applied system innovation

MDPI

Review

Buzzing Through Data: Advancing Bee Species Identification with Machine Learning

Ashan Milinda Bandara Ratnayake [1,2], Hartini Mohd Yasin [3], Abdul Ghani Naim [4] and Pg Emeroylariffion Abas [1,*]

1 Faculty of Integrated Technologies, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong BE1410, Brunei Darussalam; 22h1316@ubd.edu.bn
2 Department of Computer Science & Informatics, Uva Wellassa University, Badulla 90000, Sri Lanka
3 Faculty of Science, Universiti Brunei Darussalam, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong BE1410, Brunei Darussalam; hartini.yasin@ubd.edu.bn
4 School of Digital Science, Universiti Brunei Darussalam, Universiti Brunei Darussalam, Jalan Tungku Link, Gadong BE1410, Brunei Darussalam; ghani.naim@ubd.edu.bn
* Correspondence: emeroylariffion.abas@ubd.edu.bn

Supplementary Document

This Supplementary Document accompanies the paper "Through Data: Advancing Bee Species Identification with Machine Learning" and is designed to provide detailed insights into specific aspects of the study for enthusiastic readers seeking deeper understanding.

Figure S1 presents a focused view of only four of the nineteen tribes within the Apinae sub-family, highlighting some of the most significant groups within this classification. Among these, the Apini tribe is notable for encompassing the genus Apis, commonly known as the true honeybees. This tribe is renowned for its members' ability to produce honey and beeswax, both of which play crucial roles in commerce and agriculture due to their extensive use in food products and pollination services. Additionally, the Meloponine tribe, which consists of various stingless bee species, is distinctive for its members' lack of a stinging mechanism. These bees are celebrated not only for their production of honey, propolis, and beeswax but also for their cultural and medicinal value across diverse societies. The unique social structures and behaviors observed within the Apini and Meloponine tribes are of considerable interest for both ecological research and entomological studies, providing insights into the complex interactions within bee communities and their environments.

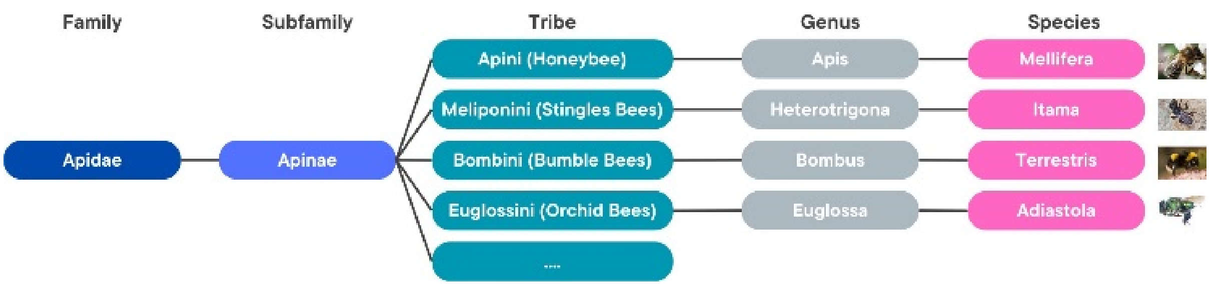

Figure S1. Four tribes among 19 tribes of the Apinae subfamily with an example species for each tribe.

The summaries of each study included in this systematic literature review are meticulously organized into tables based on the machine learning (ML) methods applied: Shallow Learning, Deep Learning, and a combination of both. These are respectively cataloged in Tables S1, S2, and S3. For clarity and ease of reference, each study has been assigned a unique publication number. This identifier is consistently used throughout the manuscript to enable quick referencing of the studies. The tables provide comprehensive details on each paper, including:

1. Publication No. - Unique identifier for each study.
2. Reference - Citation of the paper.
3. Citations - Number of times the paper has been cited.
4. Paper Title - Title of the study.
5. Publication Year - Year the study was published.
6. Data Type - Types of data used in the study (e.g., Images, Acoustic, Movement).
7. Species Focus - Focus of the identification (e.g., Stingless Bees (SB), Honeybees (HB), Other Bees, Other Species).
8. Species Count - Number of species analyzed.
9. ML Algorithms Used - Machine learning algorithms utilized in the study.
10. Study Location - Country or region where the study was conducted.
11. Includes Flying Insects - Indicates if flying insects are considered in the study.
12. Data Source - Origin of the dataset (e.g., Collected, Crowdsourced, Other Dataset).
13. Model Performance - Effectiveness of the developed model.
14. Notes - Additional remarks or important information.

These titles aim to provide clear, concise descriptions of each data point, ensuring easy navigation and comprehension of the table's contents.

Table S1. Summary of Publications Related to Shallow-Learning

Publication No.	**References**	**Cited by**	**Paper title**	**Year**	**Dataset (Images /Acoustic/ Movement)**	**Identification Focus (Stringless Bees(SB)/ Honeybees(HB)/ Other Bees/ Others)**	**Number of Species selected**	**Utilized Machine Learning algorithms**	**Target Country**	**Is consider flying insect**	**Dataset (Collected/Crows-Source/Other Dataset)**	**Performance**	**Remarks**
2	[20]	44	Biodiversity informatics in action: identification and monitoring of Bee species using ABIS	2001	Images	Other Bee (Bombus bee)	7	SVM and KDA	N/A	No	Collected	Acc.=95 %	Mobilize the bee using ice for data collection.

6	[49]	6	Machine learning approach for automatic recognition of tomato-pollinating bees based on their buzzing-sounds	2021	Acoustic	SB, Other Bees and Others	15	LR, SVM, RF, DT, and an ensemble classifier	Brazil	Yes	Collected	Acc. =73.39% MacF1= 59.06%	Uses SongMeter SM2 to collect data
7	[51]	7	Identifying Bee Species by Means of the Foraging Pattern Using Machine Learning	2018	Movement	SB	2	MLP, SVM, RF	Brazil	Yes	Collected	Acc. = 87.41%	IR transponders stick to thoracic of the bee
11	[47]	31	Evaluating classification and feature selection techniques for honeybee subspecies identification using wing images	2015	Images	HB	26	LDA, KNN,Logistic, NB, DT, MLP and SVM	N/A	No	N/A	Acc.= 65.15%	
13	[12]	23	Automated classification of bees and hornet using acoustic analysis of their flight sounds	2019	Acoustic	HB, Other Bee and Other	4	SVM	Japan	Yes	Collected	Avg. Precision 0.89	Record sound using AT9905, Audio-Technica)
16	[50]	2	Automated detection of the yellow-legged hornet (Vespa velutina) using an optical sensor with machine learning	2023	Movement	HB, Other Bee and Other	7	RF	Spain	Yes	Collected	Avg. Acc.= 80.1 ± 13.9%	
19	[46]	24	A reference process for automating bee species identification based on wing images and digital image processing	2014	Image	Other Bee (Orchid bee)	5	LDA,Logic Regression,SVM,KNN, Other Methods	N/A	No	Collected	Acc. 87.68%	
20	[44]	4	A fully automatic classification of bee species from wing images	2021	Images	HB ,SB , Other Bee , Other	48	KNN	N/A	No	Collected	Acc. Species = 96% Genus = 99%	

25	[48]	1	Hierarchical classification of pollinating flying insects under changing environments	2022	Movement	HB, SB and Other	7	KNN, NB, RF, and SVM	Brazil	Yes	Collected	Avg. hierarchical F-measure= 94.08	

Table S2. Summary of Publications related to Deep-Learning

Publication No.	Reference	Cited by	Paper title	Year	Dataset (Images /Acoustic/ Movement)	Identification Focus (Stringless Bees (SB)/ Honeybees (HB)/ Other Bees/ Others)	Number of Species se-	Utilised Machine Learning algorithms	Target Country	Is consider flying insect	Dataset (Collected/Crowssource/Other Dataset)	Performance	Remarks
1	[42]	0	DY-RetinaNet Based Identification of Common Species at Beehive Nest Gates	2022	Images	HB and Others	3	Improved RetinaNet(DY RetinaNeT)	China	No	Collected	Avg. Acc. : 97.38%	Do Object Detection
4	[43]	5	Classification of Ecological Data by Deep Learning	2020	Images	Other Bee and Other	29	LeNet-5, AlexNet, ResNet50, Inception v3, Inception-ResNetV2, VGG-16 and VGG-19	N/A	No	N/A	Acc. Of Butterfly:96.88 %, Bee : 92.90%	
5	[13]	7	Image recognition using convolutional neural networks for classification of honeybee subspecies	2022	Images	HB	7	InceptionResNet V2, InceptionNet V3, MobileNet V2 and ResNet 50	Italy	No	Other dataset	Avg. F1 Score: 0.99	
9	[38]	1	A concatenated approach based on transfer learning and PCA for	2021	Images	HB and other	4	VGG, ResNet, XceptionNet, and EfficientNet	N/A	No	Other datasets	Acc. 96.89% AUC: 0.9973	Model based on transformer.

			classifying bees and wasps										
10	[16]	31	Assessing the potential for deep learning and computer vision to identify bumble bee species from images	2021	Images	Other Bee (Bumble Bee)	36	ResNet, Wide ResNet, InceptionV3, and MnasNet	United States and Canada	No	Other datasets	Acc. : 91.7 %	BeeMachine web application utilized this model
15	[40]	33	Image-based species identification of wild bees using convolutional neural networks	2020	Images	HB , Other Bee ,Other	129	B-CNN, MobileNet and Inception-ResNet	Germany,Brazil, United States and China.	No	Other Dataset	top-1 Acc. of 93.95% and a top-5 Acc. of 99.61%	
17	[34]	2	A Deep Learning Approach to classify the Honeybee Species and health Identification	2021	Images	HB	7	Inception v3	N/A	No	Other dataset	Acc. : 86%	
18	[3]	0	Automatic acoustic recognition of pollinating bee species can be highly improved by Deep Learning models accompanied by pre-training and strong data augmentation	2023	Acoustic	HB and Other Bee	16	Pre-trained Audio Neural Networks(PANNs), and EfficientNet V2	Chile	Yes	N/A	Macro F1-Score: 58.04 (± 2.47)	Zoom H4n Pro Handy Recorder is used to collect data
21	[32]	0	Honey sources: Neural network approach to bee species classification	2021	Images	HB, Other Bee	3	Faster R-CNN with Resnet 101+FPN backbone	N/A	No	Collected	Acc. : 91%	Do Object Detection
23	[37]	0	BeeNet: An End-To-End Deep Network For Bee Surveillance	2023	Images	HB	11	ResNest, EfficientNet, and NoisyStudent,BeeNet	N/A	No	Crowed sourced dataset and	Acc. : 92.45%	

											Other dataset		
24	[33]	0	Research on Mini-EfficientDet Identification Algorithm Based on Transfer Learning	2022	Images	Other Bee and other	3	EffientDet with one layer of BiFPN	N/A	No	Other dataset	Acc. of Chinese Bee: 98.66%	Do Object Detection
26	[36]	1	Image segmentation of meliponine bee using Mask-RCNN	2020	Images	SB	1	Faster R-CNN	Malaysia	No	Collected and Other dataset	Acc. :74%	Do Object Detection to differentiate SB and background

Table S3. Summary of Studies related to combination of Shallow- and Deep-Learning

Publication No.	References	Cited by	Paper title	Year	Dataset (Images /Acoustic/ Movement)	Identification Focus (Stringless Bees(SB)/ Honeybees(HB)/ Other Bees/ Others)	Number of Species se-	Considered Machine Learning algorithms	Target Country	Is consider flying insect	Dataset (Collected/Crows-Source/Other Dataset)	Performance	Remarks
3	[39]	103	A Vision-Based Counting and Recognition System for Flying Insects in Intelligent Agriculture	2018	Images	HB and Others	6	YOLO and SVM	N/A	No	Collected	Acc.: 90.18%	Do Object Detection
8	[35]	5	AI in apiculture: A novel framework for recognition of invasive insects under unconstrained flying conditions for smart beehives	2023	Images and Movement	HB and Others	3	DT,Ensemble, SVM, KNN, xception,googlenet and bagged trees	Pakistan	Yes	Collected	Acc.: 97.1%	Intel's Real sense D435 depth camera use to capture data and calculate the 3D trajectories
12	[41]	6	Using the Software DeepWings© to Classify Honey Bees across	2022	Images	HB	3	DeepWing	Portugal, Spain,	No	Collected	M-lineages: 71.4%	

			Europe through Wing Geometric Morphometrics						Belgium, France, Ireland, Poland, Russia, Sweden, Switzerlan, United Kingdom, Croatia, Hungary, Moldavia, Romania, and Slovenia			C-lineages: 97.6% Iberian honey: 89.7% dark honey: 41.1% Carniolan honey: 88.3%
14	[45]	7	DeepWings©: Automatic Wing Geometric Morphometrics Classification of Honey Bee (Apis mellifera) Subspecies Using Deep Learning for Detecting Landmarks	2022	Images	HB	26	Faster R-CNN NAS,Faster R-CNN Inception Resnet v2 Atrous Coco, YoloV3, SSD MobileNet v1 FPN coco, Unet and SVM	N/A	No	Collected and Other Dataset	SSD MobileNet v1 FPN coco model achieve 0.975 of mAP@0.5, landmark detection(U-

												NET) is 0.943 SVM achieved 86.6%	
22	[52]	7	Image-Based Automated Species Identification: Can Virtual Data Augmentation Overcome Problems of Insufficient Sampling?	2022	Images	Other Bee , Other	4	VGG16 and SVM	N/A	No	Other Dataset	Accuracy of each class: Pleophylla: 80.07% Schizonycha: 85.16% Osmia:85.90% Parides: 98.06%	Utilized Style-GAN to generate new data from the existing images and Generated heatmaps of specimen images to show contribution to decision making