BeeML: BAN MachLA
Bee Annotation Machine Learning Algorithm

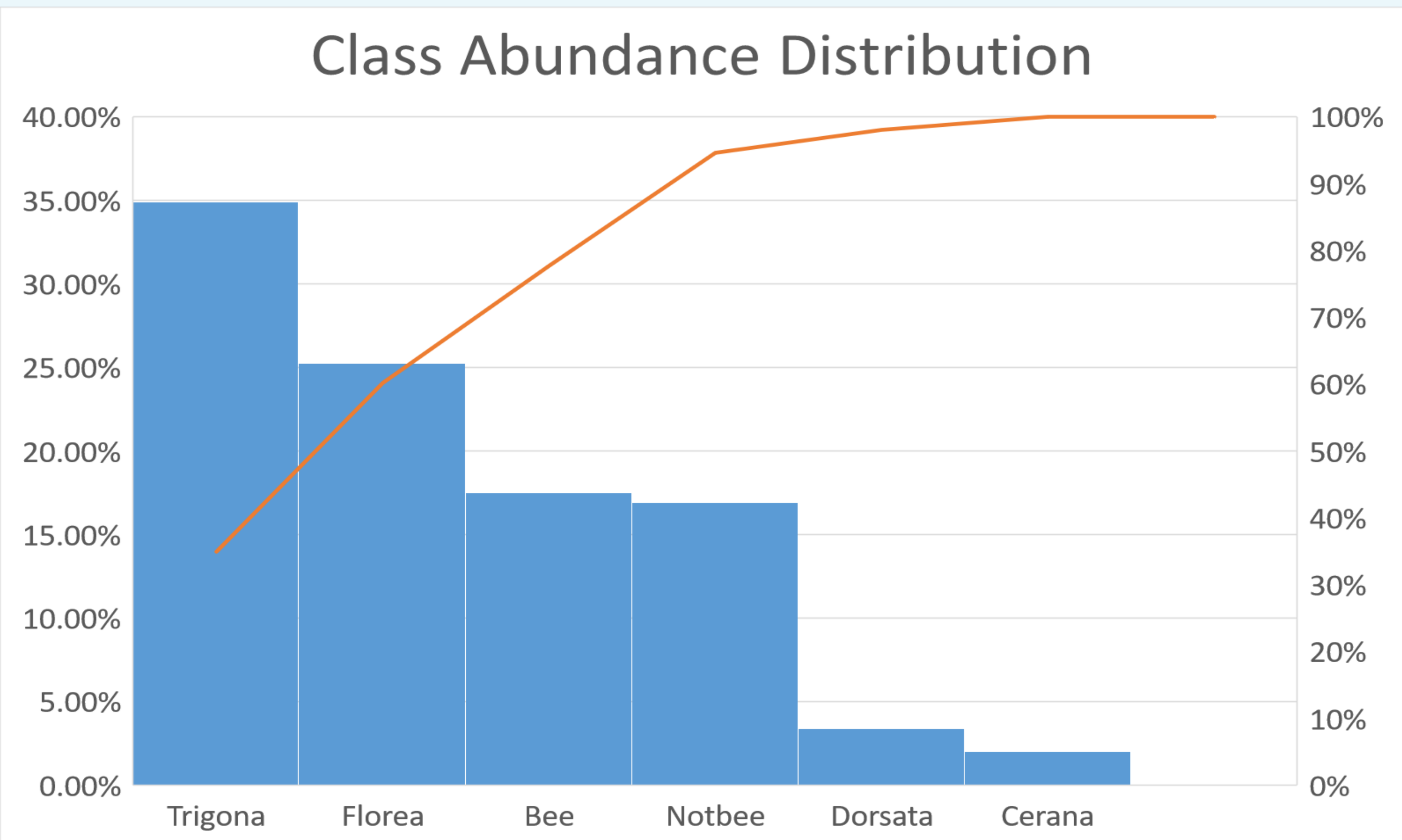
M Sreerag & C L Srinivas , TA – Aritra Mukhopadhyay
National Institute of Science Education and Research

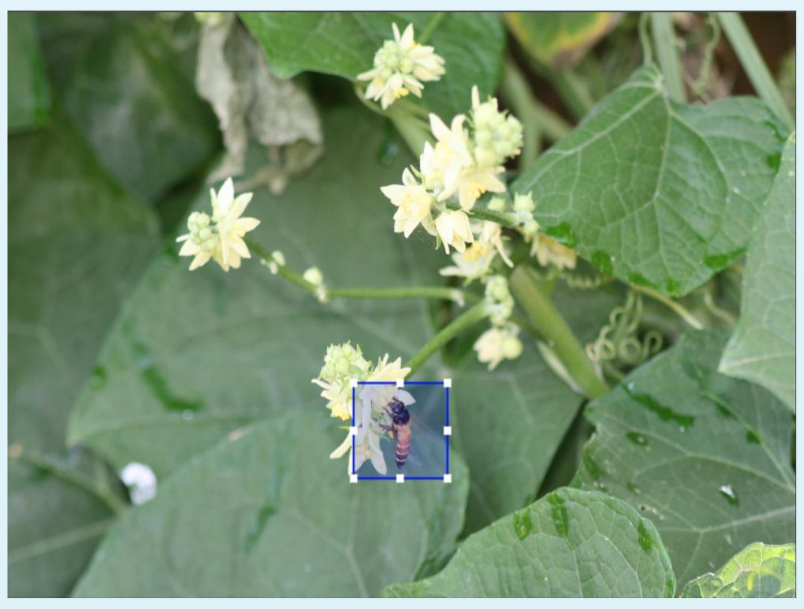
Introduction

I. Identifying and classifying living organisms is a fundamental problem in biology. An entire field of study is devoted to this particular problem, called systematics.
II. With advancement in machine learning, there has been a quantum leap in the scalability, efficiency and accessibility of classification methods.
III. The literature on object detection and image recognition machine learning algorithms is well established, with neural networks performing exceptionally in visual classification tasks.
IV. One of the major limitations of artificial intelligence is the requirement of large training data. In the particular case of bee species identification, there have scarcely been any reports of datasets on Indian bee species. The existing datasets pertain to European and American bee species. Given the morphological variation in geographically distant genera and the nature of machine learning, these model have very good chance of performing sub-optimally on an Indian dataset.

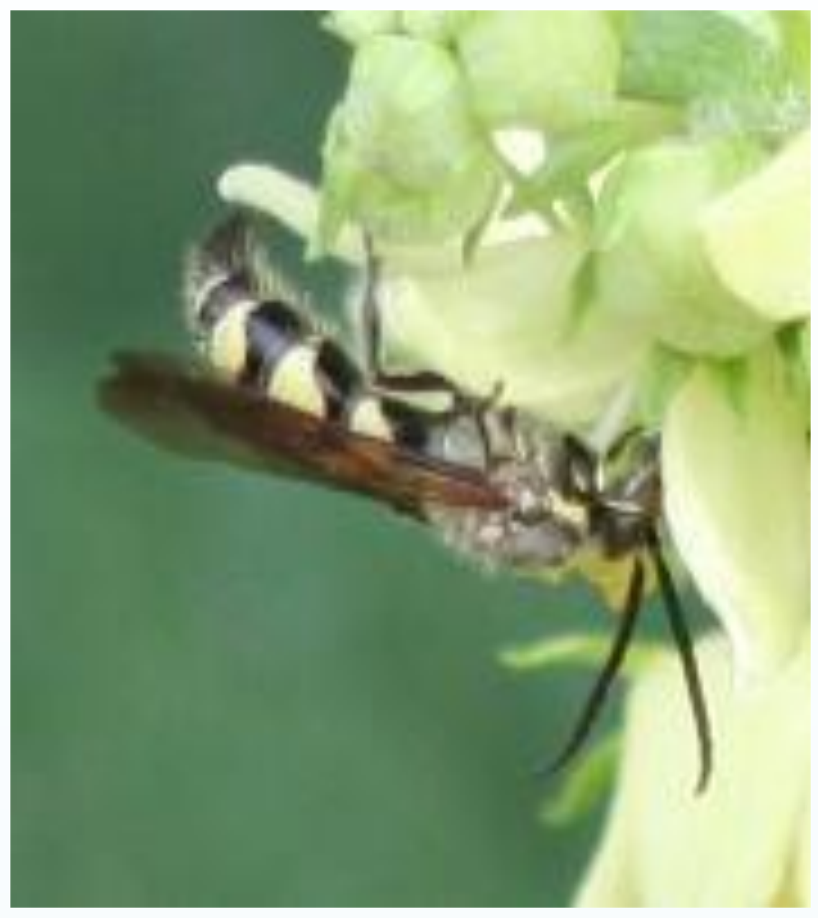
Dataset

Class	Number of Data Points	Percentage Abundance
Dorsata	111	3.40%
Cerana	66	2.03%
Florea	821	25.25%
Trigona	1135	34.90%
Bee	569	17.50%
Notbee	550	16.91%
Total	**3252**	**100%**

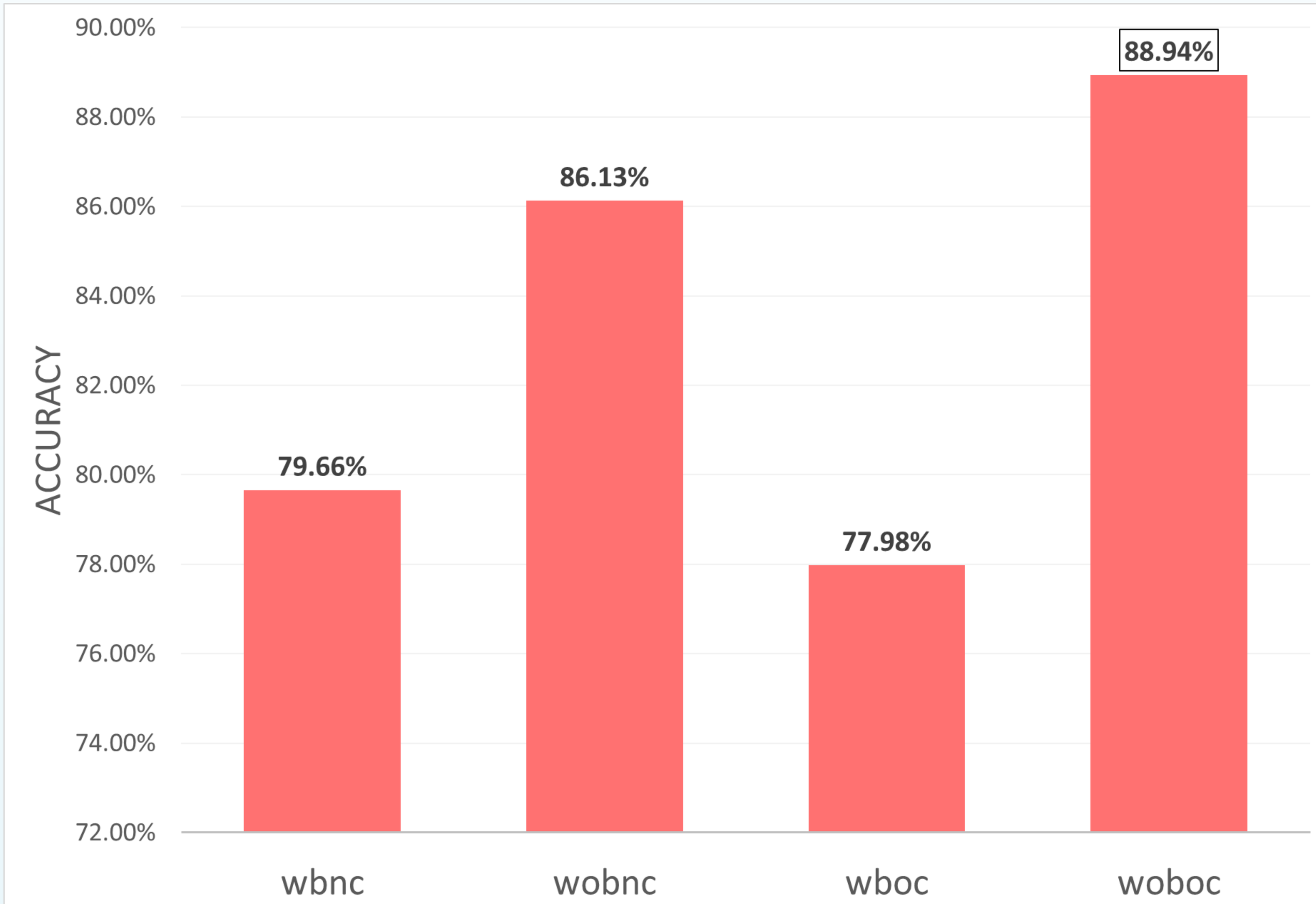








Models and Results



Dataset	Accuracy
wbnc	79.66%
wobnc	86.13%
wboc	77.98%
woboc	88.94%

- wbnc – with bee not clear
- wobnc – without bee not clear
- wboc – with bee only clear
- woboc – without bee only clear

I. Data Loader
II. Data Augmentation
III. Class Equalization
IV. Hyperparameters
   i. Model Architecture
   ii. Activation Function
   iii. Learning Rate
   iv. Number of Epochs
   v. Loss Function
   vi. Optimizer
   vii. Batch Size
   viii. Kernal Size

References

1. De Nart, D., Costa, C., Di Prisco, G. et al. Image recognition using convolutional neural networks for classification of honey bee subspecies. Apidologie 53, 5 (2022). https://doi.org/10.1007/s13592-022-00918-5
2. Kelley, W., Valova, I., Bell, D. H., Ameh, O., & Bader, J. (2021). Honey sources: neural network approach to bee species classification. Procedia Computer Science, 192, 650–657. https://doi.org/10.1016/j.procs.2021.08.067
3. Spiesman, BJ, C Gratton. RG Hatfield, WH Hsu, S Jepsen, B McCornack, K Patel, G Wang. 2021. Assessing the potential for deep learning and computer vision to identify bumble bee species from images. Scientific Reports 11:7580.

sreerag.m@niser.ac.in
srinivas.cl@niser.ac.in