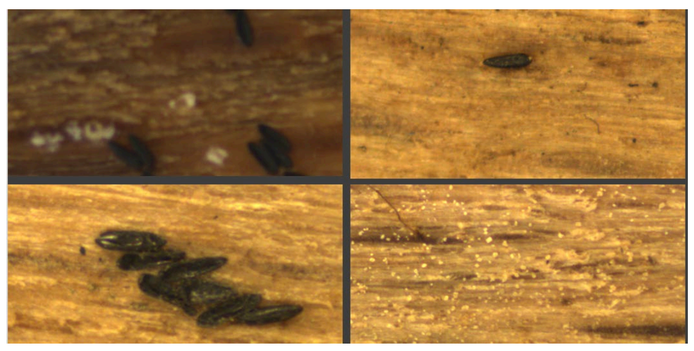
HSLU Hochschule Luzern

Technik & Architektur

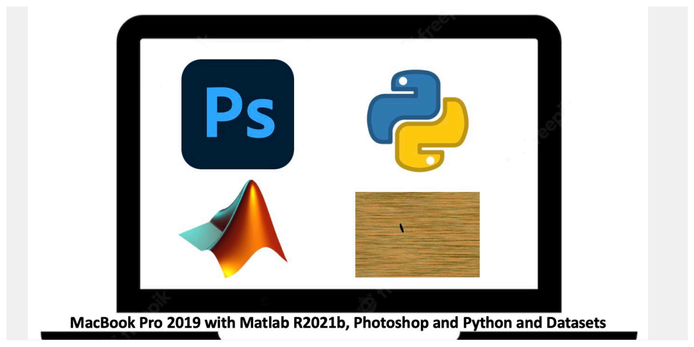
Master-Thesis Studiengang ICT

Identification of Tiger Mosquito Eggs using Machine Learning

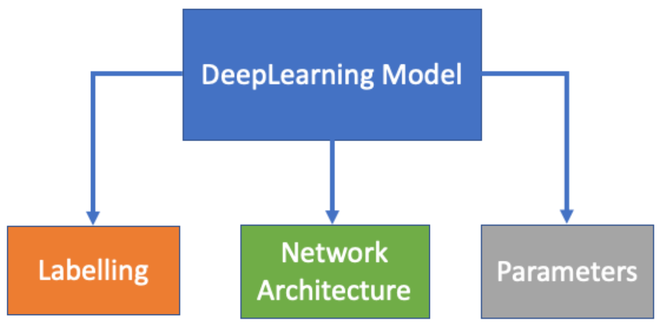
Overview of some sample wooden pieces and eggs



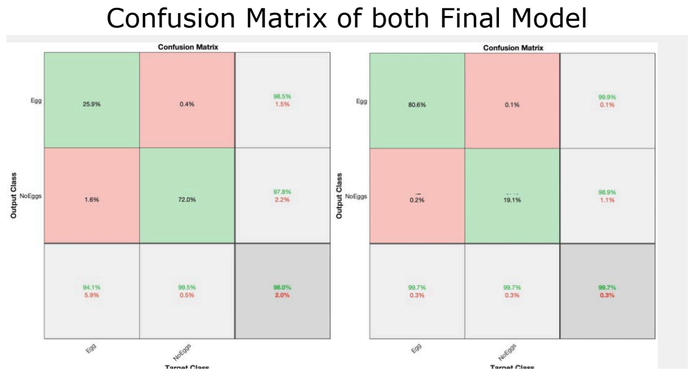
Overview of computational experiment setup



Experiments Categorization

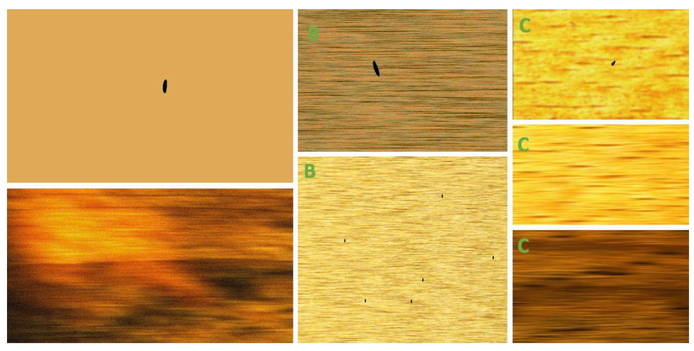


Confusion Matrix of both Final Model

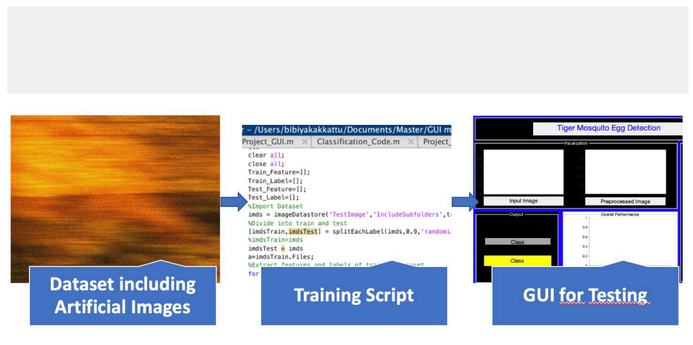


Left Model: only Real Data, Right Model: including Artificial Dataset

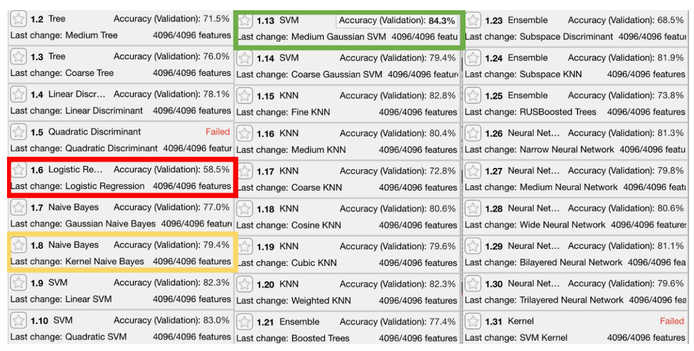
Overview of some sample artificial datasets



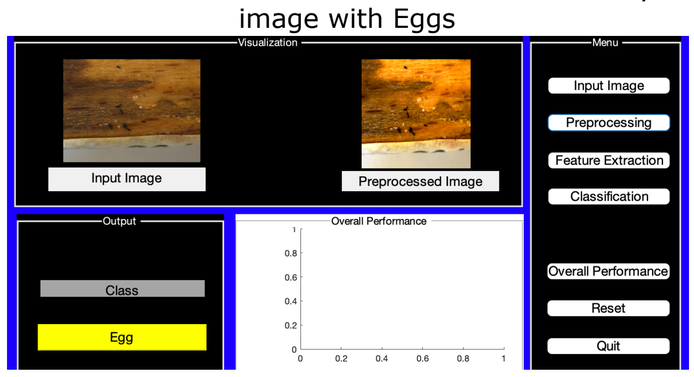
Overview of computational experiment setup



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

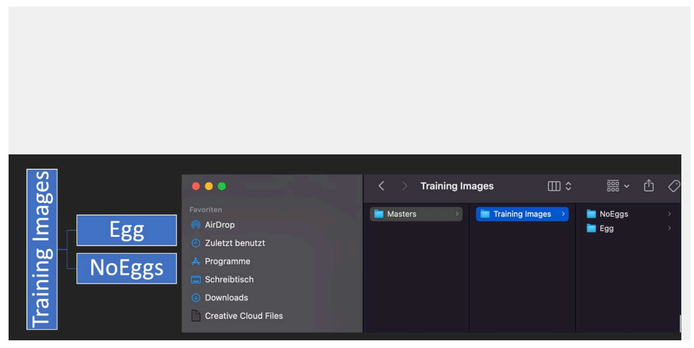


Correct Prediction of Second model for a noisy image with Eggs

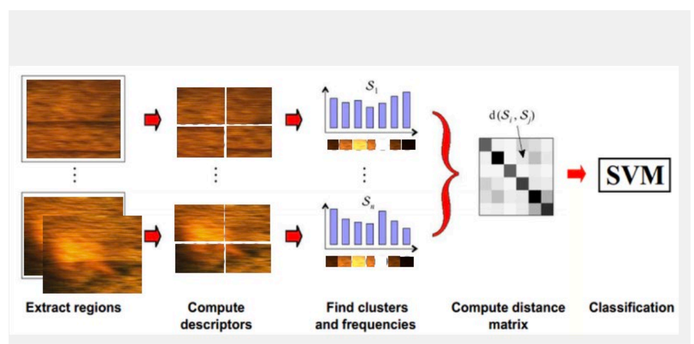


Artificial Tiger Mosquito Eggs Drawings from the Adobe Photoshop

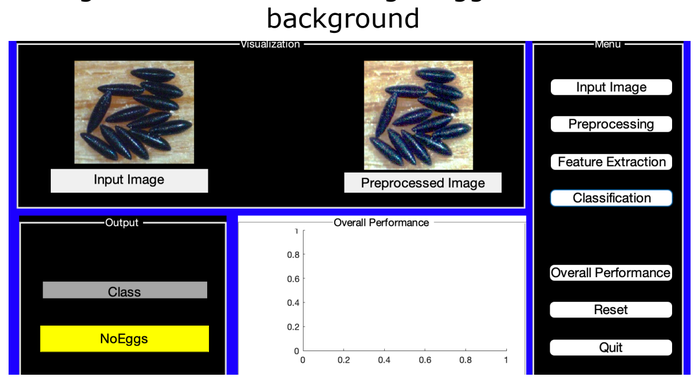
Hierarchy of the folders containing labelled training datasets



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



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.

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