Bee Image Classification

using a CNN to detect *varroa mites*

By: Mahdi Shadkam-Farrokhi

The Problem

- **Honey bees contribute** nearly **$20 billion** to the value of U.S. crop production [1]
- Beekeepers continue to **lose** up to **45 percent of their hives** annually [2]
- ***Varroa mites*** are a major culprit [2]
- *Varroa mites* are **super effective vs. honey bees**
  - Drain body fluids directly
  - Transmit deadly virus

varroa destructor

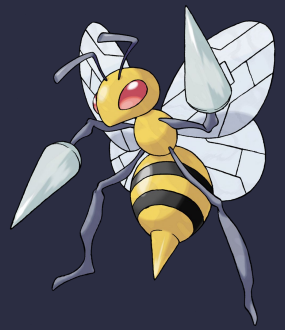
How can we help?

- Improve beekeeper's ability to detect *varroa* infestation
  - **Quicker detection = quicker treatment**
  - Taking a picture is quick and convenient!
- Plan of attack:
  - Use **Honey Bee Annotated Image Dataset** (on Kaggle) [3]
  - Train a Convolutional Neural Network (CNN) to **classify a bee image as having a *varroa mite* or not** (binary classification)
  - Using accuracy as metric
  - **BONUS:** Create a practical application with the model

Glimpse of the data

- ~4,400 bee images
  - 25% *varroa*
  - **75% Healthy (baseline)**
- Observations:
  - Different image sizes
  - Different orientations
  - Dark Exoskeletons
  - Background confusion
  - Different resolutions
  - Partial bee images
  - Different photographers?

Model Preparation



- Addressing issues:
  - ~~Different image sizes~~ **Resized to a standard WxH**
  - ~~Different orientations~~ **Rotation + Mirror Transformation**
  - ~~Dark Exoskeletons~~ **Brightness Transformation**
  - ~~Background confusion~~ **Brightness Transformation**
  - **Different resolutions**
  - **Partial bee images**
  - **Different photographers?**

- Network topography:
  - [3 x (2 Conv + 1 MaxPool)] + [2 Dense]

Model Evaluation

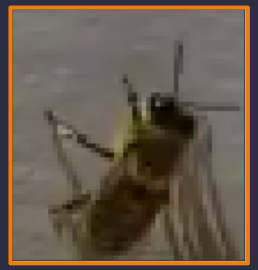
- 98.9% training accuracy
- **99.5% testing accuracy**
  - Background confusion still an issue
- Is the CNN actually detecting *varroa*?
  - Neural Networks are “black box” models
  - We can look at convolutional layer weights!

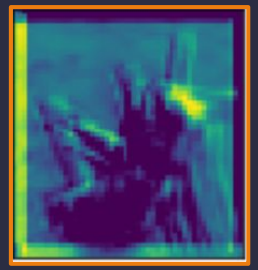
Exploring the CNN

Healthy bee

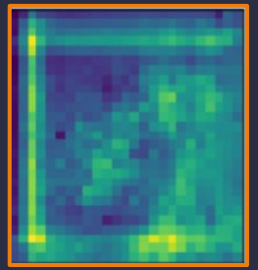
Conv #1
(one filter shown)

Conv #3
(one filter shown)



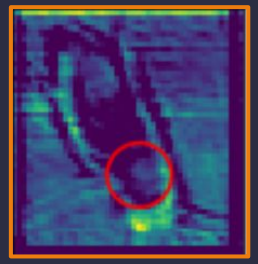


Conv #2 +
MaxPool

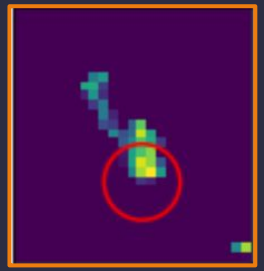


Conv #4 + ...

Varroa bee



Conv #2 +
MaxPool



Conv #4 + ...

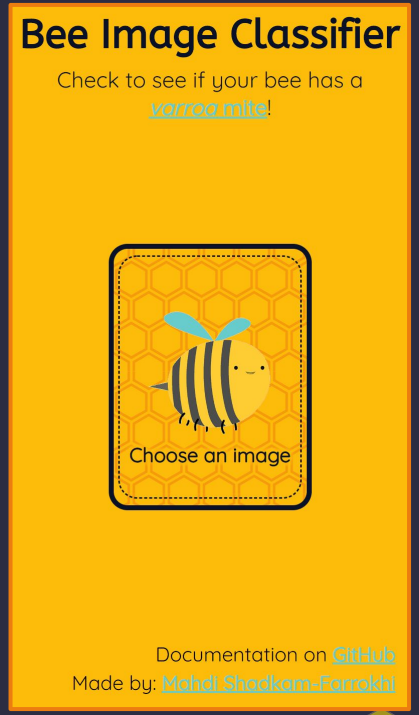
Conclusion

- Successfully trained a CNN to perform better than baseline
  - 98.9% training and **99.5% testing accuracy**
- The model seems to actually detect *varroa*!
- One weakness of the CNN!
  - Background confusion
  - Take photos using contrasting background
- Grain of salt...
  - Small dataset!
  - Single source of *varroa* images = bias
  - Resizing = reduced data quality

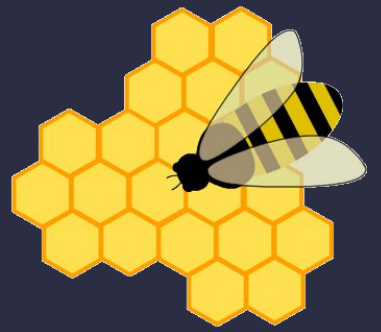
Recommendations + Future Work

- Revisit with more robust dataset w/same WxH
- Bee Classifier mobile web application [4]
  - Validate and test model in production-like setting
- Impractical for beekeepers to photograph individual bees!
  - Train CNN to crop from photos with many bees
  - Feed those images into our CNN

https://bit.ly/2KYNPlP



Sources



1. ABF - Pollination Facts
2. CMS business - Agriculture and Bees: What Consumers Need to Know
3. Honey Bee Annotated Image Dataset (Kaggle)
4. Bee Image Classifier (proof of concept)


bee
HAPPY