CS3 Hook Document

Henry Allen DS4002 - 12PM, SpotCheck (PS3)

Have you ever been convinced that the ladybug you just picked up peed on you? Smelled foul? Or that there were more and more of them by the window every year? You may not be dealing with a ladybug after all. The invasive Asian lady beetle (*Harmonia axyridis*) has become a far more prevalent home intruder in recent years. Squashing the infamous and recognizable Spotted Lanternfly without second thought is pretty easy, but crushing what could be an innocent, native ladybug (*Coccinella novemnotata*) carries a lot more guilt.

Imagine being the person who can finally settle the question: Is this a harmless native species, or an invasive look-alike? With our model, SpotCheck, you step into that role. SpotCheck removes the guesswork from invasive species identification. Using a Convolutional Neural Network (CNN) trained on iNaturalist image data, it can detect which "ladybug" species appears in a photo with 81% accuracy. By exploring the provided notebooks and code in our GitHub repository, you will be able to reproduce the model that could later be applied to images of your own.

Beginning with scripts/1_scraping.ipynb, you will scrape image data directly from the iNaturalist citizen science platform. From there, the workflow guides you through scripts/2_exploratory plots, scripts/3_preprocessing, then model building and evaluation in scripts/4_model. Each step reveals how raw images become a functioning species-identifier through feature extraction, normalization, and iterative training.

In the model, the CNN will detect species-specific patterns automatically, no entomology background required. Wing shape, head markings, and spot geometry are all learned from pixel data, enabling the model to help determine when to squash and when to save. Your task is to use our repository, engage with the dataset and workflow, and see how a real species-classification model comes together.

GitHub Repository: https://github.com/henry5250/SpotCheck/tree/main