MULTIFUNCTIONAL PLATFORM FOR PLANT DISEASE DETECTION

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

INTRODUCTION

Crop losses are a major threat to the wellbeing of rural families, to the economy and governments, and to food security worldwide. CIP [Centro Internacional de la Papa (CIP)] , the international research center with a 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'.
Globally, about 16% of all crops are lost to plant diseases each year.
Dr. Caitilyn Allen Department of Plant Pathology, University of Wisconsin–Madison

The aim of our research is to facilitate the detection and prevention of diseases of agricultural plants by both deep learning and programming services. The idea is to develop a multifunctional platform, which will use modern organization and deep learning technologies to provide a new level of service to farmers' community.
As an end product we are going to develop a mobile application allowing users to send photos and a text description of sick plants and get the cause of the illness.

IMAGE DATABASE

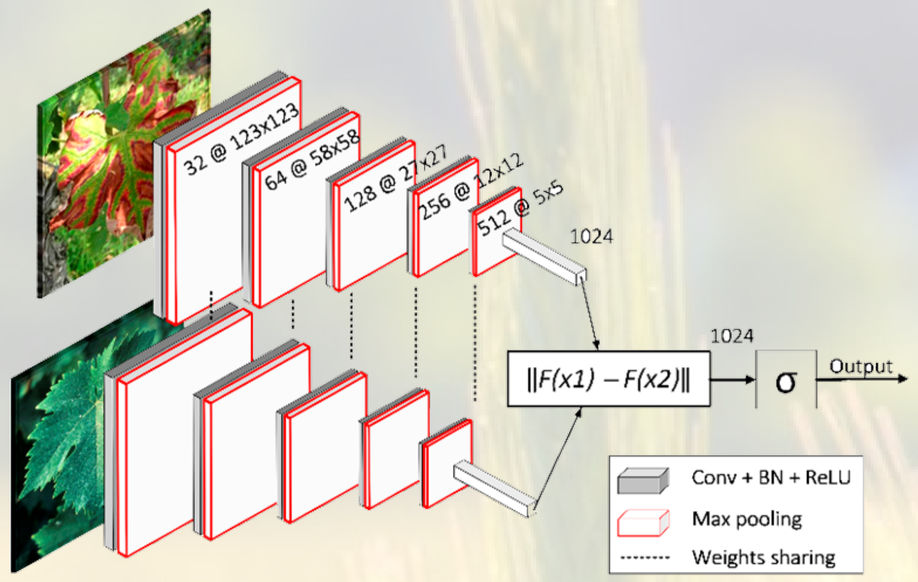
To train a deep learning model, a great amount of data is required.
Unfortunately, there is not a well-suitable database for our purposes. The existing image datasets are rather synthetic.
We have to create our own database.
10 classes: 5 for Grape and 5 for Wheat
~ 460 images

SIAMESE NETWORKS FOR DISEASE DETECTION

We considered different models, which were used in the related works, to understand what would be the best option. The results obtained on real life images were very poor – 48% accuracy. We collected our own database. But, how could we get good features from a very small amount of data? We addressed this problem to the so-called one-shot approach offering a solution by siamese neural networks.

The Siamese network consists of twin networks joined by the similarity layer with the energy function at the top. Weights of twins are tied, thus the result is invariant and in addition guarantees that very similar images cannot be in different locations in the feature space, because each network computes the same function. The similarity layer determines some distance metric between so-called embeddings, i.e. high-level feature representations of the input pair of images. Training on pairs means that there are quadratically many possible samples to train the model on, thus making it hard to overfit.

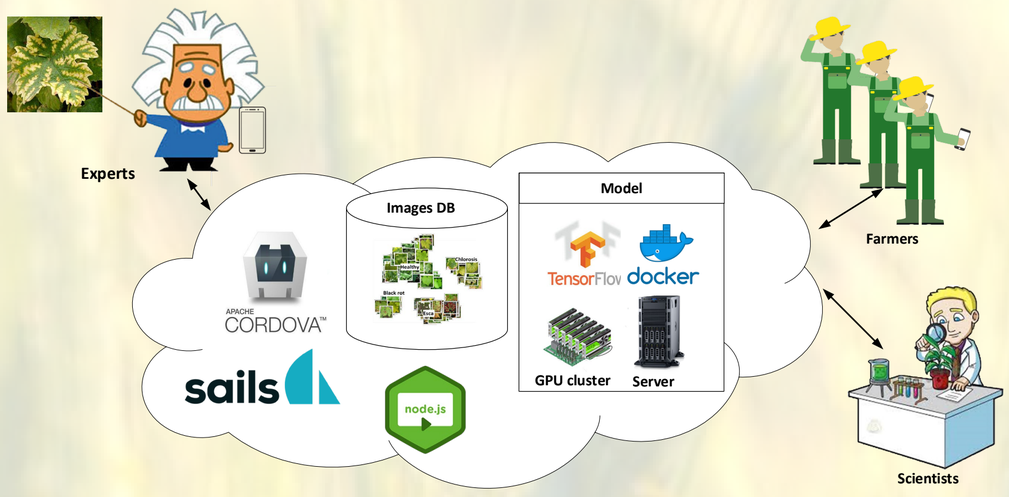
Our Siamese network unites the twins within the L1 distance layer followed by the sigmoid activation in order to train the net with a cross-entropy objective. Detailed information can be found at [Disease detection on the plant leaves by deep learning, Springer].



Our best siamese convolutional architecture. «Conv» means a convolutional operation, «BN» is a batch normalization, «32 @ 123x123» – 32 filters with a particular size

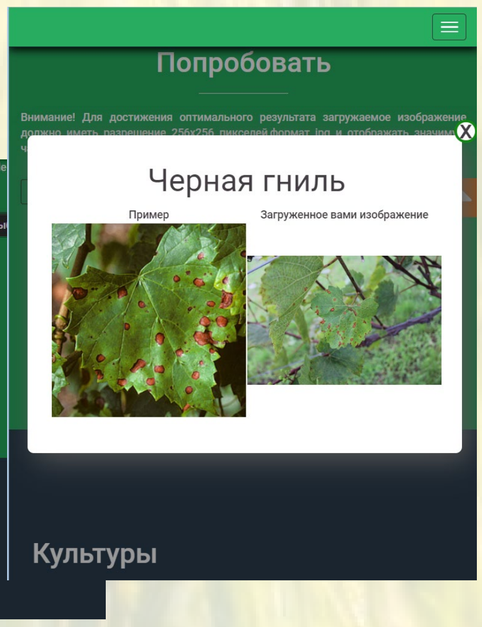
The obtained accuracy with a value of 93% proves that this architecture can be used for our task .

PDDP ARCHITECTURE



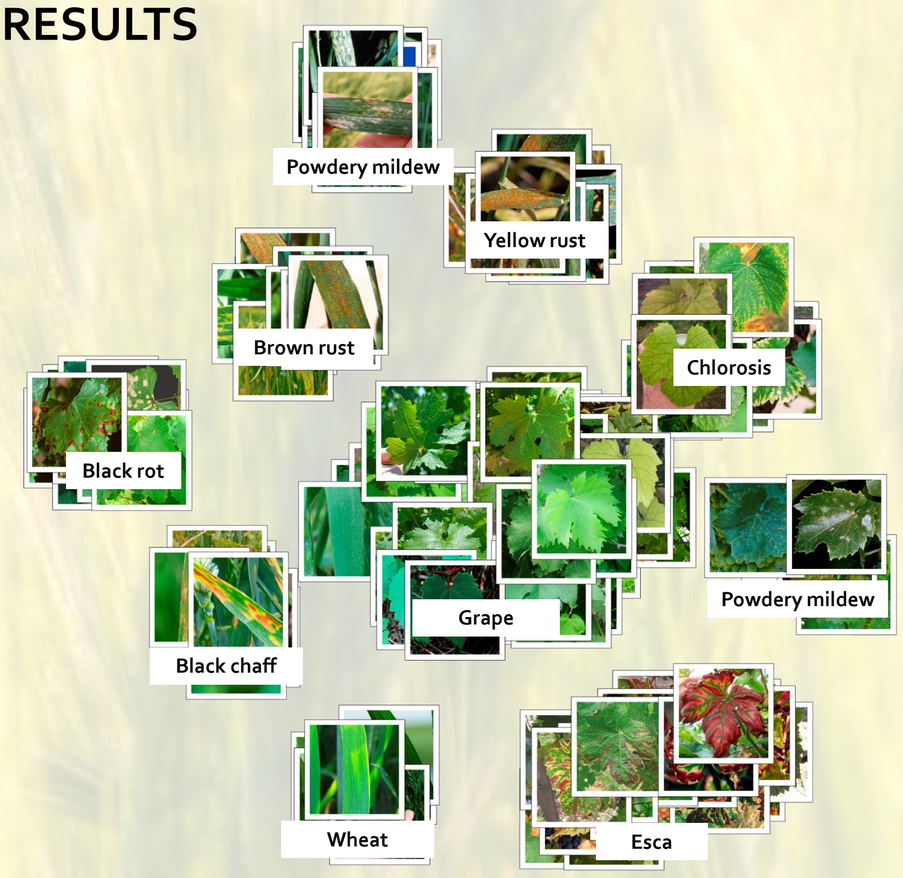
PDDP consists of a set of interconnected services and tools developed, deployed and hosted with the help of the JINR cloud infrastructure. Our web portal (pdd.jinr.ru) was developed with the Node.js (Sail.js). It provides not only a web interface, but also API for third-party services. We have the TensorFlow model in the Docker realized as a service. The model can work at a virtual server or at a GPU cluster. Right now, we are storing images directly on the local drive, but if their number increases dramatically, we will use cloud storage like disk.jinr.ru. We will use Apache Cordova to create a Mobile App for Android, IOS and Windows platforms.

Currently, we have a web site with general information - pdd.jinr.ru, the open image database and the model running in the Docker container at the virtual server. One has the ability to submit disease detection jobs at the website and get results.



Example of the portal work

RESULTS



T-SNE visualization of the high-level features extracted by the siamese twin in 2D space

CONCLUSION AND PLANS

- It is not enough to have many images to recognize diseases. The quality of the image database is extremely important for the results of detection.
- The Siamese neural network is a very perspective research field for plant disease detection projects.
- It is clear that the unambiguous detection of diseases is an unsolvable task, especially at first stages of the plant illness.

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