NIVA Geotagged photos – Lithuanian AI Challenge

Jappe Franke, Arun Pratihast, Rob Knapen

Wageningen Environmental Research

November 2022

Introduction

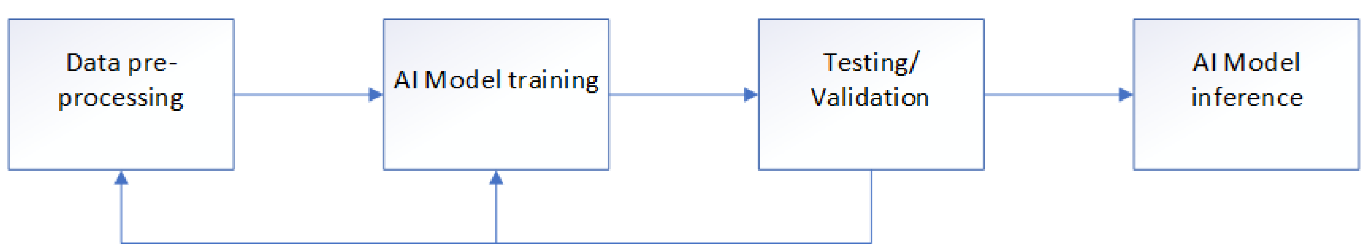
As part of the NIVA project, a platform for AI powered auto review and pre-validation of geo-tagged agriculture images for IACS (Integrated Administration and Control System) will be developed. This platform will be hosted at the Walton Institute on behalf of NIVA. The platform is split into two main aspects, a general facility for image uploading and tagging, and a facility for reviewing by a human or auto-reviewing by an AI model.

For the targeted initial three auto-review (AI) tasks, suitable auto-review challenges have been proposed by Paying Agencies. Including the Lithuanian challenge for which the model training and validation is described in this document.

In Lithuania currently farmers are required to capture and send geotagged photos with season's growth just before harvest for Coupled support for the fruits, berries and vegetables scheme. However, this is resulting in many images being sent in and needing manual inspection for confirmation of fruits, berries and vegetables to be present in the photos. A possible pre-selection by an AI model is helpful. The initial challenge for the AI model is straight-forward, a classification task of images with evidence of fruit, berry or vegetable bearing plants.

Workflow

The functional features of the workflow include: 1) Data pre-processing; 2) AI Model training; 3) Testing/Validation; and 4) AI Model inference for the proper operationalization. The individual modules are shown below:



Data pre-processing

After obtaining the dataset from Lithuania, the dataset of ~9700 images of fruit, berry or vegetable bearing plants (Compliant) and ~800 plants without (Non-compliant) was prepared for training a Convolutional Neural Network (CNN) deep learning AI model.

Manual image check

The dataset contained quite a few images that were not of plants;

Also, many images were recorded at a 90-degree angle and unfortunately those images didn't contain any [EXIF] information on their orientation;

During examination of the available dataset, it was noted that it only contained about 800 photos from the Non-compliant category. The clear cause of this is that the farmers do not take such photos and send them in for validation, so they simply are not plentiful.

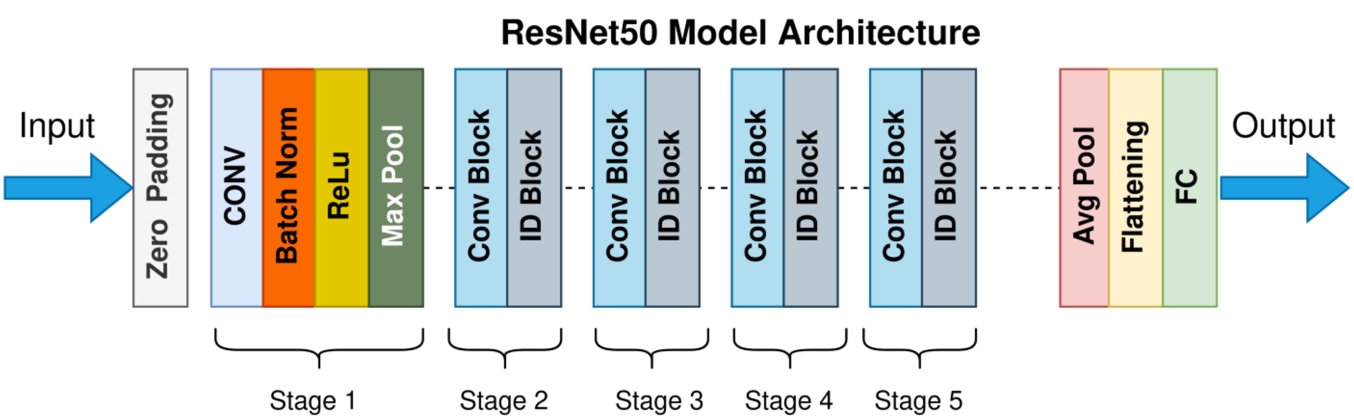
Some images in the Non-compliant non fruit/vegetable bearing plants still contained fruits/vegetables;

Automated image cropping

All images selected for use were automatically cropped to a center-bottom image of 500 x 500 pixels, so that some noise was removed at the borders of the landscape and portrait images in the dataset. The reasoning behind it was that this would force the model to focus on the (expected) region of the photo that typically contains most of the 'plant' area, where it could learn the most usable patterns.

Model selection

Considering that the challenge is a computer vision task, the well-known ResNet50 convolutional neural network (CNN) model was selected for this initial experiment. The ResNet architecture (2015) is based on a series of convolutional and max pooling steps, and the first model to introduce skip connections. In general, it has good performance on classification tasks with a limited number of trainable parameters. Instead of training from scratch it is also possible to load ImageNet (https://www.image-net.org) based pre-trained parameters, which gives the model some initial knowledge about features visible in real-world photos.



The PyTorch (https://pytorch.org) open source machine learning framework has been used to train the model on the available images, and for inference (predictions). The Albumentations (https://albumentations.ai) Python package has been used as well for data augmentation.

Training

A ResNet50 model was trained on a GeForce RTX 2080 GPU (Graphics Processing Unit) for 40 epochs (iterations).

Unfortunately, the Non-compliant collection was too similar to the Compliant collection; the Non-compliant images could not be recalled by the freshly trained model.

After consultation with the Lithuanian partner, it was decided that the dataset would be improved by only adding clear example images of the 2 categories. This improved image set contained 977 Compliant and 578 Non-compliant images.

The training took about 4 hours, with a training : validation : test split of 6 : 3 : 1

Image augmentation (by adding additional variation in orientation, hue, saturation, brightness) was done on all the images, but augmentation did not improve the classification and validation; an indication the dataset is still noisy.

After training the results were as follows;
The training dataset achieved an internal accuracy of 88.7%

Table 1 Model training confusion matrix

Training	**Compliant**	**Non-compliant**
Compliant	897	80
Non-compliant	95	483

Table 2 Model training performance metrics

Metrics	**Precision**	**Recall**	**F1-score**	**Support**
Compliant	0.904	0.918	0.911	977
Non-compliant	0.858	0.836	0.847	578
Total Accuracy			0.887	1555
Macro Average	0.881	0.877	0.879	1555
Weighted Average	0.887	0.887	0.887	1555

Below are a few example images; red labels are falsely predicted images

Validation with Non-compliant Irish clean pasture images

Validation against 872 Non-compliant Irish clean pasture images showed 81.4% accuracy. Using the Irish images was done because of the very low amount of Non-compliant images in the original dataset.

Table 3 Confusion matrix for model predictions on Compliant - Irish Non-compliant images

Validation	**Compliant**	**Irish - Non-compliant**
Compliant	872	105
Irish - Non-compliant	238	634

Table 4 Model prediction performance on Compliant - Irish Non-compliant images

Metrics	**Precision**	**Recall**	**F1-score**	**Support**
Compliant	0.786	0.893	0.836	977
Irish - Non-compliant	0.858	0.727	0.787	872
Total Accuracy			0.814	1849
Macro Average	0.822	0.810	0.811	1849
Weighted Average	0.820	0.814	0.813	1849

Testing on the whole original Lithuanian dataset

Testing against 830 Non-compliant and 9753 Compliant images from the original dataset showed 56.2% accuracy.

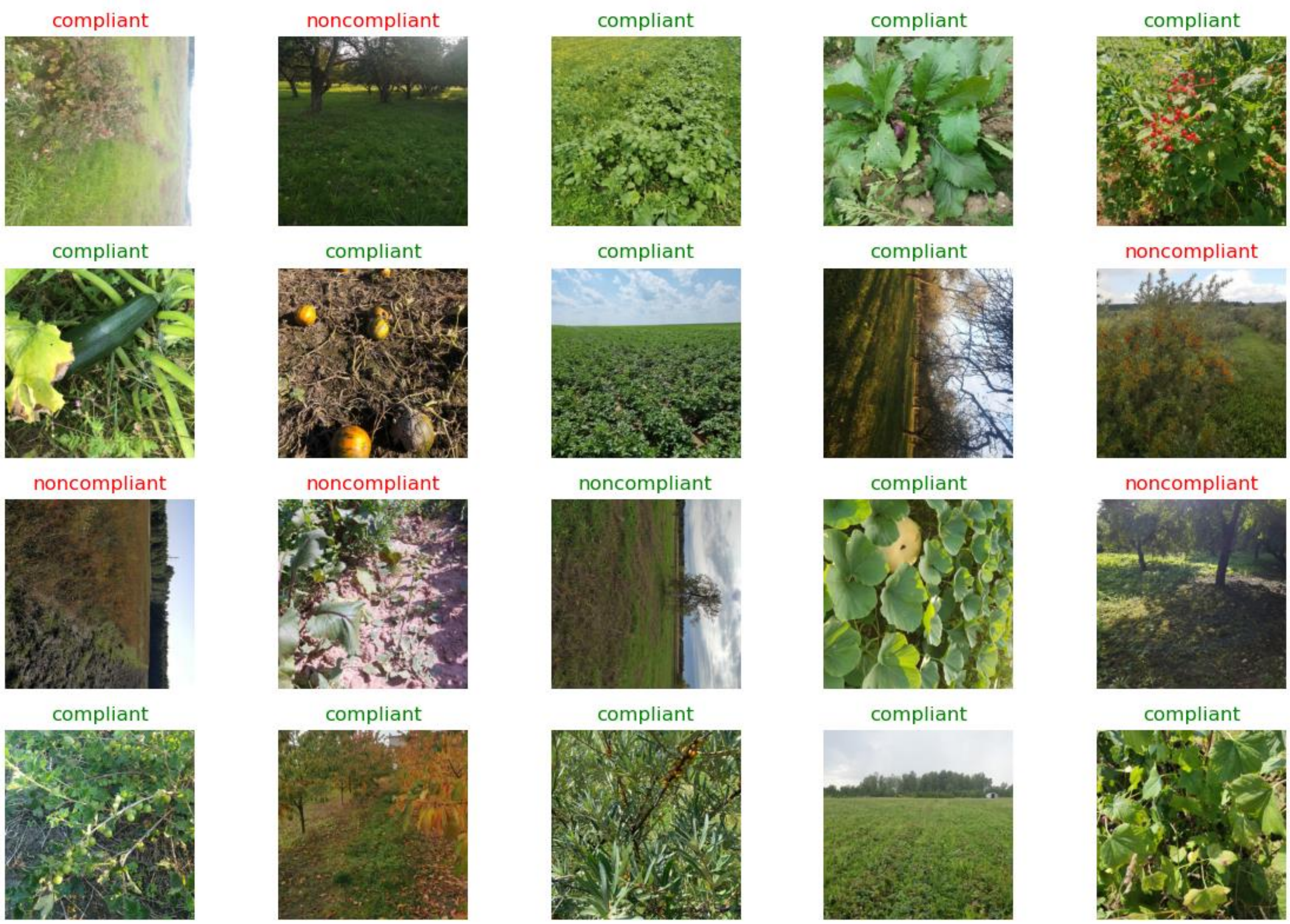
This shows the original dataset is quite confusing to the ai model trained on the improved dataset. It could be interesting to see if the model can correctly detect wrongly labeled images as a follow-up experiment.

Table 5 Confusion matrix for model predictions on Original Compliant - Non-compliant images

Validation	**Original Compliant**	**Original Non-compliant**
Original Compliant	5536	4217
Original Non-compliant	414	416

Table 6 Model prediction performance on Original Compliant - Non-compliant images

Metrics	**Precision**	**Recall**	**F1-score**	**Support**
Original Compliant	0.930	0.568	0.705	9753
Original Non-compliant	0.090	0.501	0.152	830
Total Accuracy			0.562	10583
Macro Average	0.510	0.534	0.429	10583
Weighted Average	0. 864	0.562	0.662	10583



Conclusions and lessons learned

The AI model was developed and trained on 1555 field images from the Lithuanian Challenge, the classification of Compliant and Non-compliant fruit/berry/vegetable plants was achieved. The results show that the AI model (using a ResNet 50 CNN (Convolutional Neural Network) architecture) trained reasonably well on the available Lithuanian Compliant and Non-compliant plant images with a training accuracy of 88.7%, weighted F1 0.887, and a validation accuracy of 81.4%, weighted F1 0.813.

However, this model's performance drops significantly when used to predict the original images (accuracy 56.2%, weighted F1 0.662). This shows that the original dataset contains many pictures that still prove to confuse the AI model. The original dataset contained quite a few images of fruits and berries in the Non-compliant section.

The expectation is that these results can be improved substantially by a more tailored training dataset. E.g., the model must be told what aspects in the image make sense to learn about, so that it will be able to make the expected distinction between Compliant and Non-compliant situations. Currently it looks at a lot of noisy / confusing photos. Pre-processing images were already zoomed in on a 500x500 pixel center-bottom crop of the image, but that does not help enough.

After consultation with the plant vision expert colleagues at the Agro Food Robotics group, they also advised basically the same. It is their expectation that annotating the areas of interests (on what causes compliant vs non-compliant classification) will improve the model accuracy quite significantly. Annotation, however, is very time consuming and perhaps a good idea for follow-up.

Looking at the results overall; more clear and higher quality images are needed for both categories. Currently the visual characteristics for Compliant images and Non-compliant images are still very noisy.