



## Essa aranha é peçonhenta? Eficiência de um aplicativo de reconhecimento de imagem em reconhecer espécies de aranhas sinantrópicas

**Palavras-chave:** iEcology, inteligência artificial, aranhas peçonhentas.

Autores:

Isabela Maria Tosta Fernandes Pereira, IB - UNICAMP

Prof<sup>o</sup> Dr<sup>o</sup> Raul Costa Pereira (orientador), IB- UNICAMP

Departamento de Biologia Animal (DBA)

Laboratório de Ecologia das Variações Individuais (LEVIIn)

---

### Introdução

Identificar taxonomicamente ou funcionalmente organismos na natureza é um desafio complexo que humanos enfrentam desde os primórdios de nossa espécie. Por exemplo, a capacidade de detectar se um determinado animal é ou não perigoso (e.g., peçonhento ou um predador) certamente foi importante na evolução de nossa espécie e em como exploramos ambientes naturais (CARO, 2005). É verdade que esse desafio se tornou menos importante para humanos com a vida em cidades, onde apenas uma parcela pequena das espécies animais potencialmente perigosas à saúde humana pode ser encontrada. Apesar disso, por exemplo, cidades tropicais, em particular, abrigam um número não desprezível de animais com interesse médico, como por exemplo, *Latrodectus curacaviensis*, *Loxosceles amazonica* e *Phoneutria nigriventer*. Democratizar o acesso a informações confiáveis sobre organismos perigosos à vida humana é ainda um desafio vigente que repousa na intersecção da ecologia, taxonomia, saúde pública e comunicação.

Ecólogos recentemente propuseram o novo campo, denominado *iEcology*, que estuda padrões e processos ecológicos utilizando bancos de dados de redes sociais ou outras plataformas na internet (BARTA, 2023; JARIĆ, 2020). A expansão do acesso a imagens digitais, combinada com técnicas avançadas de inteligência artificial (IA) para classificação de imagens, oferece uma oportunidade promissora para que ecólogos tenham acesso a novos conjuntos de dados sobre, por exemplo, a ocorrência de espécies (AUGUST *et al*, 2020).

Uma vez que ferramentas de IA's podem ser usadas para reconhecer espécies em diferentes situações e ambientes, essas tecnologias emergentes podem ser aliadas importantes não só no estudo sobre a distribuição da biodiversidade, mas também para buscar entender as relações entre organismos sinantrópicos e vidas humanas em cidades, inclusive em um contexto de prever e prevenir casos de acidentes com organismos perigosos comuns em cidades, como aranhas peçonhentas.

Aranhas de interesse médico foram responsáveis por aproximadamente 170.000 casos notificados por animais peçonhentos entre 2017 e 2021, com a maioria dos registros nas regiões sul e sudeste do país (MINISTÉRIO DA SAÚDE, 2022). Neste mesmo período, houve um aumento nos casos de acidentes com aranhas peçonhentas, frequentemente em ambientes urbanos e periurbanos. Esses acidentes não só causam desconforto e sofrimento para as vítimas, mas também apresentam um desafio para o desenvolvimento de

estratégias de saúde pública em cidades brasileiras. Muitas vezes, a identificação precisa das aranhas é crucial para um tratamento adequado e para estratégias eficazes de controle. Nesse sentido, popularizar o acesso às informações sobre como identificar aranhas peçonhentas é fundamental para prevenir e tratar eficientemente acidentes.

O objetivo do presente trabalho é testar a eficiência de um aplicativo popular de reconhecimento de imagens baseado em métodos de IA, o *Google Lens*, como ferramenta para identificar as espécies de aranhas mais comuns em cidades brasileiras, tanto peçonhentas quanto não peçonhentas. Para isso, usei fotos identificadas por especialistas em uma plataforma de ciência cidadã, e testei se a IA do *Google Lens* era capaz de chegar a mesma ou similar identificação, tanto em termos taxonômicos quanto em termos do interesse de saúde pública. Comparei ainda se a precisão na identificação do *Google Lens* variava entre aranhas peçonhentas e não peçonhentas. Em um contexto geral, avaliei se a tecnologia de reconhecimento de imagem pode auxiliar na identificação rápida e precisa de aranhas, contribuindo para a redução dos acidentes e favorecendo estratégias modernas de saúde pública.

### **Metodologia**

O primeiro passo foi fazer obter fotos de aranhas tiradas por cidadãos não especialistas, mas que passaram por identificação taxonômica e validação por especialistas. Para isso, usei fotos do *iNaturalist*, uma plataforma de ciência cidadã, com coleções de dados de ocorrência de espécies documentadas em fotos e, normalmente, georreferenciadas. A identificação de espécies na plataforma é realizada por especialistas e tem uma precisão próxima de 80% (VAN HORN, 2018).

Busquei imagens validadas (i.e., listadas como *Research grade* no aplicativo) com base na lista de espécies de aranhas mais comuns dentro ou no entorno de residências no Brasil, (INSTITUTO BUTATAN; SÃO PAULO; SPERANZA *et al.*, 2020). A lista contém 7 espécies, sendo três delas de importância médica (i.e., peçonhentas para humanos): viúva flamenguinha (*Latrodectus curacaviensis*), aranha-marrom (*Loxosceles amazonica*) e armadeira (*Phoneutria nigriventer*), e quatro delas que não são peçonhentas para humanos: papa-moscas (*Menemerus bivittatus*), aranha-caseira-vermelha (*Nesticodes rufipes*), aranha treme-treme (*Pholcus phalangioides*) e aranha-cuspideira (*Scytodes fusca*).

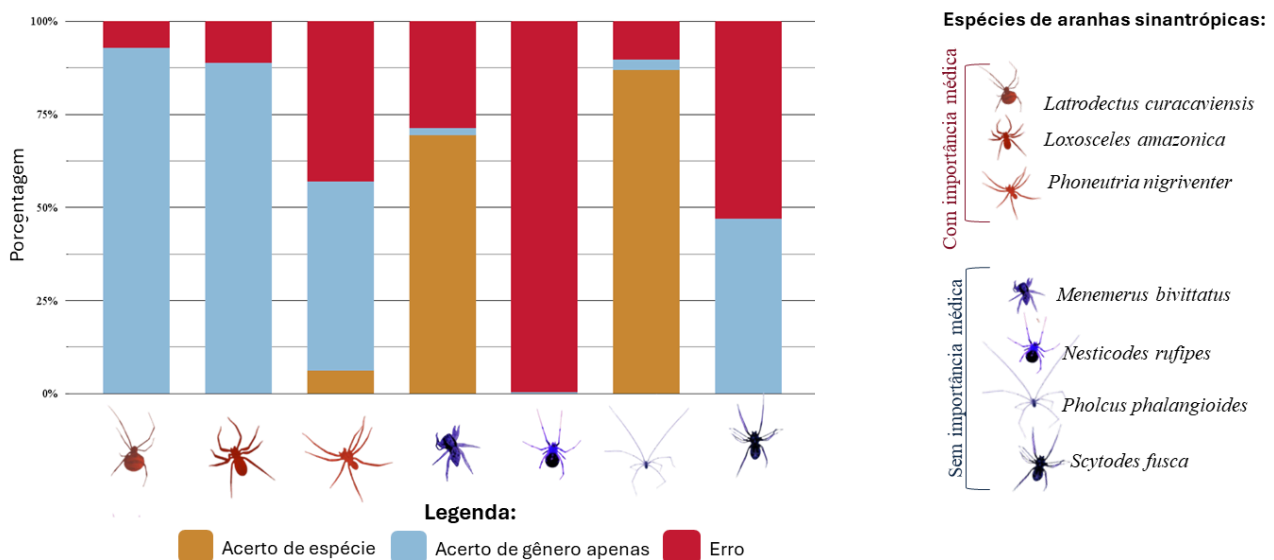
Coletei fotos de 100 postagens no *iNaturalist* de cada espécie, totalizando 700 postagens e 1157 fotos, já que uma mesma postagem pode apresentar mais de uma foto de uma mesma aranha. Apenas coletei e contabilizei fotos diferentes, fotos iguais ou repetitivas foram excluídas da amostra. Após a coleta das fotos com identificação validada, fiz a busca de cada foto na pesquisa por imagem do *Google*, utilizando o *Google Lens*. Classifiquei os resultados gerados pelo aplicativo para cada foto como: (i) identificação correta da espécie, (ii) identificação incorreta, mas com acerto apenas do gênero, (iii) se o nome identificado era de aranha e quando era, (iv) se era uma aranha de importância médica. Para testar se a precisão na identificação do *Google Lens* variava entre aranhas peçonhentas e não peçonhentas, fiz um teste qui-quadrado, com nível de significância de 5%, no software estatístico R.

### **Resultados e Discussão**

Considerando todas as fotos coletadas (n = 1157), 21,35% tiveram suas espécies devidamente identificadas pelo *Google Lens*, em nível de espécie. A maioria desses acertos foi registrada para a aranha-pernuda *P. phalangioides* (87% de acerto), enquanto para quatro espécies, duas de interesse médico (*Latrodectus curacaviensis* e *Loxosceles amazonica*) e duas não peçonhentas (*Nesticodes rufipes* e *Scytodes fusca*), o aplicativo não acertou a identificação a nível de espécie de nenhuma foto (Figura 1). Nesse contexto, a porcentagem de erros de identificação, que se aproxima de 80%, sugere que ferramentas de IA ainda estão longe de ser úteis para esse fim. Porém, uma parte considerável desses erros está associada a acerto no gênero, mas erro no epíteto específico.

Considerando as aranhas de interesse médico, ao qual pertencem todas as aranhas dos gêneros *Latrodectus*, *Loxosceles* e *Phoneutria* e não apenas as documentadas como mais comuns no Brasil, o aplicativo, apesar de errar a identificação taxonômica, indica uma espécie congênérica (ou apenas o gênero) que também é peçonhento. Nesse sentido, para um usuário em dúvida sobre a aranha encontrada ser ou não peçonhenta, a informação mais importante é o gênero, pouco importando o epíteto específico. Sendo assim, 77,42% das fotos de aranhas peçonhentas tiveram a identificação correta a nível de gênero no *Google Lens*.

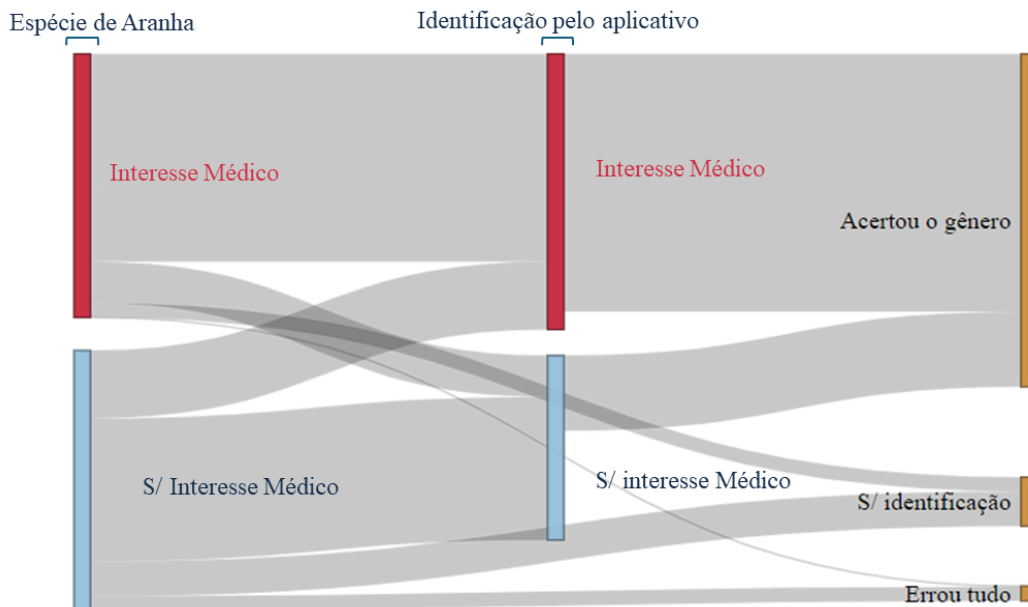
A Figura 1 revela ainda que há uma evidente discrepância da precisão dos resultados do *Google Lens* entre espécies, evidenciando que esse aplicativo varia muito em seu potencial de gerar identificações taxonômicas confiáveis dependendo da espécie. Isso pode ser reflexo de similaridades das espécies alvo do estudo com outras espécies morfológicamente similares, as quais podem ser confundidas entre si quando apenas características da foto são consideradas. Outra fonte de erro importante deve ser a qualidade da foto. Por exemplo, a posição que as aranhas se encontram na fotografia, deve influenciar a habilidade do *Google Lens* em identificá-la, o que deve depender não somente do usuário que fotografou o animal, mas também de atributos de comportamento e história de vida da espécie. Investigar como a biologia da espécie e a qualidade da foto influenciam nas chances de identificação taxonômica precisa por ferramentas de reconhecimento de imagem é um dos próximos passos do presente projeto.



**Figura 1:** Eficiência do *Google Lens* na identificação de espécies de aranhas encontradas dentro ou no entorno de construções habitáveis.

Para entender com mais detalhes os tipos de erros encontrados, classificamos os registros equivocados em (i) acerto de gênero, (ii) sem identificação ou (iii) errou tudo (i.e., gênero e epíteto errados ou nem mesmo uma aranha identificada). É interessante notar que uma porcentagem não trivial das espécies de interesse médico foi confundida com espécies sem interesse médico (15,9%) (Figura 2), o que pode ser perigoso para alguém tentando uma rápida identificação da espécie encontrada em casa usando o *Google Lens*. Por outro lado, há também uma confusão no grupo de espécies sem interesse médico, que por vezes foram confundidas por espécies de interesse médico (26,1%), o que pode gerar não apenas desinformação aos usuários, mas também colocar em perigo espécies de aranhas não peçonhentas que foram identificadas erroneamente. De modo geral, o *Google Lens* teve pior desempenho em identificar espécies sem interesse médico do que as com interesse médico ( $\chi^2 = 211.74$ , g.l. = 1,  $p < 0.01$ ), o que pode ser resultado de maiores buscas e pesquisas com espécies peçonhentas, que acabam treinando o algoritmo da plataforma. O mesmo

padrão é observado quando consideramos apenas identificações corretas no nível taxonômico de gênero, ou seja, houve uma maior proporção de acertos para espécies de interesse médico.



**Figura 2:** Tipos de erros encontrados na identificação do Google Lens. Google Lens por aranhas com interesse médico.

### Conclusões

Este parece ser o primeiro trabalho testando a eficiência de ferramentas de reconhecimento de imagens baseada em IA para a identificação de espécies de interesse médico. Poucos estudos prévios testaram a eficiência do Google Lens e outros aplicativos similares para outros organismos, principalmente plantas, mas fora de um contexto de aplicação para uso por cidadãos em um contexto de saúde pública (ZHADAN & KOLKIN, 2022; BANIA, 2022; PHILLIPS, 2022) De modo geral, nossos resultados evidenciam que o Google Lens possui limitações consideráveis na identificação precisa de aranhas, especialmente em nível de espécie, com uma taxa de acerto de apenas 21,35%.

No entanto, para aranhas de interesse médico, a identificação correta em nível de gênero foi mais eficaz, com uma taxa de acerto de 77,42%, o que sugere que a ferramenta pode ser útil em situações em que a identificação rápida do gênero é suficiente para medidas preventivas. Apesar disso, a variabilidade dos resultados, especialmente a confusão entre espécies peçonhentas e não peçonhentas, indica que a ferramenta não deve ser utilizada como única fonte de identificação em situações críticas de saúde pública. Tal variabilidade nos resultados sugere que o treinamento do algoritmo pode ser influenciado pela quantidade de dados disponíveis sobre espécies de maior interesse médico, destacando a necessidade de aprimoramento da ferramenta para fornecer identificações mais confiáveis.

## Referências

- AUGUST, Tom A.; PESCOTT, Oliver L.; JOLY, Alexis; BONNET, Pierre. AI Naturalists Might Hold the Key to Unlocking Biodiversity Data in Social Media Imagery. *Patterns*, v. 1, n. 7, 2020. DOI: <https://doi.org/10.1016/j.patter.2020.100116>. Disponível em: <https://www.sciencedirect.com/science/article/pii/S2666389920301574>. ISSN 2666-3899. Acesso em: 2 ago. 2024.
- BANIA, Allif et al. Identification of Ornamental Plants Via Google Lens Based on Intersemiotic. 2022. Disponível em: [https://www.researchgate.net/profile/Allif-Bania/publication/362380679\\_Identification\\_of\\_Ornamental\\_Plants\\_Via\\_Google\\_Lens\\_Based\\_on\\_Intersemiotic/links/62e8a0a43c0ea8788775b959/Identification-of-Ornamental-Plants-Via-Google-Lens-Based-on-Intersemiotic.pdf](https://www.researchgate.net/profile/Allif-Bania/publication/362380679_Identification_of_Ornamental_Plants_Via_Google_Lens_Based_on_Intersemiotic/links/62e8a0a43c0ea8788775b959/Identification-of-Ornamental-Plants-Via-Google-Lens-Based-on-Intersemiotic.pdf). Acesso em: 6 ago. 2024.
- BARTA, Z. Deep learning in terrestrial conservation biology. *Biologia Futura*, v. 74, p. 359–367, 2023. DOI: <https://doi.org/10.1007/s42977-023-00200-4>. Acesso em? 2 ago. 2024.
- CARO, T. Antipredator defenses in birds and mammals. Chicago: University of Chicago Press, 2005.
- JARIĆ, I.; ROLL, U.; ARLINGHAUS, R.; BELMAKER, J.; CHEN, Y.; CHINA, V.; et al. Expanding conservation culturomics and iEcology from terrestrial to aquatic realms. *PLoS Biology*, v. 18, n. 10, p. e3000935, 2020. DOI: <https://doi.org/10.1371/journal.pbio.3000935>.
- INSTITUTO BUTANTAN. Aranhas em casa, na escola, no transporte: devo ter medo delas? Disponível em: <https://butantan.gov.br/bubutantan/aranhas-em-casa-na-escola-no-transporte-devo-ter-medo-delas>. Acesso em: 6 ago. 2024.
- INSTITUTO BUTANTAN. Maioria das aranhas que vive nas cidades não é venenosa; conheça as mais perigosas. Disponível em: <https://butantan.gov.br/noticias/maioria-das-aranhas-que-vive-nas-cidades-nao-e-venenosa--conheca-as-mais-perigosas>. Acesso em: 6 ago. 2024.
- MINISTÉRIO DA SAÚDE. Brasil. *Boletim Epidemiológico*. v. 53, n. 31, 2022. Disponível em: <https://www.gov.br/saude/pt-br/centrais-de-conteudo/publicacoes/boletins/epidemiologicos/edicoes/2022/boletim-epidemiologico-vol-53-no31>. Acesso em: 2 ago. 2024.
- PHILLIPS, Stephen F. Assessment of the Diagnostic Skills and Needs of Nebraska Extension Entomology Diagnosticians. 2022. Disponível em: <https://digitalcommons.unl.edu/cgi/viewcontent.cgi?article=1090&context=entodistmasters>. Acesso em: 6 ago. 2024.
- SÃO PAULO (Município). Secretaria Municipal da Saúde. Animais Sinantrópicos. Disponível em: [https://www.prefeitura.sp.gov.br/cidade/secretarias/saude/vigilancia\\_em\\_saude/controlado\\_de\\_zoonoses/animais\\_sinantronicos/index.php?p=4470](https://www.prefeitura.sp.gov.br/cidade/secretarias/saude/vigilancia_em_saude/controlado_de_zoonoses/animais_sinantronicos/index.php?p=4470). Acesso em: 6 ago. 2024.
- SPERANZA, E. I. et al. Aranhas sinantrópicas em áreas urbanas: implicações para a saúde humana e o controle de pragas. *Biota Neotropica*, v. 20, n. 4, 2020. Disponível em: <https://www.scielo.br/j/bn/a/Tkq8GV5XZ8jxDqKbxm6F5pr/>. Acesso em: 6 ago. 2024.
- VAN HORN, G.; MAC AODHA, O.; SONG, Y.; CUI, Y.; SUN, C.; SHEPARD, A.; ADAM, H.; PERONA, P.; BELONGIE, S. The iNaturalist species classification and detection dataset. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, 2018, p. 8769–8778. Disponível em: [https://openaccess.thecvf.com/content\\_cvpr\\_2018/html/Van\\_Horn\\_The\\_iNaturalist\\_Species\\_CVPR\\_2018\\_paper.html](https://openaccess.thecvf.com/content_cvpr_2018/html/Van_Horn_The_iNaturalist_Species_CVPR_2018_paper.html). Acesso em: 5 ago. 2024.
- ZHADAN, Sergey; KOLKIN, Stanislav. Comparing Google Lens Recognition Accuracy with Other Plant Recognition Apps. 2022. Disponível em: [https://www.researchgate.net/profile/Sergey-Zhadan/publication/360662195\\_Comparing\\_Google\\_Lens\\_Recognition\\_Accuracy\\_with\\_Other\\_Plant\\_Recognition\\_Apps/links/628798af6e41e5002d3501cb/Comparing-Google-Lens-Recognition-Accuracy-with-Other-Plant-Recognition-Apps.pdf](https://www.researchgate.net/profile/Sergey-Zhadan/publication/360662195_Comparing_Google_Lens_Recognition_Accuracy_with_Other_Plant_Recognition_Apps/links/628798af6e41e5002d3501cb/Comparing-Google-Lens-Recognition-Accuracy-with-Other-Plant-Recognition-Apps.pdf). Acesso em: 6 ago. 2024.