Hierarchical machine learning classifiers better predict source identity from marmoset vocalizations

Nikhil Phaniraj[1,2], Kaja Wierucka[1], Yvonne Zürcher[1], Judith M Burkart[1]

[1]University of Zürich, [2]Indian Institute of Science Education and Research (IISER) Pune

nikhil.phaniraj@uzh.ch

Abstract (max 300 words)

Animals living in dense habitats or with poor vision that have little visual contact between group members often have to signal identity in their vocalizations along with other social, emotional, and contextual information (Boughman & Wilkinson, 1998; Fukushima et al., 2015; Prat et al., 2016; Soltis et al., 2005; Tooze et al., 1990). In such cases, it is possible that animals use broader-category cues that can convey information such as age or sex of the source for effective individual recognition. Determining source identity from vocalizations is important not only for receiver individuals but also for researchers studying these animals. As manual vocal data acquisition and real-time labeling by researchers is a cumbersome task, automizing this process is beneficial (Blumstein et al., 2011; Rickwood & Taylor, 2008). This has led to the development of a plethora of machine learning techniques for source identification, each competing for higher accuracy (Cheng et al., 2010, 2012; Stowell et al., 2019; Vieira et al., 2015). However, whether machine learning algorithms can use broader-category cues present in vocalizations for efficient source identification remains poorly explored. In this study, we used common marmosets, an arboreal primate species relying mainly on vocal cues to communicate, to test this hypothesis (Eliades & Miller, 2017). Because marmosets tend to live in smaller groups with an average of around eight individuals and have a social structure that includes helpers and breeders (Erb & Porter, 2017), it is plausible that most of the individual variability of calls within groups is explained by variation in sex and social structure. We found that a hierarchical classification approach in which the machine first learns to determine the social status, then sex, and finally the source identity outperformed the non-hierarchical classifier. This provided evidence that marmoset calls contain information about sex, social status, and individual identity of the caller. Finally, we assessed the impact of sample size on classifier accuracy and provided sample size guidelines for future studies. Hierarchical classifiers appear to be a promising tool for automatic source identification from animal vocalizations.

[Abstract Word count: 289]

References

Blumstein, D. T., Mennill, D. J., Clemins, P., Girod, L., Yao, K., Patricelli, G., Deppe, J. L., Krakauer, A. H., Clark, C., Cortopassi, K. A., Hanser, S. F., McCowan, B., Ali, A. M., & Kirschel, A. N. G. (2011). Acoustic monitoring in terrestrial environments using microphone arrays: Applications, technological considerations and prospectus. *Journal of Applied Ecology*, *48*(3), 758–767. https://doi.org/10.1111/j.1365-2664.2011.01993.x

Boughman, J. W., & Wilkinson, G. S. (1998). Greater spear-nosed bats discriminate group mates by vocalizations. *Animal Behaviour*, *55*(6), 1717–1732.

Cheng, J., Sun, Y., & Ji, L. (2010). A call-independent and automatic acoustic system for the individual recognition of animals: A novel model using four passerines. *Pattern Recognition*, *43*(11), 3846–3852. https://doi.org/10.1016/j.patcog.2010.04.026

Cheng, J., Xie, B., Lin, C., & Ji, L. (2012). A comparative study in birds: Call-type-independent species and individual recognition using four machine-learning methods and two acoustic features. *Bioacoustics*, *21*(2), 157–171.

Eliades, S. J., & Miller, C. T. (2017). Marmoset vocal communication: Behavior and neurobiology. *Developmental Neurobiology*, *77*(3), 286–299.

Erb, W. M., & Porter, L. M. (2017). Mother's little helpers: What we know (and don't know) about cooperative infant care in callitrichines. *Evolutionary Anthropology: Issues, News, and Reviews*, *26*(1), 25–37.

Fukushima, M., Doyle, A. M., Mullarkey, M. P., Mishkin, M., & Averbeck, B. B. (2015). Distributed acoustic cues for caller identity in macaque vocalization. *Royal Society Open Science*, *2*(12), 150432.

Prat, Y., Taub, M., & Yovel, Y. (2016). Everyday bat vocalizations contain information about emitter, addressee, context, and behavior. *Scientific Reports*, *6*(1), 1–10.

Rickwood, P., & Taylor, A. (2008). Methods for automatically analyzing humpback song units. *The Journal of the Acoustical Society of America*, *123*(3), 1763–1772. https://doi.org/10.1121/1.2836748

Soltis, J., Leong, K., & Savage, A. (2005). African elephant vocal communication II: Rumble variation reflects the individual identity and emotional state of callers. *Animal Behaviour*, *70*(3), 589–599.

Stowell, D., Petrusková, T., Šálek, M., & Linhart, P. (2019). Automatic acoustic identification of individuals in multiple species: Improving identification across recording conditions. *Journal of The Royal Society Interface*, *16*(153), 20180940. https://doi.org/10.1098/rsif.2018.0940

Tooze, Z. J., Harrington, F. H., & Fentress, J. C. (1990). Individually distinct vocalizations in timber wolves, Canis lupus. *Animal Behaviour*, *40*(4), 723–730.

Vieira, M., Fonseca, P. J., Amorim, M. C. P., & Teixeira, C. J. (2015). Call recognition and individual identification of fish vocalizations based on automatic speech recognition: An example with the Lusitanian toadfish. *The Journal of the Acoustical Society of America*, *138*(6), 3941–3950.