Bug Soup: Using Computer Vision to Automate Analysis of Insects Samples Preserved in Liquid

Aubrey Moore

June 3, 2021

Contents

1 Introduction

Insect sampling methods used in biodiversity studies (pan traps, Burlese funnels, Malaise traps, etc.) often result in large numbers of insects preserved in alcohol or other liquid. I refer to this as *bug soup*. Often, individual specimens are sorted by hand for identification and further study. This process can be very tedious. Modern digital macrophotography coupled with computer vision can relieve the tedium (Hoye et al. 2021).

The objective of this study is to develop a highly automated method that will count individual insects in a bug soup and identify each to species or morphospecies.

This is a *living document* documenting my progress.

2 Imaging

- Images are acquired using an Olympus TG-5 camera using the microscope mode with in-camera focus stacking. Illumination comes from the built-in flash unit fitted with and optional diffuser.

- Each *bug soup* will be contained in an unvented tissue culture flask. These flasks are designed for microscopic examination of cell cultures but suit this purpose well.

3 Object detection

- *Object detection* is the process of identifying and extracting areas of interest in an image. In this case areas of interest are images of individual insects. **ImageJ** has an excellent **particle analysis plugin** which can be used for this purpose. We can start with ImageJ but I think I will roll my own in Python.

4 Analysis

- A deep learning model, similar to the one used by iNaturalist for automated identification of organisms (Ken -ichi Ueda 2020) will be developed to identify insects in the bug soup.

- It will be possible to build sufficiently large training sets by taking multiple images of each flask with agitation between images. This will result in multiple images of individuals in many orientations.

- The deep learning model can be trained using two different methods: **supervised learning** or **unsupervised learning**. In **supervised learning**, a subset of insect images is selected as a training set and each one is labeled (to species or pseudospecies) by a human expert. In **unsupervised learning**, the training set is automatically grouped into pseudospecies using a clustering algorithm.

5 References

Hoye, Toke T., Johanna Ärje, Kim Bjerge, Oskar L. P. Hansen, Alexandros Iosifidis, Florian Leese, Hjalte M. R. Mann, Kristian Meissner, Claus Melvad, and Jenni Raitoharju (Jan. 12, 2021). "Deep Learning and Computer Vision Will Transform Entomology". In: *Proceedings of the National Academy of Sciences* 118.2. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.2002545117. pmid: 33431561. URL: https://www.pnas.org/content/118/2/e2002545117 (visited on 05/16/2021).

Ken -ichi Ueda (Oct. 29, 2020). *An Overview ofComputer Vision on iNaturalist.* YouTube; TDWG 2020. URL: https://www.youtube.com/watch?v=xfbabznYFV0 (visited on 05/07/2021).