Review of: "Maximizing citizen scientists' contribution to automated species recognition"

shaun winterton

Potential competing interests: The author(s) declared that no potential competing interests exist.

An excellent study. The use of AI in image identification is becoming ever more important and widely used on very popular platforms. Still, garbage in means garbage out, so methods that can assist in understanding the limitations of the baseline image data and identifications that the AI uses as it's reference database is important in understanding were improvements (training) must be made. I would be also interested in spatial patterns in over-represented versus under-represented taxa. It is clear that observations that populate such image databases are biased towards serendipitous encounters near urban areas (especially obvious when mapped), charismatic and large taxa, and diurnal taxa (for animals at least). It also makes sense that flying groups with a large amateur interest bases are over-represented (i.e. butterflies, birds, dragonflies). The conclusions that prioritised data collection of under-represented groups should be a focus for these image databases in the future.