

# **Spider Species Identification: Bridging Traditions with Molecular and Deep Learning Approaches**

## **Abstract:**

This article provides an in-depth review of various methods employed in the identification and sequencing of spiders, highlighting the advancements and challenges in the field. With the increasing importance of spiders in ecological studies, medical research, and biodiversity conservation, accurate identification and genetic analysis have become crucial. This review discusses traditional and modern techniques, shedding light on their applications, limitations, and future prospects.

The exploration begins with an analysis of taxonomists' etymological choices, examining patterns in naming conventions across continents and centuries. Traditional morphological identification, anchored in backbone taxonomy, dichotomous keys, and statistical analyses, highlights the advantages and challenges of relying on observable features. The study transitions to molecular techniques, elucidating the applications and challenges of DNA barcoding, Next-Generation Sequencing (NGS), and metabarcoding in spider identification. The integration of deep learning models, exemplified by the YOLOv7-based Spider Identification APP, represents a landmark in computer vision for efficient and user-friendly spider species recognition. The study's multifaceted approach provides a nuanced understanding of spider taxonomy, bridging historical practices with state-of-the-art technologies, and lays the groundwork for future advancements in the field.

**Keywords:** Spider identification, Morphological taxonomy, DNA barcoding, Next-Generation Sequencing, Metabarcoding, Deep learning models, Biodiversity analysis.

## **Introduction:**

Spiders play a vital role in ecosystems, and their identification and sequencing have gained significance in recent years. This section introduces the importance of spider identification and the role of genetic sequencing in understanding their biology, behavior, and evolutionary relationships. Understanding the complexities of the natural world has been a longstanding

human endeavor, characterized by the systematic classification of elements[43]. This tradition, evident in Mendeleev's periodic table and stellar categorization, extends to biological taxonomy. Linnaeus's binomial nomenclature, proposing species identification through genus and species epithets, remains a lasting system. This study explores taxonomists' etymological choices in spider taxonomy, scrutinizing patterns across time and continents, with a focus on potential cultural influences. The investigation traverses' traditional methods, emphasizing the significance of morphological identification and molecular techniques, offering a holistic perspective on spider taxonomy.

## **Different Experimental Approach for Species Identification:**

### **1.Traditional Methods:**

Throughout history, humans have organized natural elements like stars and species into systematic frameworks. Linnaeus introduced binomial nomenclature, defining species through genus and species names. Taxonomists creatively name species, drawing inspiration from traits, habitats, individuals, or wordplay (Scerri, 2019; Langer & Kudritzki, 2014; Linnaeus, 1751, 1758; Vendetti & Garland, 2019; Lehman, 1967; Hey, 2001; Slater, 2016). Our study examines these naming patterns in spider taxonomy to reveal potential cultural influences.

**2.Morphological Identification:** Traditional taxonomic methods rely on morphological characteristics, including body structure, coloration, and genital morphology. In this study, spider samples were collected from Ratapani Wildlife Sanctuary situated in the Indian state of Madhya Pradesh. The sanctuary spans across portions of the Raisen and Sehore districts, with approximate geographical coordinates of Latitude: 23.4346° N and Longitude: 77.4980° E.



Image1: Ratapani Wildlife Sanctuary Latitude: 23.4346° N Longitude: 77.4980° E.(Image taken by Google Map)

### Sample Organism:

The specimen I collected belongs to the Nephila family. The Nephila family of spiders is commonly found in the area surrounding Ratapani Wildlife Sanctuary. The Nephila spider, also called the Giant Wood Spider or Golden Orb Weaver, belongs to a family known for its large, colorful spiders found in tropical and subtropical areas of Asia and Australia. These spiders are famous for their impressive size, intricate webs, and striking appearance. Despite their imposing size, they pose little threat to humans and are important for managing insect numbers in their environments.

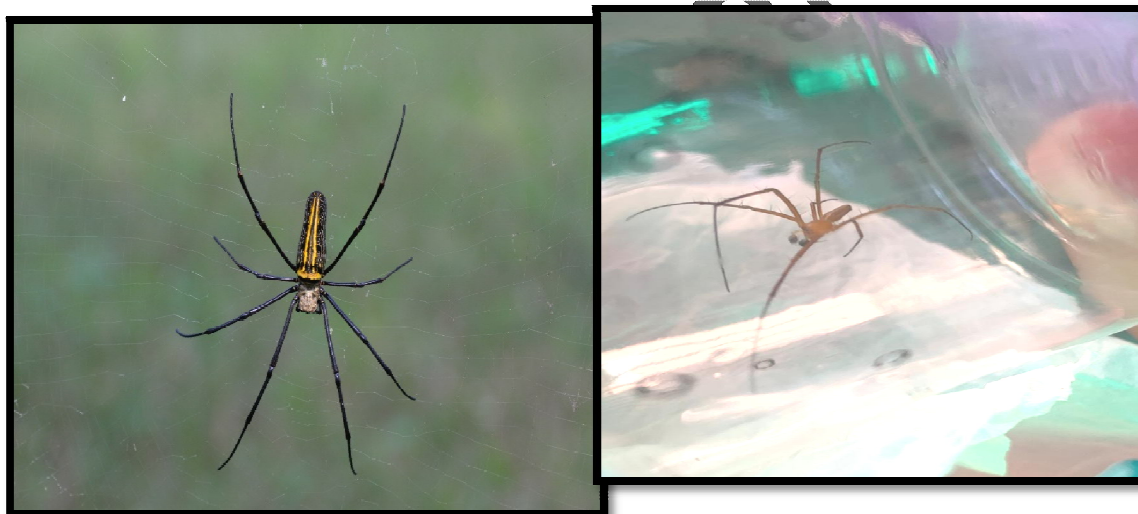


Image 2:Image of *Nephila pilipes* from Ratapani Wildlife Sanctuary, Bhopal (M.P.), highlighting the morphological features that illustrate sexual dimorphism between the female and male.

*N. pilipes* exhibit female gigantism and male dwarfism, with significant size differences between males and females, particularly notable among terrestrial animals. This sexual dimorphism is likely a result of evolutionary selection favoring females with higher fecundity (Kuntner, M., et al2014). Female *N. pilipes* invest heavily in parental care, including egg production and web

construction, demonstrating substantial commitment to their progeny.

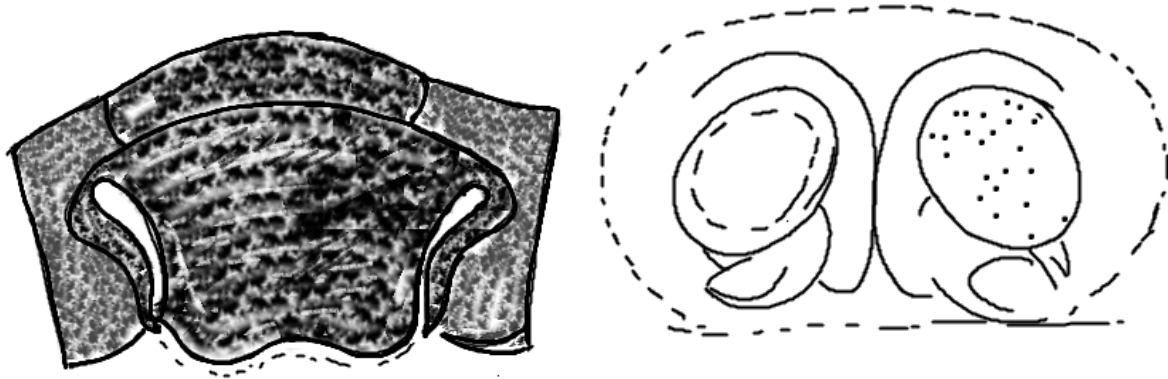


Image 3: Create a ray diagram of the female genital organ of *Nephila pilipes* as described by Kuntner et al. (2009).

Female: Female *N. pilipes* are usually 30–50 mm in size, with a cephalothorax measuring about 15 mm long and 10 mm wide, and an abdomen around 30 mm long and 15 mm wide. They are mostly dark yellow-brown with yellow stripes, black or brown tergum covered in dense hairs, and long black and yellow legs lacking apparent hairs. The color contrast aids in foraging success with visually-oriented prey.

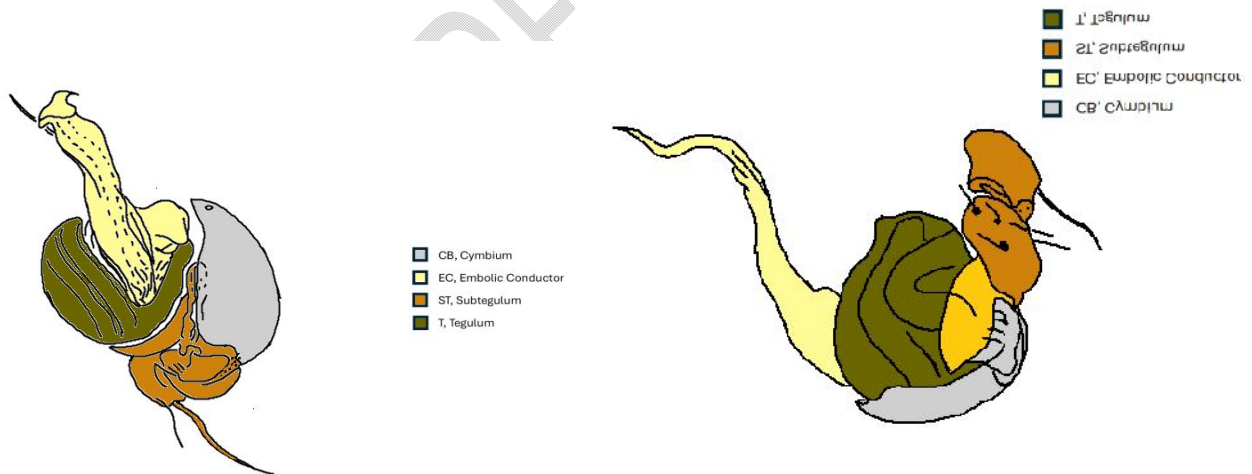


Image 4: Recreate the diagram of the male genital organ of *Nephila pilipes*, using different colors to label the following parts: CB (cymbium), E (embolus), EC (embolic conductor), Ecp (proximal part of the embolic conductor), Ecd (distal part of the embolic conductor), ST (subtegulum), and T (tegulum) as shown by Kuntner et al. (2009).

Male: Male *N. pilipes* typically measure 5–6.5 mm, with a cephalothorax about 2.5 mm long and 2 mm wide, and an abdomen approximately 4 mm long and 1.5 mm wide. They have light brown legs with some hairs, larger front eyes than rear eyes, and a yellow carapace with very few hairs.

**2.1 Backbone Taxonomy:** The backbone taxonomy was derived from the World Spider Catalog (2020) with 48,464 taxonomic entities, spanning back to Carl Alexander Clerck's 1757 work. To align with zoological nomenclature, Clerck's publication date was adjusted to January 1, 1758 (Clerck, 1757; Linnaeus, 1758).

**2.2 Etymology Classification and Validation:** Species etymologies were categorized into six groups, covering morphology, ecology, geography, people, modern/past culture, and others. Etymologies were assigned based on original descriptions, supplemented by inference using language expertise, etymological references, and collaboration (Bonnet, 1961; Nilsson, 2010). Cross-validation of 400 randomly selected inferred etymologies demonstrated internal consistency, with 57.75% matching original descriptions.

**2.3. Statistical Analyses:** All analyses were conducted in R v.4.0.3 (R Core Team, 2021) using 'ggplot2' v.3.3.4 (Wickham, 2016). Hypotheses were tested through regression models with a significance threshold set at 0.001. Temporal trends in etymology choice were analyzed using a generalized additive model. A second model assessed temporal trends across continents, considering a binomial distribution. Generalized linear models examined the likelihood of taxonomists selecting a specific etymology category across continents (Zuur & Ieno, 2016; Wood, 2004; Zuur et al., 2009).

**2.4 Species Geographic Distribution:** Species distributions were classified at the continental level using World Spider Catalog data and converted into ISO country codes (Cardoso & Pekar, 2022; Arel-Bundock et al., 2018).

**2.5 Dichotomous Keys:** Dichotomous keys, such as Levi's orb-weaving spider key (1957), guide users through choices leading to species identification based on observable characteristics. Their value lies in structured decision-making, aiding accurate identifications (Levi, 1957).

### **3. Molecular Techniques:**

**3.1. DNA Barcoding:** DNA barcoding uses specific DNA sequences for species identification. Spider specimens were collected, and their genomic DNA was extracted. The COI gene was

amplified via PCR and sequenced. Sequences were compared to BOLD and GenBank databases using BLAST for species identification, following Hebert et al. (2003) protocols.

**3.2. Next-Generation Sequencing (NGS):** NGS technologies have revolutionized genomics, enabling large-scale spider sequencing studies. Specimens were collected, and DNA extracted using a commercial kit (Manufacturer's Manual). Libraries were prepared with the Nextera XT Kit and sequenced on an Illumina HiSeq platform. Data were processed with Trimmomatic, assembled with SPAdes, and annotated using BLAST (Bolger et al., 2014; Bankevich et al., 2012; Altschul et al., 1990).

**3.3. Metabarcoding:** Metabarcoding enables simultaneous identification of multiple species from environmental DNA samples. In spider community analysis, DNA is non-invasively collected, extracted, amplified using universal primers, and sequenced. Bioinformatics pipelines process the data for community analysis using R or QIIME software (Elbrecht & Leese, 2017; Deagle et al., 2014; Edgar, 2013; Caporaso et al., 2010).

**3.4 Deep Learning Models:** Our YOLOv7-based Spider Identification APP (SpiderID\_APP) efficiently identifies spider genera in Taiwan, leveraging deep learning. Trained on 24,000 iNaturalist images, it provides 120 genus classifications. Complementary methods like DNA barcoding are needed for species-level identification. The APP features a user-friendly GUI, integrating YOLO detection scripts (Elbrecht & Leese, 2017; Deagle et al., 2014).

## **Result and Discussion**

The comprehensive study delves into spider identification and sequencing methods, encompassing traditional approaches and cutting-edge technologies. The analysis of taxonomists' etymological choices revealed temporal trends and internal consistency, reflecting a balance between historical adherence and evolving conventions. Backbone taxonomy exploration and dichotomous keys highlighted the significance of morphological identification, recognizing its value while emphasizing the need for continuous updates. In this study, I observed *Nephila* spiders found in Ratapani Wildlife Sanctuary and conducted morphological identification based on previous scientific research. Molecular techniques, including DNA barcoding, Next-Generation Sequencing (NGS), and metabarcoding, showcased their efficacy in spider genomics

and biodiversity studies. The YOLOv7-based Spider Identification APP represented a milestone in computer vision, seamlessly integrating with traditional methods. Challenges such as sample contamination and incomplete databases were identified, emphasizing the ongoing need for methodological refinement. In summary, the results provide a multifaceted understanding of spider taxonomy, bridging historical practices with state-of-the-art technologies and paving the way for future advancements in the field.

### **Conclusion:**

This review explores various methods used to identify and sequence spiders. As technology advances, combining traditional and molecular approaches improves accuracy and efficiency in spider taxonomy and genomics. Integrating these methods is crucial for understanding spider diversity, ecology, and evolution. Despite challenges like sample contamination and limited databases, the field progresses. Future advancements, including artificial intelligence and portable sequencing devices, promise greater strides. Embracing these technologies will deepen our understanding of spider diversity and evolution. The interdisciplinary nature of spider identification ensures a strong foundation for uncovering the complexities of these fascinating creatures.

Disclaimer (Artificial intelligence)

Option 1:

Author(s) hereby declare that NO generative AI technologies such as Large Language Models (ChatGPT, COPILOT, etc) and text-to-image generators have been used during writing or editing of manuscripts.

Option 2:

Author(s) hereby declare that generative AI technologies such as Large Language Models, etc have been used during writing or editing of manuscripts. This explanation will include list the

name, version, model, and source of the generative AI technology and as well as the all input prompts provided to a generative AI technology

Details of the AI usage are given below:

- 1.
- 2.
- 3.

### References:

1. Altschul, S. F., Gish, W., Miller, W., Myers, E. W., & Lipman, D. J. (1990). Basic local alignment search tool. *Journal of Molecular Biology*, 215(3), 403-410.
2. Arel-Bundock, V., Enevoldsen, N., & Yetman, C. J. (2018). Country code: Convert country names and country codes. *Journal of Open Source Software*, 3(28), 1025.
3. Bennett, A. F., & Saunders, D. (2011). Habitat fragmentation and landscape change. In N. S. Sodhi & P. R. Ehrlich (Eds.), *Conservation Biology for All* (Chapter 5). Oxford University Press.
4. Birkhofer, K., Gavish-Regev, E., Endlweber, K., Lubin, Y. D., von Berg, K., Wise, D. H., & Scheu, S. (2008). Cursorial spiders retard initial aphid population growth at low densities in winter wheat. *Bulletin of Entomological Research*, 98(3), 249-255.
5. Bonnet, P. (1961). *Bibliographia Araneorum. Analyse méthodique de toute la littérature aranéologique jusqu'en 1939. Tome III*. Toulouse: Imprimerie Douladoure.
6. Cain, A. J. (2010). The Methodus of Linnaeus. *Archives of Natural History*, 19(2), 231-250.
7. Cardoso, P., & Pekár, S. (2022). wscmap: A world spider catalog map. *Biodiversity Data Journal*, 10, e79648.
8. Cardoso, P., Pekár, S., Jocqué, R., & Coddington, J. A. (2011). Global patterns of guild composition and functional diversity of spiders. *PLoS ONE*, 6(6), e21710.
9. Chaplin-Kramer, R., O'Rourke, M. E., Blitzer, E. J., & Kremen, C. (2011). A meta-analysis of crop pest and natural enemy response to landscape complexity. *Ecology Letters*, 14(9), 922-932.
10. Chiao, C. C., Wu, W. Y., Chen, S. H., & Yang, E. C. (2009, June 26). Visualization of the spatial and spectral signals of orb-weaving spiders, *Nephila pilipes*, through the eyes of a honeybee. *Journal of Experimental Biology*, 212(14), 2269–2278. <https://doi.org/10.1242/jeb.030734>
11. Coddington, J. A., Hormiga, G., & Scharff, N. (1997). Giant female or dwarf male spiders? *Nature*, 385(6618), 687–688.

12. Danielson-François, A., Hou, C., Cole, N., & Tso, I. M. (2012). Scramble competition for moulting females as a driving force for extreme male dwarfism in spiders. *Animal Behaviour*, *84*(4), 937–945.
13. Deagle, B. E., Thomas, A. C., Shaffer, S. A., Trites, A. W., & Jarman, S. N. (2014). DNA metabarcoding and the cytochrome c oxidase subunit I marker: Not a perfect match. *Biological Letters*, *10*(9), 20140562.
14. Dormann, C. F., McPherson, J. M., Araújo, M. B., Bivand, R., Bolliger, J., Carl, G., ... & Kühn, I. (2007). Effects of landscape structure and land-use intensity on similarity of plant and animal communities. *Global Ecology and Biogeography*, *16*(6), 774–787.
15. Elbrecht, V., Taberlet, P., Dejean, T., Valentini, A., Usseglio-Polatera, P., Beisel, J. N., ... & Leese, F. (2016). Testing the potential of a ribosomal 16S marker for DNA metabarcoding of insects. *PeerJ*, *4*, e1966.
16. Elbrecht, V., & Leese, F. (2017). PrimerMiner: An R package for development and in silico validation of DNA metabarcoding primers. *Methods in Ecology and Evolution*, *8*(5), 622–626.
17. Harvey, M. S., Austin, A. D., & Adams, M. (2007). The systematics and biology of the spider genus *Nephila* (Araneae: Nephilidae) in the Australasian region. *Invertebrate Systematics*, *21*(5), 407–451.
18. Hebert, P. D., Ratnasingham, S., & deWaard, J. R. (2003). Biological identifications through DNA barcodes. *Proceedings of the Royal Society of London. Series B: Biological Sciences*, *270*(1512), 313–321.
19. Jäger, P., Arnedo, M. A., Azevedo, G. H. F., Baehr, B., Bonaldo, A. B., Haddad, C. R., ... & Ramírez, M. J. (2021). Twenty years, eight legs, one concept: Describing spider biodiversity in Zootaxa (Arachnida: Araneae). *Zootaxa*, *4979*(1), 131–146.
20. Jozwiak, P., Rewicz, T., & Pabis, K. (2015). Taxonomic etymology – in search of inspiration. *ZooKeys*, *513*, 143–160.
21. Kuntner, M., Coddington, J. A., & Schneider, J. (2009). Intersexual arms race? Genital coevolution in nephilid spiders (Araneae, Nephilidae). *Evolution*, *63*(6), 1451–1463.
22. Kuntner, M., Kralj-Fišer, S., Schneider, J. M., & Li, D. (2009). Mate plugging via genital mutilation in nephilid spiders: An evolutionary hypothesis. *Journal of Zoology*, *277*(4), 257–266.
23. Kuntner, M., Zhang, S., Gregorič, M., & Li, D. (2012, November). *Nephila* female gigantism attained through post-maturity molting. *Journal of Arachnology*, *40*(3), 345–347.
24. Kuntner, M., & Elgar, M. A. (2014). Evolution and maintenance of sexual size dimorphism: Aligning phylogenetic and experimental evidence. *Frontiers in Ecology and Evolution*, *2*, Article 26.
25. Lee, V. M. J., Kuntner, M., & Li, D. (2015, January 21). Ballooning behavior in the golden orbweb spider *Nephila pilipes* (Araneae: Nephilidae). *Frontiers in Ecology and Evolution*, *3*.
26. Legrand, R. S., & Morse, D. H. (2000). Factors driving extreme sexual size dimorphism of a sit-and-wait predator under low density. *Biological Journal of the Linnean Society*, *71*(4), 643–664.
27. Levi, H. W. (1957). The spider genera *Enoplognatha*, *Theridion*, and *Paidisca* in America north of Mexico. *American Museum Novitates*, *1832*, 1–39.

28. Luong, C. T., Tsai, C. S., Lin, S. Y., & Chang, C. M. (2023). SpiderID\_APP: A user-friendly APP for spider identification in Taiwan using YOLO-based deep learning models. *Insects*, 14(2), 160.
29. Mammola, S., Viel, N., Amiar, D., Mani, A., Hervé, C., Heard, S. B., Fontaneto, D., & Pétillon, J. (2023). Taxonomic practice, creativity and fashion: What's in a spider name? *Zoological Journal of the Linnean Society*, 198(2), 494–508.
30. Natural History Museum Bern. (2017, May 15). Taxon details Nephila pilipes (Fabricius, 1793). World Spider Catalog. Retrieved from [URL]
31. Nilsson, A. N. (2010). All diving beetle specific and subspecific names explained. *Skörvnöpparn, Umeå Supplement 1*, 1–42.
32. Platnick, N. I. (1989). Advances in spider taxonomy 1981-1987: A supplement to Brignoli's A catalog of the Araneae described between 1940 and 1981 (Edited by P. Merrett). Manchester University Press. Distributed exclusively in the United States and Canada by St. Martin's Press.
33. Ratnasingham, S., & Hebert, P. D. N. (2007). The Barcode of Life Data System. *Molecular Ecology Notes*, 7(3), 355-364.
34. Tso, I. M. (2004, July 1). Colourful orb-weaving spiders, *Nephila pilipes*, through a bee's eyes. *Journal of Experimental Biology*, 207(15), 2631–2637.
35. Tso, I. M. (2005, March 15). Giant wood spider *Nephila pilipes* alters silk protein in response to prey variation. *Journal of Experimental Biology*, 208(6), 1053–1061.
36. Tso, I. M., Chiang, S. Y., & Blackledge, T. A. (2007, March 27). Does the giant wood spider *Nephila pilipes* respond to prey variation by altering web or silk properties? *Ethology*, 113(4), 324–333. <https://doi.org/10.1111/j.1439-0310.2007.01318.x>
37. Vendetti, J., & Garland, R. (2019). Species name formation for zoologists: A pragmatic approach. *Journal of Natural History*, 53(47-48), 2999-3018.
38. Vendetti, J., Vendetti, J., & Garland, R. (2019). Species name formation for zoologists: A pragmatic approach. *Journal of Natural History*, 53(47-48), 2999-3018.
39. Wang, Z., Cang, Y., Kremer, F., Thomas, E. L., & Fytas, G. (2020, January 14). Determination of the complete elasticity of *Nephila pilipes* spider silk. *Biomacromolecules*, 21(3), 1179–1185. <https://doi.org/10.1021/acs.biomac.9b01607>
40. Wise, D. H. (1981). Inter- and intraspecific effects of density manipulations upon females of two orb-weaving spiders (Araneae: Araneidae). *Oecologia*, 48(2), 252–256.
41. Zhang, S., Kuntner, M., & Li, D. (2011). Mate binding: Male adaptation to sexual conflict in the golden orb-web spider (Nephilidae: *Nephila pilipes*). *Animal Behaviour*, 82(6), 1299–1304.
42. Zuur, A. F., & Ieno, E. N. (2016). A protocol for data exploration to avoid common statistical problems. *Methods in Ecology and Evolution*, 7(10), 1121–1131.
43. Bjerge K, Mann HM, Høye TT. Real-time insect tracking and monitoring with computer vision and deep learning. *Remote Sensing in Ecology and Conservation*. 2022 Jun;8(3):315-27.