

Dear Editors,

I am pleased to submit a proposal for a Methods article entitled “SegColR: Deep Learning for Automated Segmentation and Color Extraction.” This proposed article introduces SegColR, a new software package designed to address the challenges of extracting phenotypic information from the vast amount of biodiversity data generated by citizen science platforms such as iNaturalist.

The rapid accumulation of observational data on these platforms presents significant hurdles for extracting accurate and meaningful phenotypic details, such as color information. Traditional manual methods of extraction are often subjective and labor-intensive. To overcome these limitations, SegColR employs advanced deep learning models, specifically GroundingDINO and SegmentAnything, to automate the processes of image segmentation and color extraction. This approach not only improves efficiency but also enhances the precision of phenotypic data extraction from non-standardized images.

This work represents a significant contribution to the field by integrating cutting-edge deep learning tools into the R programming environment. While deep learning advancements are predominantly driven by developments in Python, R remains a cornerstone of academic research in ecology and evolutionary biology. By making these sophisticated image processing techniques accessible through a user-friendly R-based interface, SegColR democratizes access to advanced data analysis tools for a broader audience of biologists and ecologists. The automation and flexibility provided by SegColR facilitate large-scale biodiversity studies and ecological research, bridging a critical gap in the analysis of citizen science data.

SegColR offers a streamlined interface for object detection and segmentation, leveraging lightweight, pretrained models to enable use on personal computers. The package’s prompt-based object detection capabilities allow for the extraction of specific image components, as demonstrated by excluding particular features, such as fish fins, in the color extraction process. Attached below, you will find figures generated directly by the SegColR package, showcasing its core functionalities.

Thank you for considering this proposal. I am eager to submit the full article for review and to contribute to the ongoing advancements in biodiversity research and data analysis.

Best regards,

James Boyko