada decodificador. Após o treinamento da rede, foi realizada uma inferência e coletado o dado comprimido na saída do codificador, com uma dimensão de tamanho 20. Ao final das três etapas de pré-processamento, foi obtido um conjunto de dados com 8.827 exemplos e cada um com dimensão de tamanho 20.

### 4.3. Extração de padrões

Após a compressão inteligente, os dados foram agrupados utilizando o *framework* FOSC. A extração da hierarquia de grupos foi realizada utilizando os dendrogramas produzidos pelos métodos de ligação mínima (*Single linkage*), ligação completa (*Complete linkage*), por média (*Average linkage*), por média ponderada (*Weighted Average*) e o método de ligação de Ward. Cada uma das hierarquias geradas foi apresentada ao *framework*, que realizou a extração ótima dos grupos com a seguinte variação do parâmetro  $min_{ClSize}$ : 32, 48, 64, 96, 128, 192 e 256.

Os resultados das instâncias dos algoritmos foram avaliados pelas métricas de avaliação Área Sob Curva para Agrupamento (AUCC) [Jaskowiak et al. 2022] e Pontuação de Silhueta [Rousseeuw 1987]. Estas métricas revelam se há alta similaridade entre objetos classificados em um mesmo grupo e baixa similaridade entre objetos classificados em grupos distintos, de forma a averiguar se o agrupamento se deu de forma lógica de acordo com as distâncias entre os valores dos objetos. Quanto maior os valores dos índices, maior

o indicativo de uma partição válida. Também foi calculado o número de ruídos definidos pelo FOSC e utilizado como penalidade nos resultados obtidos pelos índices, conforme detalhado nos trabalhos de [Moulavi et al. 2014, Gertrudes et al. 2019].

#### 4.4. Pós-processamento

A Tabela 1 mostra os melhores resultados do FOSC usando diferentes combinações de métodos de ligação e  $min_{ClSize}$ . Apenas o método *Average linkage* apresenta  $min_{ClSize} = 96$ , enquanto os outros têm  $min_{ClSize} = 32$ . Levando em conta índices de validação e ruídos, o método *Average linkage* com  $min_{ClSize} = 96$  foi escolhido para análise, apresentando melhores índices AUCC e Silhueta e menos ruídos que o método *Weighted*.

Modelo de ligação	Melhor $min_{ClSize}$	AUCC	Silhueta	Num. Ruídos
Average	96	<b>0.97</b>	0.43	86
Weighted	32	0.91	<b>0.55</b>	288
Complete	32	0.92	0.49	127
Single	32	0.93	0.34	0
Ward	32	0.86	0.31	0

**Tabela 1. Performance dos algoritmos de agrupamento obtidos sobre os sinais transformados via autoencoders.**

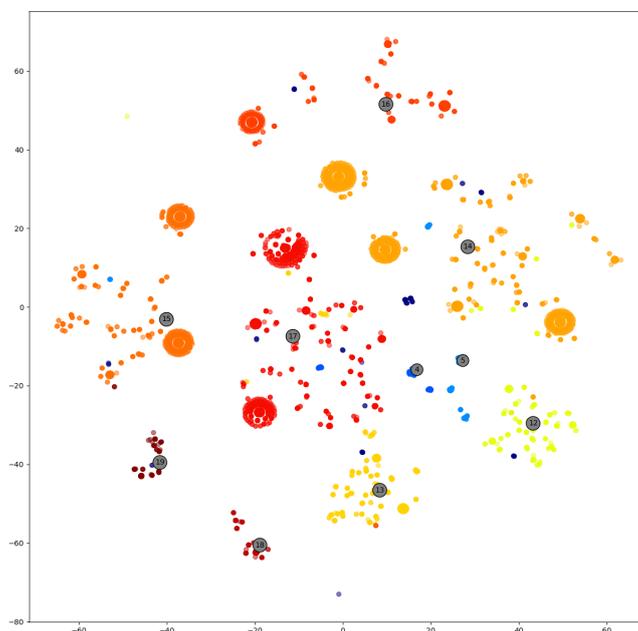
O resultado obtido pelo método de *Average linkage* produziu 10 grupos que podem ser visualizados na Figura 3. Para visualizar os grupos, foi utilizado o método incorporação de vizinhos estocásticos T-Distribuídos (*t-Distributed Stochastic Neighbor Embedding* t-SNE) [Van der Maaten and Hinton 2008] que diminuiu a dimensão dos dados comprimidos pela rede *autoencoder* para duas dimensões.

No contexto de agrupamento de dados, os resultados obtidos apresentam-se extremamente promissores, indicando a existência de grupos que compartilham características semelhantes e que se distinguem claramente dos demais grupos. Esses resultados podem sugerir a existência de padrões nos sinais emitidos pelos lagartos, que podem estar relacionados a palavras ou frases isoladas emitidas pelos animais.

Com o intuito de simplificar a análise desses resultados por especialistas no domínio, realizou-se a projeção de sinais exemplares de cada *cluster*, conforme proposto por [McInnes et al. 2017]. Esses sinais foram selecionados levando em consideração o tempo de vida do *cluster*, antes que todos os objetos fossem classificados como ruído ou antes que o *cluster* dividisse em novos *clusters* válidos. O resultado obtido é ilustrado na Figura 4. Através da análise das Figuras 4(a) a 4(j), é possível constatar que de fato existem padrões de sinais presentes em cada um dos grupos definidos pelo método *Average linkage*. No entanto, até o momento, o significado desses sinais permanece desconhecido. Portanto, uma investigação mais aprofundada sobre os padrões de sinais caberá ao especialista na área.

## 5. Conclusões

O propósito deste estudo foi contribuir para a identificação de possíveis padrões nos sinais emitidos por lagartos do gênero *Tropidurus*. Para alcançar esse objetivo, foram executadas diversas etapas do processo de análise de dados, incluindo extração de sinais de vídeos,



**Figura 3. Visualização dos agrupamentos obtidos pelo método de ligação Average com parâmetro  $m_{ClSize} = 96$ .**

segmentação de sinais e análise dos sinais segmentados. Os resultados obtidos demonstraram que é possível identificar padrões de sinais visuais utilizando técnicas de análise de sinais e agrupamento em algoritmos computacionais. Isso poderá ser de grande valia para os pesquisadores na área, uma vez que permite evitar a análise manual dos sinais, melhorando a qualidade e quantidade de análises realizadas, possibilitando a observação de milhares de dados simultaneamente.

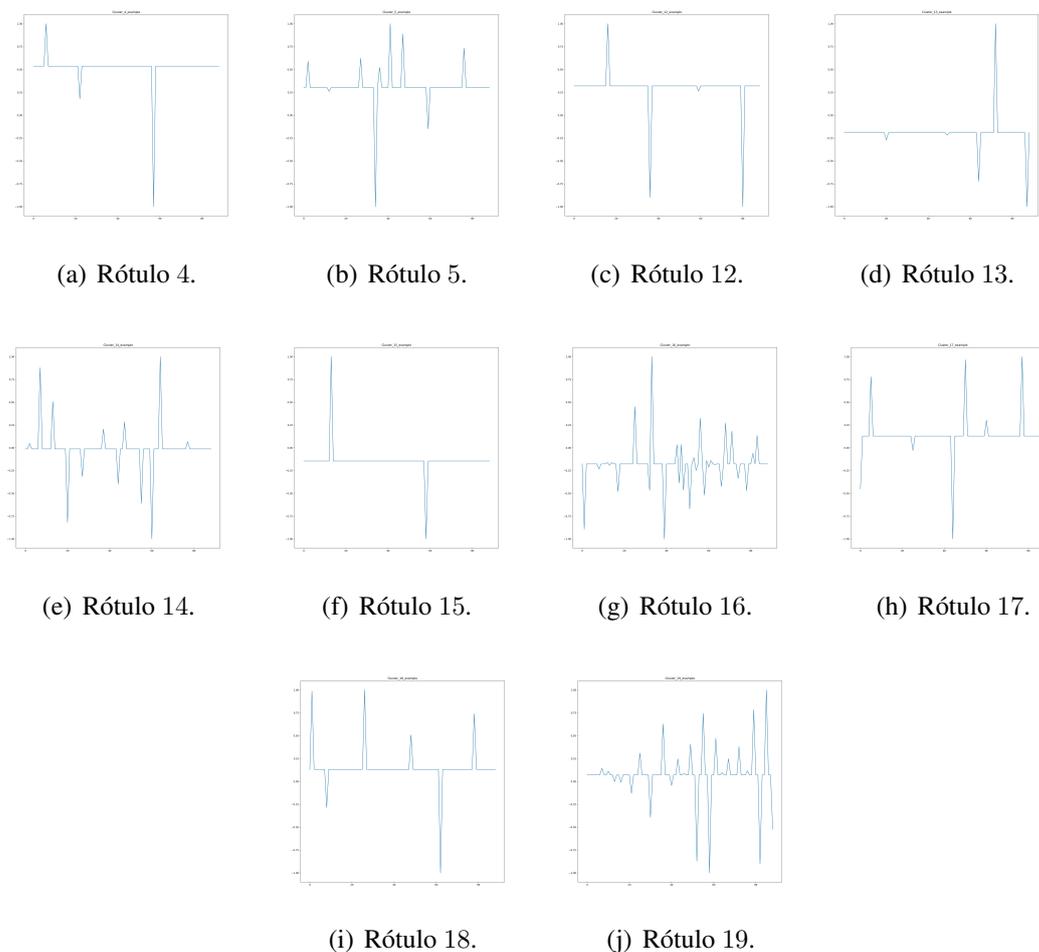
Para trabalhos futuros, pretendemos aprofundar os estudos nas etapas definidas neste projeto. Na etapa de extração dos sinais, buscaremos refinamentos para o processo considerando a movimentação do lagarto no vídeo. Na etapa de segmentação dos sinais, exploraremos a possibilidade de segmentar os sinais em diferentes tamanhos, considerando que possam existir padrões emitidos pelos lagartos com diferentes tempos de movimento, conforme sugestão do especialista no domínio. Por fim, no pós-processamento, realizaremos estudos para apresentar os resultados de forma mais clara ao especialista.

### Agradecimentos

Gostaríamos de agradecer ao Programa de Iniciação à Pesquisa (PIP) da Universidade Federal de Ouro Preto, à CAPES e à FAPEMIG pelo apoio financeiro concedido.

### Referências

- Alcock, J. (2009). *Animal behavior: An evolutionary approach*. Sinauer associates.
- Bochkovskiy, A. (2021). Yolo v4, v3 and v2 for Windows and Linux.



**Figura 4. Exemplos de objetos extraídos dos *clusters*, obtidos pelo FOSC a partir do dendrograma produzido pelo método *Average linkage*.**

- Briechele, K. and Hanebeck, U. D. (2001). Template matching using fast normalized cross correlation. In *Optical pattern recognition XII*, volume 4387, pages 95–102. tex.organization: International Society for Optics and Photonics.
- Campello, R. J. G. B., Moulavi, D., Zimek, A., and Sander, J. (2013). A framework for semi-supervised and unsupervised optimal extraction of clusters from hierarchies. *Data Min. Knowl. Discov.*, 27(3):344–371. tex.xxxdoi: 10.1007/s10618-013-0311-4.
- Carpenter, C. C. and Ferguson, G. W. (1977). Variation and evolution of stereotyped behavior in reptiles. *Biology of the Reptilia*, 7:335–554.
- Chiarot, G. and Silvestri, C. (2023). Time series compression survey. *ACM Computing Surveys*, 55(10):1–32. Publisher: ACM New York, NY.
- Das, G., Lin, K.-I., Mannila, H., Renganathan, G., and Smyth, P. (1998). Rule discovery from time series. In *KDD*, volume 98, pages 16–22. Number: 1.
- Fu, T.-c. (2011). A review on time series data mining. *Engineering Applications of Artificial Intelligence*, 24(1):164–181.

- Gertrudes, J. C., Zimek, A., Sander, J., and Campello, R. J. G. B. (2019). A unified view of density-based methods for semi-supervised clustering and classification. *Data Mining and Knowledge Discovery*, 33(6):1894–1952.
- Goodfellow, I. J., Bengio, Y., and Courville, A. C. (2016). *Deep learning*. Adaptive computation and machine learning. MIT Press.
- Hochreiter, S. and Schmidhuber, J. (1997). Long short-term memory. *Neural Computation*, 9(8):1735–1780. tex.eprint: <https://direct.mit.edu/neco/article-pdf/9/8/1735/813796/neco.1997.9.8.1735.pdf>.
- Jain, A. K. and Dubes, R. C. (1988). *Algorithms for clustering data*. Prentice-Hall.
- Jaskowiak, P. A., Costa, I. G., and Campello, R. J. G. B. (2022). The area under the ROC curve as a measure of clustering quality. *Data Mining and Knowledge Discovery*, 36(3):1219–1245.
- Kaplan, G. (2014). Animal communication. *Wiley Interdisciplinary Reviews: Cognitive Science*, 5(6):661–677. Publisher: Wiley Online Library.
- Martins, E. P. (1991). Individual and sex differences in the use of the push-up display by the sagebrush lizard, *Sceloporus graciosus*. *Animal Behaviour*, 41(3):403–416. Publisher: Elsevier.
- Martins, E. P. (1993). Contextual use of the push-up display by the sagebrush lizard, *Sceloporus graciosus*. *Animal Behaviour*, 45(1):25–36. Publisher: Elsevier.
- Martins, E. P. (1994). Structural complexity in a lizard communication system: the *Sceloporus graciosus* "push-up" display. *Copeia*, pages 944–955. Publisher: JSTOR.
- McInnes, L., Healy, J., and Astels, S. (2017). hdbscan: Hierarchical density based clustering. *J. Open Source Softw.*, 2(11):205.
- Montgomery, D. C., Jennings, C. L., and Kulahci, M. (2015). *Introduction to time series analysis and forecasting*. John Wiley & Sons.
- Moulavi, D., Jaskowiak, P. A., Campello, R. J. G. B., Zimek, A., and Sander, J. (2014). Density-based clustering validation. In *Proc. SDM*, pages 839–847. tex.xxxxdoi: 10.1137/1.9781611973440.96.
- Nelson, C. M. and Ord, T. J. (2022). Identifying potential cues of species identity in complex animal signals. *Animal Behaviour*, 186:121–136. Publisher: Elsevier.
- Ord, T. J. and Martins, E. P. (2006). Tracing the origins of signal diversity in anole lizards: phylogenetic approaches to inferring the evolution of complex behaviour. *Animal Behaviour*, 71(6):1411–1429. Publisher: Elsevier.
- Passos, D. C. (2016). *Área de vida, organização social e comunicação visual de Tropidurus do grupo semitaeniatus (Squamata: Tropiduridae)*. Tese de doutorado, Universidade do Estado do Rio de Janeiro.
- Redmon, J., Divvala, S. K., Girshick, R. B., and Farhadi, A. (2015). You only look once: Unified, real-time object detection. *CoRR*, abs/1506.02640. tex.eprint: 1506.02640.
- Rousseeuw, P. J. (1987). Silhouettes: a graphical aid to the interpretation and validation of cluster analysis. *Journal of computational and applied mathematics*, 20:53–65. Publisher: Elsevier.

- Silva, A. O. (2018). *Framework para extração de sinais na comunicação visual de lagartos*. Trabalho de conclusão de curso, Universidade Federal de Ouro Preto.
- Spong, M., Hutchinson, S., and Vidyasagar, M. (2005). *Robot modeling and control*. Wiley select coursepack. Wiley.
- Van der Maaten, L. and Hinton, G. (2008). Visualizing data using t-SNE. *Journal of machine learning research*, 9(11).