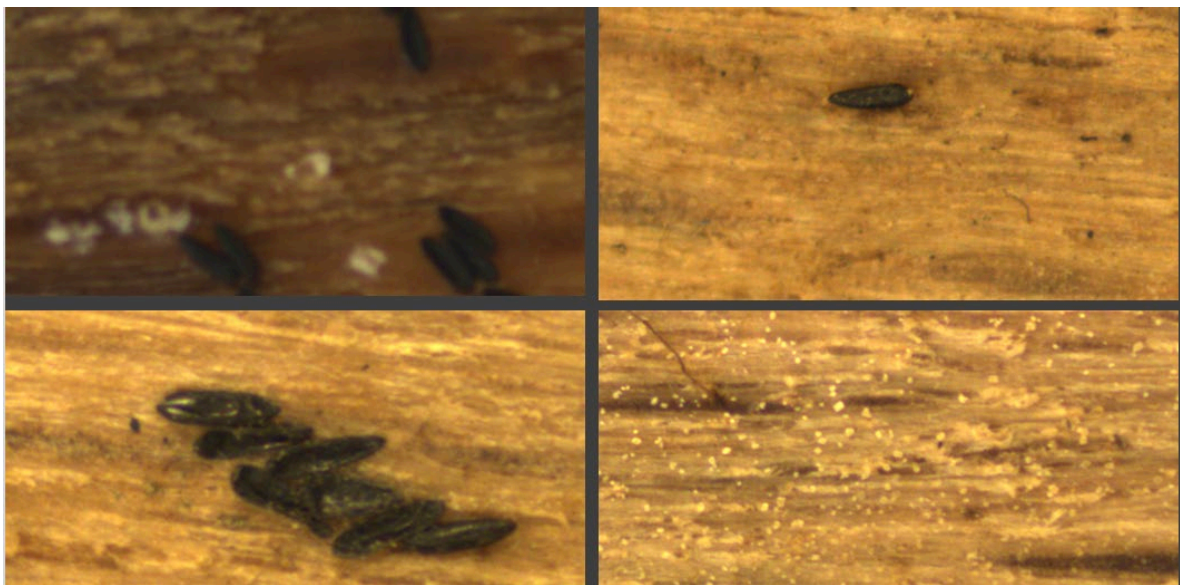
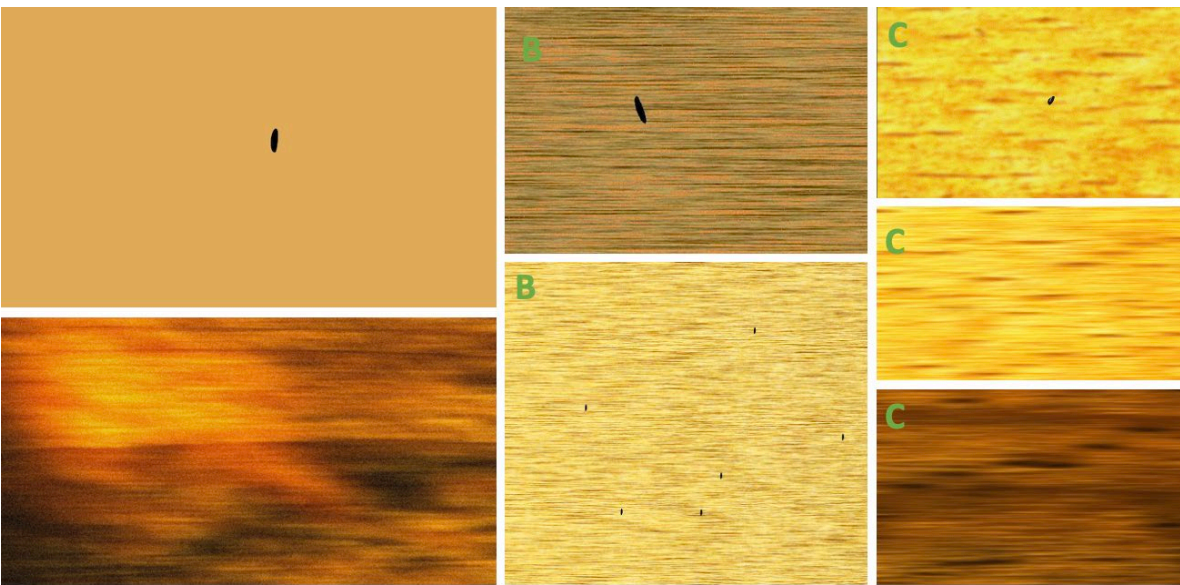


Identification of Tiger Mosquito Eggs using Machine Learning

Overview of some sample wooden pieces and eggs



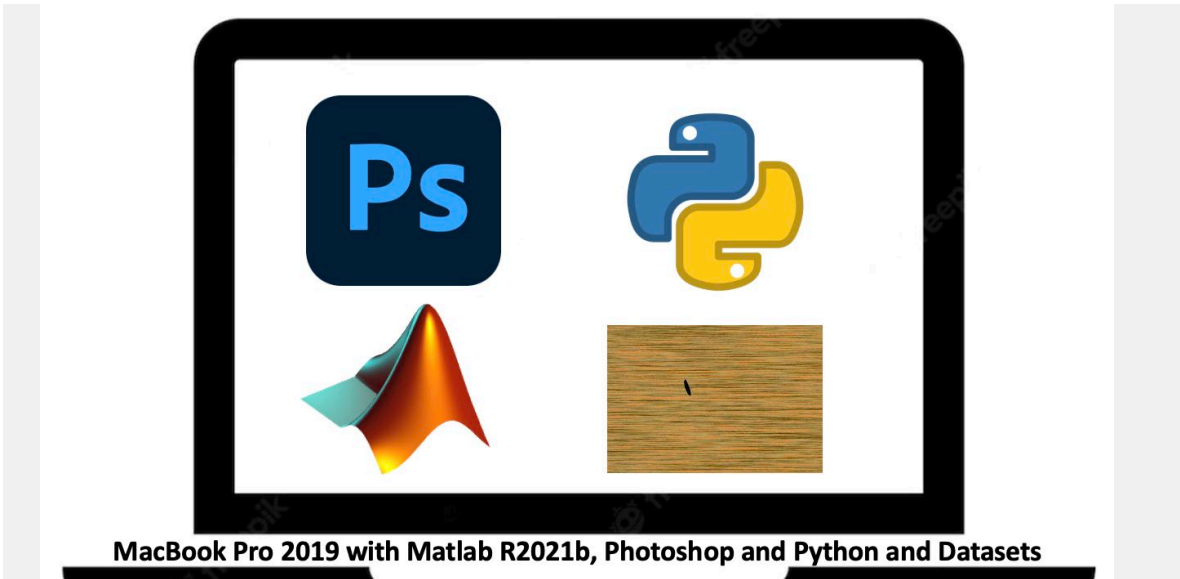
Overview of some sample artificial datasets



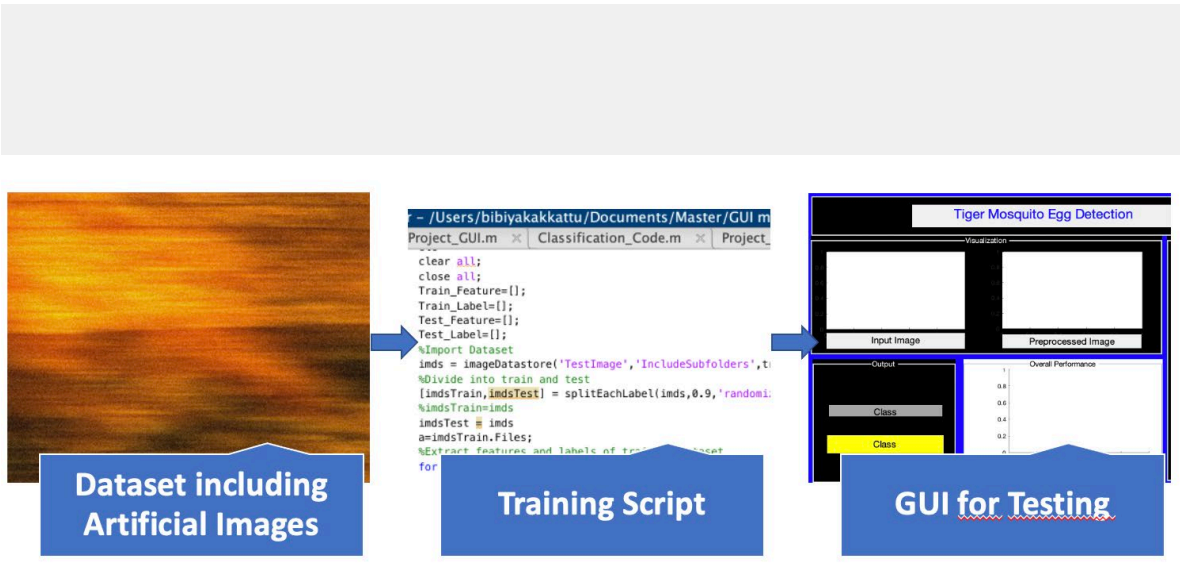
Artificial Tiger Mosquito Eggs Drawings from the Adobe Photoshop



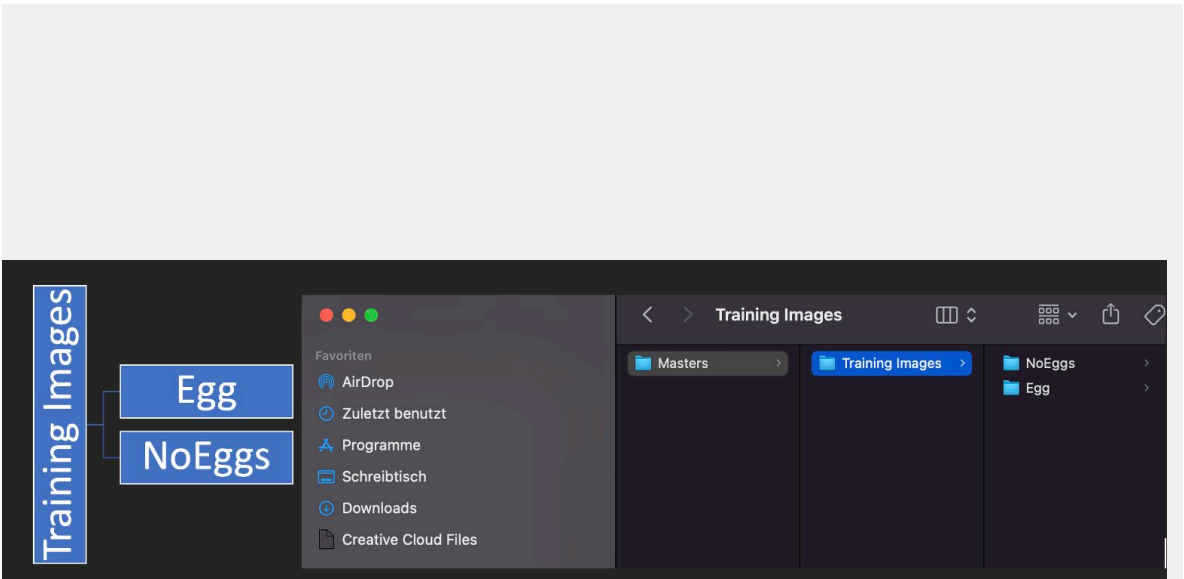
Overview of computational experiment setup



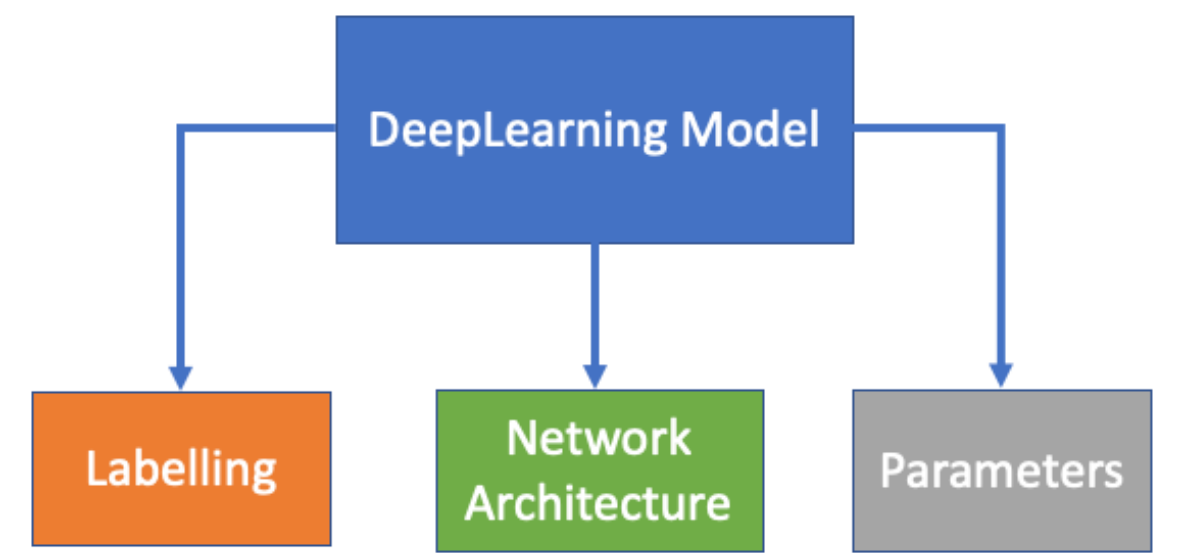
Overview of computational experiment setup



Hierarchy of the folders containing labelled training datasets



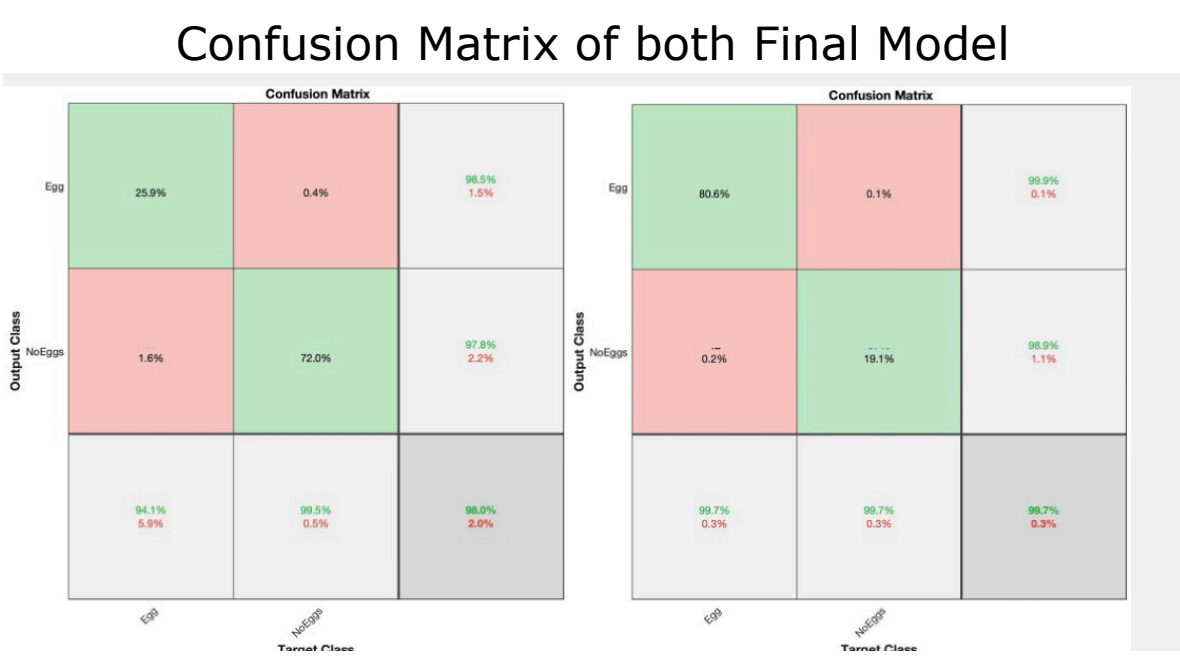
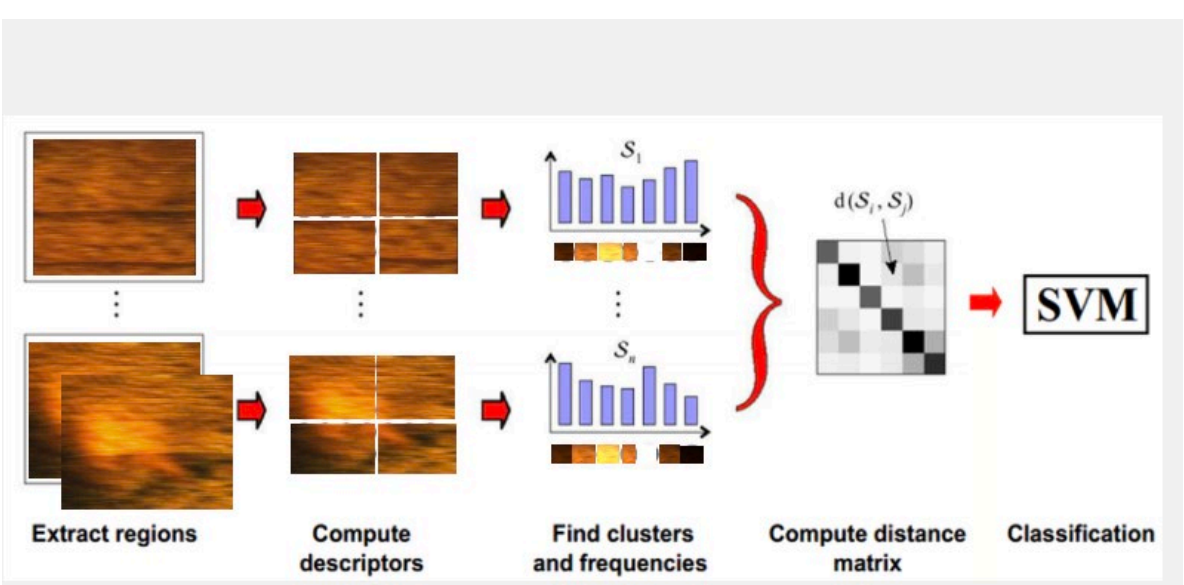
Experiments Categorization



Comparison of accuracies of different models in MATLAB's classification learner app

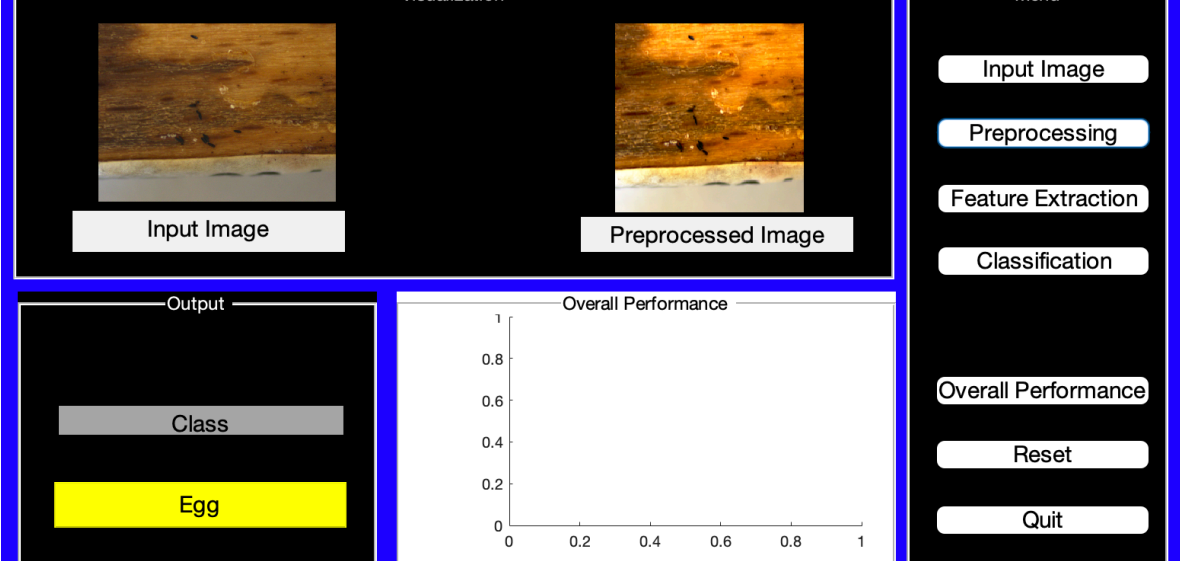
1.2 Tree	Accuracy (Validation): 71.5%	1.13 SVM	Accuracy (Validation): 84.3%	1.23 Ensemble	Accuracy (Validation): 68.5%
Last change: Medium Tree	4096/4096 features	Last change: Medium Gaussian SVM	4096/4096 featu	Last change: Subspace Discriminant	4096/4096 featu
1.3 Tree	Accuracy (Validation): 78.0%	1.14 SVM	Accuracy (Validation): 78.4%	1.24 Ensemble	Accuracy (Validation): 81.9%
Last change: Coarse Tree	4096/4096 features	Last change: Coarse Gaussian SVM	4096/4096 featu	Last change: Subspace KNN	4096/4096 featu
1.4 Linear Discriminant	Accuracy (Validation): 78.1%	1.15 KNN	Accuracy (Validation): 82.8%	1.25 Ensemble	Accuracy (Validation): 73.8%
Last change: Linear Discriminant	4096/4096 features	Last change: Fine KNN	4096/4096 features	Last change: RUSBoosted Trees	4096/4096 features
1.5 Quadratic Discriminant	Accuracy (Validation): 78.1%	1.16 KNN	Accuracy (Validation): 80.4%	1.26 Neural Net...	Accuracy (Validation): 81.3%
Last change: Quadratic Discriminant	4096/4096 featu	Last change: Medium KNN	4096/4096 features	Last change: Narrow Neural Network	4096/4096 featu
1.6 Logistic Re...	Accuracy (Validation): 58.5%	1.17 KNN	Accuracy (Validation): 72.8%	1.27 Neural Net...	Accuracy (Validation): 79.8%
Last change: Logistic Regression	4096/4096 features	Last change: Coarse KNN	4096/4096 features	Last change: Medium Neural Network	4096/4096 featu
1.7 Naive Bayes	Accuracy (Validation): 77.0%	1.18 KNN	Accuracy (Validation): 80.6%	1.28 Neural Net...	Accuracy (Validation): 80.6%
Last change: Gaussian Naive Bayes	4096/4096 featu	Last change: Cosine KNN	4096/4096 features	Last change: Wide Neural Network	4096/4096 featu
1.8 Naive Bayes	Accuracy (Validation): 79.4%	1.19 KNN	Accuracy (Validation): 79.6%	1.29 Neural Net...	Accuracy (Validation): 81.1%
Last change: Kernel Naive Bayes	4096/4096 features	Last change: Cubic KNN	4096/4096 features	Last change: Blayered Neural Network	4096/4096 fea
1.9 SVM	Accuracy (Validation): 82.3%	1.20 KNN	Accuracy (Validation): 82.3%	1.30 Neural Net...	Accuracy (Validation): 79.6%
Last change: Linear SVM	4096/4096 features	Last change: Weighted KNN	4096/4096 features	Last change: Trilayered Neural Network	4096/4096 fea
1.10 SVM	Accuracy (Validation): 83.0%	1.21 Ensemble	Accuracy (Validation): 77.4%	1.31 Kernel	Accuracy (Validation): 83.0%
Last change: Quadratic SVM	4096/4096 features	Last change: Boosted Trees	4096/4096 features	Last change: SVM Kernel	4096/4096 featu

Comparison of accuracies of different models in MATLAB's classification learner app

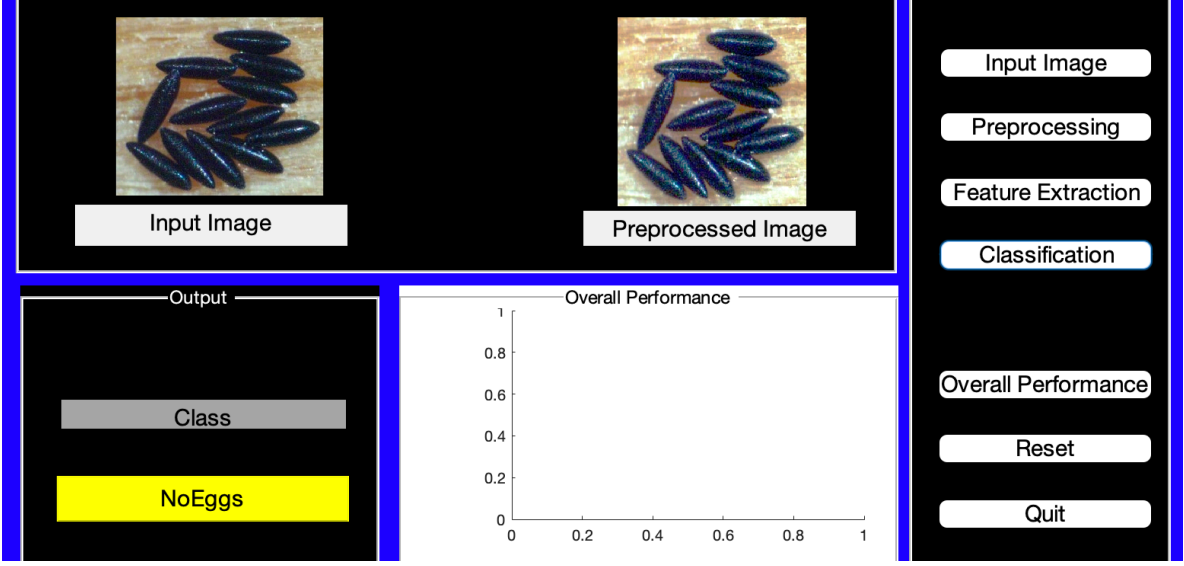


Left Model: only Real Data, Right Model: including Artificial Dataset

Correct Prediction of Second model for a noisy image with Eggs



Wrong Prediction of even larger Eggs on a wooden background



Problem statement

In order to help The Swiss Tropical and Public Health Institute (Swiss TPH) work of monitoring the population of the Tiger mosquito eggs in wooden pieces which they collect every six months from various locations in Switzerland to inspect the presence of Tiger Mosquito eggs. The current process is manual, where a human being observes each wooden pieces under the microscope to detect and count the number of eggs. This manual process is very time consuming and error prone.

Therefore, in the Master Thesis Project, we have tried to solve the sorting of images of wooden pieces into two categories with the help of a Machine Learning Model. The first category is all the wooden piece images containing the Tiger Mosquito eggs and the second category is all the wooden piece images free of any Tiger Mosquito eggs. This project aims to leverage various Machine Learning toolboxes available in MATLAB to predict and detect the eggs in the images of wooden pieces.

The Machine Learning Model must have an accuracy rate above 95% and the testing time should be less than five Minutes.

FH Zentralschweiz

Solution

We have created two models for this problem. Both the models use SURF features and a SVM based classifier for prediction. Difference between both the models is that the first model has been trained on the real-life images while the second model also leverages artificially generated dataset to overcome lack of enough training data. First model has an overall accuracy of 98% and takes 3-5 seconds for the prediction of a new image on commodity hardware. The second model trained additionally on the artificial training data has an overall accuracy of 99.7% but is much slower to predict new images on commodity hardwar with a classification time of 33-50 seconds time per image. Both the models are designed to predict Tiger Mosquito eggs on top of wooden surfaces and work well if the Egg size is less than 95x76 DIM.

Some of the limitations of the models are listed below.

- While the current model can differentiate and detect eggs when the egg size is around 95x76 pixel, it struggles to correctly detect when the egg size starts become bigger.
- The model can classify the eggs on different types of wooden patterns. Unfortunately it is not robust to identify if the egg lays on other type of material like plastic, clothes etc.

Bibiya Kakkattu

Supervisors:
Prof. Dr. Christoph Eck
Prof. Dr. Peter Gruber

Expert:
Prof. Dr. Peter Kolb

Cooperation Partner:
The Swiss Tropical and Public Health Institute