Title

Big-Bee: Towards a More Accurate Hair Quantification Pipeline

Authors

Alexander, Nicholas
Seltmann, Katja C

Publication Date

2023-05-17

Big-Bee: Towards a More Accurate Hair Quantification Pipeline

Introduction

Bees are essential for pollinating plants, as well as maintaining a significant amount of food production around the world. However, they're declining in both numbers and diversity but the known causes for these declines are limited, partly due to the lack of sufficient data on the anatomical and behavioral traits that may either increase bees' chances at survival or put them at risk under human-induced environmental changes(e.g. habitat loss, climate change).

Trait digitization is the measurement of traits through images. By teaching a computer to automate this task, trait data are efficiently collected and new correlations are discovered faster. Moreover, automating this task on bee traits allows for quicker discoveries regarding a trait's role for resilience or vulnerability.

Digitizing anatomical traits via automation first requires one or more computer vision models capable of identifying the desired trait within an image. By training a model to perform binary segmentation on a specific trait, it learns to identify the pixel regions of where that trait is located. Having only one model that identifies a specific anatomical trait can be limiting to the capabilities of trait digitization. With two or more models each trained to identify a different anatomical trait(or just one model trained to identify all desired traits) and potentially a model for identifying an entire bee, traits get digitized through a series of arithmetic operations on pixels of interest. Since bee hair is critical for several functions like pollination, staying warm, and sensing its surroundings, it's important to find efficient ways of measuring bee hairiness.

A new index for measuring bee hairiness is now possible through anatomical trait digitization. Given a binary mask of a bee and another binary mask for the same bee's hair, the hair-to-bee surface density ratio is calculated by dividing the sum of all pixels in the hair mask with the sum of all pixels in the bee mask. Since these binary masks only contain values of one for the desired regions and zero otherwise, the sum of all pixels in the hair mask should always range from zero to the sum of all pixels in the bee mask. Therefore, the index should always range from zero to one; zero meaning not a single hair on the bee and one meaning the entire bee is covered with hair. Lastly, the pixel map for the hair mask must always land within the regions of the whole bee mask. While a human can manually create these masks and run them through the index function, the goal of digitizing bee traits is to incorporate automation in the pipeline to save significant amounts of time in the long run. Hence, two binary segmentation models must be trained on different traits(whole bee and hair) in order to fully automate this task.

The datasets for developing the two models were manually created using a commercial image editor and contain binary masks of desired bee traits used for hair quantification(whole bee masks and hair masks). Creating hair masks for bees can be extremely time consuming since the hair tends to be spread out around other noisy features. The final datasets contained 315 bee masks and 199 hair masks.

Methods were inspired by previous findings on human hair segmentation in the wild. The term “in the wild” implies images with the subject(s) at an uncontrolled setting, meaning the training images contain high variance in background noise, along with other potential features that arent hair. A method called coarse-to-fine was used, where each image was sliced into squares and fed into a pretrained model that returns a white square if the input was a patch of hair or a black square if the input contained background noise. For the bee hair quantification problem, images are sliced for noise reduction and to help the hair detection model with generalizing at the lowest level.

Previous pipelines for human hair segmentation don't translate well on bee hair since images containing human hair tend to be blotched into one region. A bee generally contains hair distributed through various parts of its body. Thus, with a closeup shot of a bee it is apparent that its hair can't be identified as one blotch.

Binary segmentation on bee traits was achieved with the use of TernausNet. TernausNet modifies the U-Net architecture by using a pre-trained VGG-11 model(ImageNet) as its encoder and further trains the modified U-Net on the Carvana dataset. By using the pre-trained model as a foundation for training on the anatomical traits data, remarkable results are achieved with small amounts of data..

Methods

The first was model trained on the whole bee mask dataset and learned to identify where the bee is by performing pixel-based binary classification. The output is a binary image mask with pixel values of 1 if the pixel is part of a bee or 0 if the pixel is outside of the bee regions. Similarly, the second model learned to classify pixels that fall within the hair regions. Each training pipeline conducted a 60/20/20 split on their respective dataset for use as training/validation/testing. Since the training sets are smaller than the already small datasets, a random set of image augmentation techniques are uniquely applied to each image during every round of training. Possible combinations to choose from include vertical flipping, horizontal flipping, random rotation, random blurring, and RGB color shifting. All techniques contain a probability of 0.5 to be included in each images' set of augmentations and are applied with randomized parameters if applicable(rotation intensity, color shift intensity, blur intensity).

Since the hair detection model doesn't concern understanding the shape of a bee, the images/masks used for training are random 300x300 crops of the original image. Consequently, to receive optimal performance on new predictions of whole bee images, the prediction pipeline should slice input images into smaller pieces(at least 300x300) and pass each slice individually through the hair detection model before restitching the outputs back to the original shape.

Results

Table 1

Model	Best Iteration	Train Loss	Validation Loss
Bee	303	0.0474	0.0477
Hair	219	0.224	0.23

Table 2

Model	Train F1	Test F1	Train Acc	Test Acc
Bee	0.98	0.972	0.985	0.98
Hair	0.578	0.564	0.932	0.897
Hair > 0	0.764	0.701	0.91	0.873

$$F1 = \frac{TP}{TP + \frac{1}{2}(FP + FN)} \qquad Accuracy = \frac{(TP+TN)}{(TP+FP+FN+TN)}$$

While the bee model scored well, there's room for improvement on the hair masking model. The hair model plateaued during the 219th iteration of training and would either slowly begin to overfit or converge. The best models were chosen by identifying the last iteration where the validation and training sets stably trained together and contained the lowest validation loss that was greater than the training loss.

F1 score is the primary evaluation metric for assessing these models on their test sets. **Table 2** contains an extra row of data for the hair model labeled as, *Hair > 0*, and it reevaluated the test metrics strictly based on images with ground truth masks that sum up to at least one. In other words, *Hair > 0* filters out all images with an F1 score of zero before recalculating the mean F1 and accuracy among the rest of the images. All of the dropped images contain absolutely no hair and have binary masks containing all true negatives(wings, eyes, background). The training and testing sizes for evaluating the new scores reduced from 119/41 images to 90/33 respectively.

Discussion

To determine the skewed F1 score's validity, the false positive rate must be assessed. If images with true negative binary masks get predicted with lots of false positives, the model hasn't learned to generalize well. There isn't a standardized threshold value when assessing these metrics and different tasks will hold different expectations. A low false positive rate on masks with zero F1 along with a higher F1 score on the rest of the images suggests the model is generalizing well.

While these values did assist the model with determining which bee traits are not hair, a potential risk of training these data for long periods of time include overfitting on non-hair features, while failing to generalize on hair. Future steps include building custom loss functions to reward true negatives less, assigning low probabilities to these images to limit their frequency during training, and expanding the dataset to include more images that are well diverse in both true positives and true negatives.

Conclusion

Since bees continue to decline with insufficient data for researchers to understand why, new ways for automating trait digitization must be explored in order to expedite data collection and new findings. Modern practices in computer vision and deep learning open new doors for digitizing bee traits through images. With just a small amount of data, computers have become capable of performing human-like tasks and at much faster rates. As new pipelines are developed for digitizing bee traits, more data and models will work side by side towards exponential progress at understanding bees at a deeper level.

References:

Seltmann KC, Allen J, Brown BV, Carper A, Engel MS, Franz N, Gilbert E, Grinter C, Gonzalez VH, Horsley P, Lee S, Maier C, Miko I, Morris P, Oboyski P, Pierce NE, Poelen J, Scott VL, Smith M, Talamas EJ, Tsutsui ND, Tucker E (2021) Announcing Big-Bee: An initiative to promote understanding of bees through image and trait digitization. Biodiversity Information Science and Standards 5: e74037. https://doi.org/10.3897/biss.5.74037

M. Svanera, U. R. Muhammad, R. Leonardi and S. Benini, "Figaro, hair detection and segmentation in the wild," *2016 IEEE International Conference on Image Processing (ICIP)*, 2016, pp. 933-937, doi: 10.1109/ICIP.2016.7532494.

Iglovikov, Vladimir I. and Alexey A. Shvets. “TernausNet: U-Net with VGG11 Encoder Pre-Trained on ImageNet for Image Segmentation.” *ArXiv* abs/1801.05746 (2018): n. pag.