

# The importance and utility of machine learning for individually identifying elephants

Sydney F. Hope<sup>1\*</sup>, Sarah L. Jacobson<sup>1,2</sup>, Jan Blunk<sup>3</sup>, Paul Bodesheim<sup>3</sup>, Matthias Koerschens<sup>3</sup>, Joshua M. Plotnik<sup>1,2</sup>

<sup>1</sup>Department of Psychology, Hunter College, City University of New York, New York, USA

<sup>2</sup>Department of Psychology, The Graduate Center, City University of New York, New York, USA

<sup>3</sup>Department of Mathematics and Computer Science, University of Jena, Jena, Germany

\*sh7548@hunter.cuny.edu

**Abstract:** Analyzing animal traits at the individual level is at the core of ecology but challenging when studying wildlife populations. Our lab studies wild elephant behavior and cognition using specially-designed apparatuses, from which we can measure different aspects of behavior by observing how individuals interact with them. These apparatuses remain in the wild for long periods of time, allowing individuals to interact with them multiple times, and all behaviors are recorded using camera traps. Individual identification is necessary to study the change in or consistency in behavior within the same individual over time, and to match behavior with other factors that were observed in the same individual at a different time point or location. However, our current methods for individual identification are limited to characterizing traits (e.g., markings, ear folds, depigmentation) in photos, which takes a great deal of time and is prone to human error and bias. To solve this problem, we have partnered with the University of Jena to develop a machine-learning-based program that automatically identifies elephant individuals, which has shown good performance on an excerpt of our collected data. We are currently in the process of configuring and testing the program for its utility in our identification process. Furthermore, we are improving the identification abilities of the model by annotating new camera trap videos of individuals that we have recently identified to the training dataset (N > 250 elephants in our study population). Although an initial labeling effort for creating a training dataset is required, we think that the program is promising. It has the potential to save time and reduce human error in elephant identification, which is critical for elephant behavior, cognition, and conservation research.