2018 International Joint Conference on Agricultural and Biosystems Engineering(JCABE 2018)

Online service for detecting plant diseases

A. Nechaevskiy, G. Ososkov, P. Goncharov, A. Uzhinskiy
nechav@jinr.ru
Laboratory of Information Technologies, Joint Institute for Nuclear Research, Dubna, Russia
Sukhoi State Technical University of Gomel, Gomel, Belarus

Problem

Crop losses are a major threat to the wellbeing of rural families, to the economy and governments, and to food security worldwide. Quality of available data about the impact of plant diseases is variable, patchy and often missing, particularly for smallholders, who produce the majority of the world's food.

- CIP, the international research center with an historical mandate for potato, estimates **15% production losses each year due to late blight**
- USAblight (a national project on late blight on potato and tomato) says that (annual) **global losses 'exceed US$6.7 billion'.**

As agriculture struggles to support the rapidly growing global population, plant disease reduces the production and quality of food, fibre and biofuel crops. Losses may be catastrophic or chronic, but on average account for **32% of the production of the six most important food crops**.

Prof. Dr. David Guest. Faculty of Agriculture, Food and Natural Resources, The University of Sydney

Globally, about 16% of all crops are lost to plant diseases each year.

Dr. Caitilyn Allen Department of Plant Pathology, University of Wisconsin–Madison

Our goal

Increasing number of smartphones and advances in deep learning field opens new opportunities in the crop diseases detection.

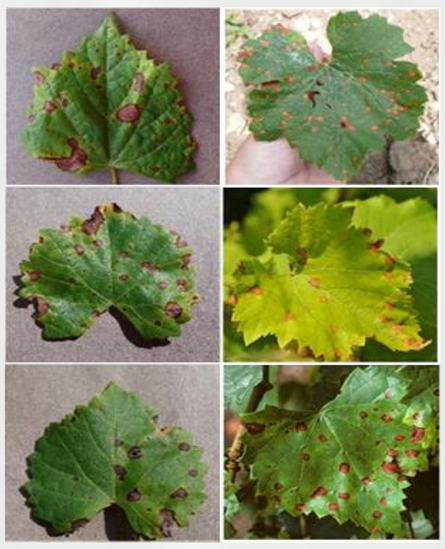
The idea is:

- **to create the plant disease detection platform (PDDP)** that will use modern organization and deep learning technologies **to provide new level of service to farmer's community;**
- provide **open access** to our image database;
- **share the code** of the models via Github;
- as end-product we are going **to develop the mobile application** allowing users to **send photos and text description of sick plants and get the cause of the illness.**

Data is a Key

Attentively look on this picture. First row – PlantVillage images, second row – real-life images from the Internet.

Do you see anything strange? **HINT**:
background



We decided to create our own dataset!
PDD dataset (http://pdd.jinr.ru/db/):

- started from **grapes images**;
- **4 diseases – Esca, Black rot, Chlorosis, Mildew**;
- **133 healthy images;**
- in total – **313 samples** with the size of 256x256;
- we are going add more crops to the dataset.

Results

Our Siamese network unites the twins within L1 distance layer followed by sigmoid activation in order to train the net with cross-entropy objective. Detailed information can be found at [*Disease detection on the plant leaves by deep learning. Goncharov P. et al.*]

In the table one can see the confusion matrix of the K-nearest neighbors algorithm on the test subset of embeddings data.

	Black rot	Chlorosis	Esca	Healthy
Black rot	7	0	1	1
Chlorosis	0	11	0	1
Esca	0	0	20	1
Healthy	0	0	0	29

T-SNE visualization of the high-level features extracted by the siamese twin

It is simple to deduce that the classification accuracy equals to **94.3%**.

We extract two components to plot them in 2D space. One can see that there are four separate clusters – one per each class.

Although, there are a few points, which wrongly got into the different set (see the table), but it is not detractive.

After we have finished with architecture we can keep on going with other parts of the platform.

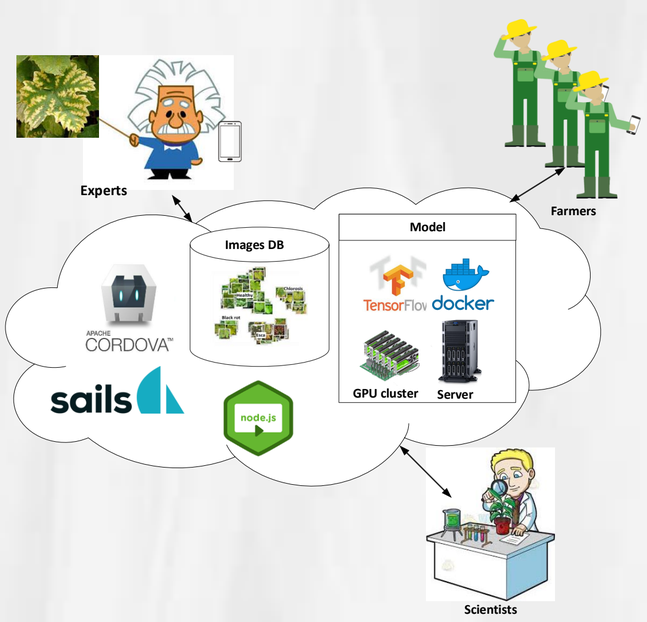
PDDP Architecture

PDDP consists of a set of interconnected services and tools developed, deployed and hosted in the cloud infrastructure.

Our web-portal pdd.jinr.ru is developed with Node.js (Sail.js). It provide not only web-interface but also an API for third-party services.

We have TensorFlow model in Docker realized as a service. The model can work at the virtual server or at GPU cluster.

Right now we are storing images directly on the local drive but if their number will increase dramatically we will use cloud storage.



PDDP basic principals

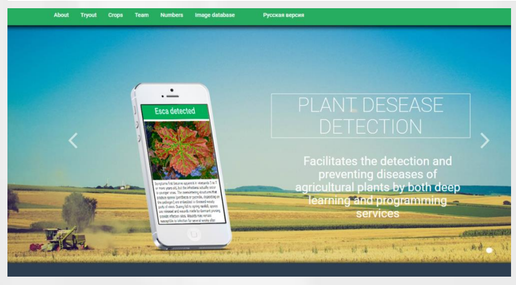
Users can:
- send photos and text description of sick plants through web-interface or mobile application and get the cause of the illness,
- browse through diseases description and galleries of ill plants,
- verify that requested disease was recognized right and treatment helps.

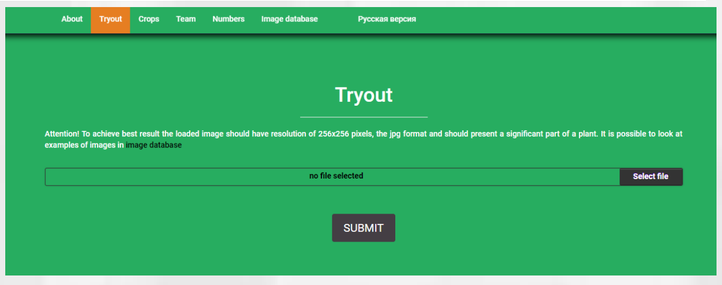
Experts can:
- browse user requests and verify recognition,
- request addition of their image or image from the user complain to the DB,
- request alter of the diseases description,
- request retraining of the model with new images.

Researchers can:
- work with images data base through web-interface or API,
- download all or only part of the base,
- obtain and API-key to submit recognition tasks to the platform.

Supervisors can:
- add new images to the data base,
- initiate retraining of the model,
- get different statistical metrics about portal users.

What we have now





Site with general information (**pdd.jinr.ru**)

Model in Docker running at the virtual server

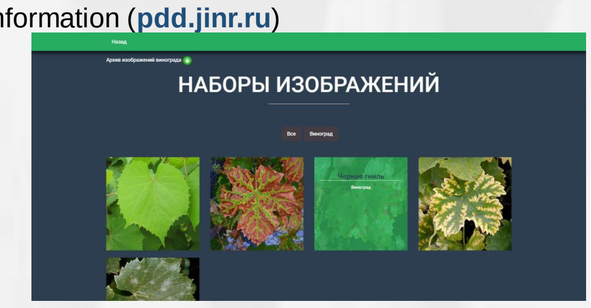
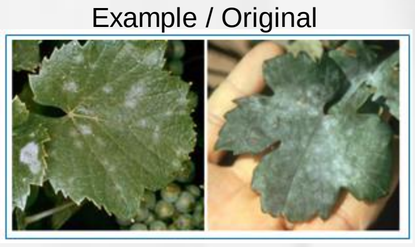


Image database

Example / Original



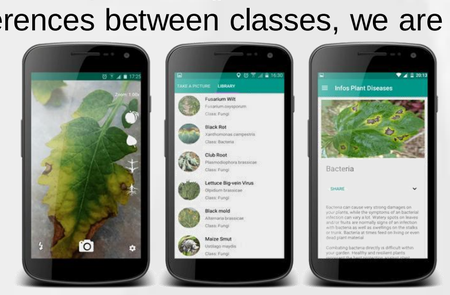
Disease detection results

Conclusion

- It is not enough to have lots of images to recognize diseases. Quality of the images database is extremely important for the results of detection.
- Siamese neural networks are very perspective research field for plant disease detection projects.
- It is clear that unambiguous detection of the diseases is unsolvable task especially at first stages of a plant illness.

We are going to use siamese neural networks as a basic deep learning architecture for PDDP and since their power consists in seeking for differences between classes, we are going to add more classes to the train dataset soon.

We continue to develop web-portal. Section for users, experts and researchers will be open soon.
We are going to present draft mobile application in second half of 2019.



Mobile Application Example

The reported study was funded by RFBR according to the research project № 18-07-00829

Disease detection on the plant leaves by deep learning. Goncharov P. et al. // Studies in Computational Intelligence, Volume 799, 2019, Pages 151-159