

# One-shot learning with triplet loss for vegetation classification tasks

A.V. Uzhinskiy<sup>1,2</sup>, G.A. Ososkov<sup>1</sup>, P.V. Goncharov<sup>1</sup>, A.V. Nechaevskiy<sup>1,2</sup>, A.A. Smetanin<sup>3</sup>

<sup>1</sup>Joint Institute for Nuclear Research, 141980, Russia, Dubna, Joliot-Curie 6,

<sup>2</sup>Russian State Agrarian University - Moscow Timiryazev Agricultural Academy,  
Russia, Moscow, Timiryazevskaya st., 49,

<sup>3</sup>National Research University ITMO, 197101, Russia, Saint-Petersburg, Kronverkskiy pr., 49

## Abstract

Triplet loss function is one of the options that can significantly improve the accuracy of the One-shot Learning tasks. Starting from 2015, many projects use Siamese networks and this kind of loss for face recognition and object classification. In our research, we focused on two tasks related to vegetation. The first one is plant disease detection on 25 classes of five crops (grape, cotton, wheat, cucumbers, and corn). This task is motivated because harvest losses due to diseases is a serious problem for both large farming structures and rural families. The second task is the identification of moss species (5 classes). Mosses are natural bioaccumulators of pollutants; therefore, they are used in environmental monitoring programs. The identification of moss species is an important step in the sample preprocessing. In both tasks, we used self-collected image databases. We tried several deep learning architectures and approaches. Our Siamese network architecture with a triplet loss function and MobileNetV2 as a base network showed the most impressive results in both above-mentioned tasks. The average accuracy for plant disease detection amounted to over 97.8% and 97.6% for moss species classification.

**Keywords:** deep neural networks; siamese networks; triplet loss; plant diseases detection; moss species classification.

**Citation:** Uzhinskiy AV, Ososkov GA, Goncharov PV, Nechaevskiy AV, Smetanin AA. One-shot learning with triplet loss for vegetation classification tasks. *Computer Optics* 2021; 45(4): 608-614. DOI: 10.18287/2412-6179-CO-856.

**Acknowledgements:** A.V.U. and A.V.N. gratefully acknowledge financial support from the Ministry of Science and Higher Education of the Russian Federation in accordance with agreement № 075-15-2020-905 dated November 16, 2020 on providing a grant in the form of subsidies from the Federal budget of Russian Federation. The grant was provided for state support for the creation and development of a World-class Scientific Center “Agrotechnologies for the Future”. The database creation part of the reported study was funded by RFBR according to the research project No 18-07-00829.

## Introduction

Deep neural networks (DNN) have started a revolution in object detection and recognition. Since 2012, many DNN architectures have been conceived and applied to different tasks including face and object recognition, plant and animal classification, image reconstruction and synthesis, etc. The transfer learning (TL) approach, within which a well-designed DNN model developed for a recognition task is reused as a feature extractor for a new model, is a common way of solving classification tasks. To get good results in a certain domain, researchers must choose an appropriate base network, fine-tune the model, select an optimizer, and a loss function. In our study, we dealt with two vegetation classification tasks. We have created the platform for plant diseases detection (PDD, pdd.jinr.ru) [1] in 2017. Crop losses is always a serious problem that costs billion dollars a year to the farmers' community. When we just started our project, there was only one real-life application allowing users to send photos and text descriptions of diseases plants and get a diagnosis and treatment recommendations (Plantix). To obtain the most efficient result we tried various deep

learning (DL) architectures and faced the problem of the poor image database [2]. Therefore, we had to collect our own database consisting of 15 classes of three crops, 611 images in total, and to use Siamese networks with cross-entropy loss that allowed us to get 95.7% accuracy [3]. It was not the worthy results, so we planned to extend the database and to look for ways to improve accuracy. Applying the triplet loss function, which showed good results in object classification [4] was one of the options. One more of our projects is focused on the development of the data management system for the ICP Vegetation Program [5]. This program is carried in the framework of the UNECE Convention on Long-Range Transboundary Air Pollution (LRTAP). The analysis is based on the processing of bio-accumulators (naturally growing mosses) collected in 39 countries of Europe and Asia every 5 years. Contributors responsible for sample collection may not always have the necessary competence to identify the moss species; however, this information is important for the analysis phase. We tried to use the approach of our PDD project to identify moss species. There are 599 images of five moss species in the database and we managed to get only 82% accuracy using the Siamese network with

cross-entropy loss and less than 60% with TL based on ResNet50. This is not a very impressive result, probably because the sample used did not allow identifying sufficiently significant properties of mosses, but even an expert finds it difficult to distinguish one type of moss from another. Our research is aimed to test whether our approach with the triplet loss function could improve the situation.

## 1. Materials and methods

### 1.1. Related works

There are many studies in which DL models are used to detect and classify plant disease symptoms. We can divide them into two types depending on the database used. The first one uses the PlantVillage database, i.e. – 14 crops and 26 diseases, a total of 54,306 images of diseased and healthy plant leaves collected under controlled conditions [6]. The second one collects a dataset in the field or on the internet. Using the PlantVillage dataset with the transfer learning approach allows one to get more than 95% accuracy on a test subset, but the accuracy decreases dramatically on field images. For example, Mohanty et al. [7] use AlexNet and GoogLeNet on the PlantVillage database to classify 26 diseases and obtain 99.35% accuracy on a test subset. However, poor accuracy of 31.4% is achieved for real-life images. Too et al. [8] use VGG16, Inception V4, ResNet with 50, 101, and 152 layers, and DenseNets on the PlantVillage database. They try a variety of optimization techniques and training approaches to obtain more than 98% accuracy for all TL models except VGG 16, for which the accuracy is over 82%. The authors have not reported tests on real-life images; however, we can assume that the accuracy could be much worse. Arguments for this assumption can be found in [9], where the author uses AlexNet, AlexNetOWTbn, GoogLeNet, Overfeat, VGG architectures on the second edition of the PlantVillage database with 87,848 images of 58 different classes. The reported validation accuracy exceeds 97%. The authors also test the model on real-life images and present an experiment with real-life images added to the training dataset. The accuracy of the model trained on the PlantVillage database on the field test set is over 33% as in [7]. Accuracy of the model trained on the mixed dataset was over 65% on the field dataset. Using the PlantVillage database to train the model for a real-life application appointed as a bad idea. Thus, studies with self-collected databases seem to us much more interesting. Fuentes et al. [10] use a self-collected tomato database of 10 classes, 5000 images in total. The authors consider three types of detectors: Faster Region-based Convolutional Neural Network (Faster R-CNN), Region-based Fully Convolutional Network (R-FCN), and Single Shot Multibox Detector (SSD) combined with a feature extractors such as VGG and ResNet along with the SmoothL1 loss function. The Faster R-CNN with VGG16 as a feature extractor gives an overall average accuracy of 83% for all classes. Türkoğlu and Hanbay [11] use various base networks such as AlexNet, VGG16, VGG19,

SqueezeNet, GoogleNet, Inceptionv3, InceptionResNetv2, ResNet50, ResNet101, and classifiers such as K-nearest neighbor (KNN), Support vector machine (SVM), Extreme learning machine (ELM). Their self-collected database comprises 1965 high-resolution images of eight different plant diseases of four crops. The highest-level accuracy, which is obtained with the ResNet50 model and SVM classifier, amounts to 97.86%. The results are very impressive. Unfortunately, there is no link to their database. Selvaraj et al. [12] focus on banana diseases. They have collected a database with 30,952 images of different parts and entire plants in which 9000 images are photos of leaves. The authors use ResNet50, InceptionV2, and MobileNetV1 architectures. The best result on leaves is over 70% accuracy with ResNet50 and InceptionV2. In general, the accuracy of models trained on field data is lower than that of models trained on PlantVillage data. Saleem et al. [13] have done a comprehensive review of works related to plant diseases and visualization techniques. It should be noted that there is another class of plant disease studies in which images are taken from aerial vehicles or satellite sensors. It is a highly interesting and promising direction; nevertheless, it is beyond the scope of our current research.

Regarding the classification of the moss species, we managed to find only one related work using the DL approach [14]. The training data is prepared by the “chopped picture” method when one high-resolution picture is “chopped” into a number of small pictures. The resulting dataset consists of 93,841 images of three moss species. The authors use the LeNet network with SGD solver on NVIDIA DIGITS 4.0. The accuracy for each moss species is 99%, 95%, and 74%. The authors suppose that some moss species are relatively large and have a relatively distinctive well-defined shape, while others are highly amorphous.

We started our work with plant disease detection architectures in 2017. Firstly, we reproduced the TL approach on the PlantVillage database considering only grape diseases. We tried four base networks such as VGG19, InceptionV3, ResNet50, and Xception. ResNet50 showed the best result, i.e. – 99.4%. The accuracy on the test subset of 30 images from the Internet was about 48%. We unsuccessfully tried different ways of data modification described in [2]. We realized that the PlantVillage database was not an option in case of the real-life application. PlantVillage images have the same background, illumination, position, and orientation. In the real-life, images taken in the different conditions under various angles and at the background can be anything. Thus, we began collecting our own database. We started from four classes of grape diseases, a total of 279 images including healthy leaves. We faced a problem of training models on a small amount of data. The solution was found with the help of a One-shot learning approach, namely Siamese network. Being already trained one of the Siamese network twins acts as a feature extractor followed by the KNN-classifier. With this

approach, we achieved the accuracy of 94%. The architecture of the proposed Siamese network presented in fig. 1. After the extension of the database to 15 classes of three crops, 611 images in total, we obtain a lower accuracy of 86% while preserving the same classification scheme. We decided that the problem was in the classifier and tried different options.

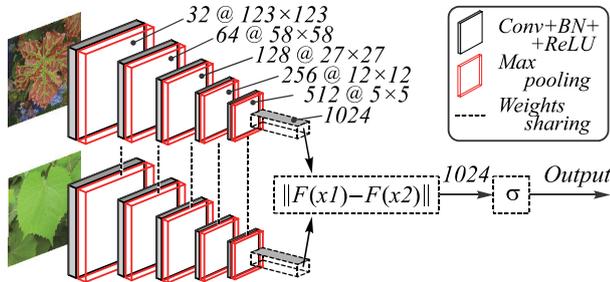


Fig. 1. The architecture of the Siamese network

With the Siamese network as a feature extractor and multi-layer perceptron (MLP) as a classifier, we managed to get 95.7% accuracy [3]. The moss classification task appeared in 2019. At that time, we already had some practices. Thus, we started from the TL approach with ResNet50. The accuracy was near 59%. The Siamese network with MLP gave only 78% accuracy. We made some changes to the convolutional layers but could not obtain more than 81% accuracy. We believe that Siamese networks is a very promising direction, and we hoped that we could improve accuracy using the triplet loss function.

1.2. The triplet loss function

In [3] we used Siamese networks with cross-entropy loss to get 95.7% accuracy on 15 classes. After increasing the database to 25 classes, the accuracy decreased to 89.6%. We wanted to check if the triplet loss could improve accuracy.

Siamese networks and the triplet loss function can be used when there is a paucity of training data available. This combination has shown impressive results in facial recognition tasks [4, 15, 16], object tracking [17], brain imaging modality recognition [18], bioacoustics classification [19], remote sensing scene classification [20] and other tasks.

The Siamese network consists of a twin network with tied weights joined by the similarity layer with the energy

function at the top. When we pass an image to the network input, we extract some features of the image in the output, so-called embeddings. Similar images cannot be in very different locations of the feature space, because each of the twins computes the same function due to weights sharing. The triplet loss function use three images during evaluations. The anchor is an arbitrary data point. The positive image belongs to the same class as the anchor. The negative image belongs to a different class from the anchor. The triple loss reduces the distance between the anchor and the positive image while increasing the distance between the anchor and the negative image (see fig. 2).

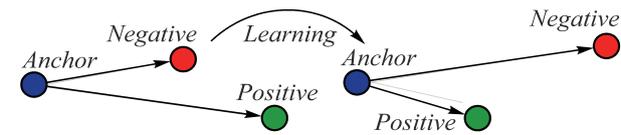


Fig. 2. Visualization of the learning process with the triplet loss function

When training the Siamese network with the triplet loss function, the input consists of three images, two of which belong to the same class, and the last one belongs to a different class. The model processes each image and produce a feature vector. In the end, we can say that the two images are from different classes if the distance between them is high and that they are from the same class if the distance is small.

1.3. Image database

Our image dataset for plant disease detection was collected from open sources on the Internet. It consists of 25 classes of four crops (cotton, wheat, corn, cucumbers, and grape), 935 images in total (see fig. 3). All images are 256x256 pixels and contain meaningful information about crops and diseases on them. Comparing with our old publication [3], we have added two new crops, i.e. – cotton and cucumbers with ten classes (Alternaria leaf blight, Healthy, Powdery mildew, Verticillium wilt, Nutrient deficiency for cotton and Anthracnose, Downy mildew, Healthy, Nutrient deficiency, Powdery mildew for cucumbers). The database and the link for downloading can be found at <http://pdd.jinr.ru/>.

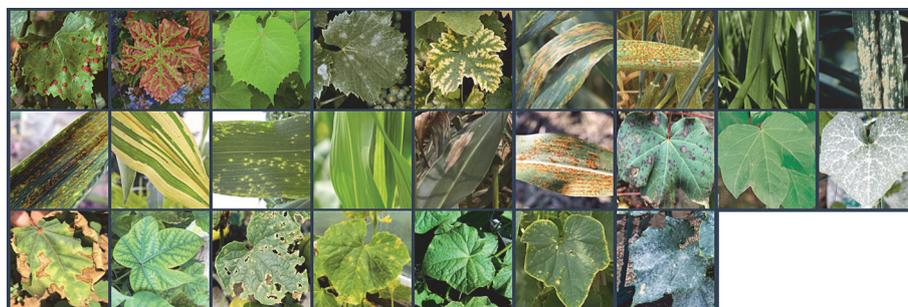


Fig. 3. Images of 20 classes of diseases from the PDD database

Our moss species dataset consists of 599 images of five species (AbietinellaAbietina, HylocomiumSplendens, HypnumCupressiforme, PleuroziumSchreberi, and

PseudoscleropodiumPurum), see fig.4. All images are 256x256 pixels. The database and the link for downloading can be found at <http://moss.jinr.ru/>.

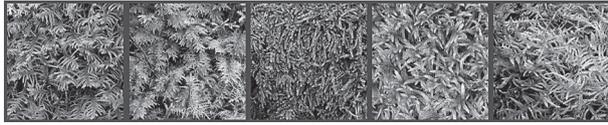


Fig. 4. Images of five moss species

1.4. Current solution

In our current architecture, we have a Siamese network with three twins (see fig. 5).

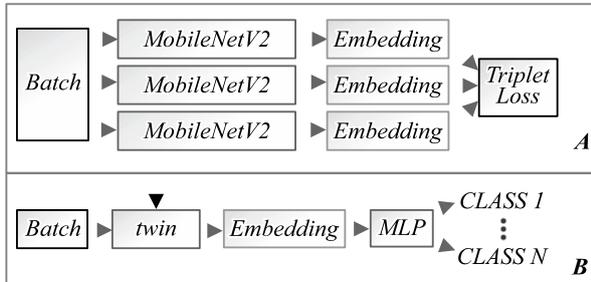


Fig. 5. Current architecture: (a) siamese network with three twins and the triplet loss function; (b) one of the twins used as a feature extractor and MLP used as a classifier

We use MobileNetV2 pre-trained on the ImageNet dataset as a base network for twins. The Siamese network is trained with the help of triplet loss using the Triplet-Torch utility. After training, one of the twins is used as a feature extractor for a multi-layer perceptron, which acts as a classifier. The dimension of the feature vector extracted from the embedding model is 1280.

The model was trained on the “HybriLIT” heterogeneous platform of the Joint Institute for Nuclear Research. The platform has various computing architectures, and we used NVIDIA Tesla K40. We trained the Siamese network for 300 epochs with a batch size 32 and a learning rate of 0.001. We used an 80 to 20 ratio to split the data into train and test subset. It took about 40 minutes to train the model. The Adam optimizer with learning rate equals to 0.001 and cross-entropy loss was used to train the perceptron. The classifier model was trained for 40 epochs, and it took a couple of minutes.

At present, each model weights over 18 Mb. We want to have a model that can be run “on board” of mobile devices; therefore, we have tried dynamic and post training quantization (PTQ). The best result was achieved by using PTQ. To perform PTQ, we have trained the feature extractor, then we fused all Conv-BN-ReLU blocks and inverted residual blocks of the original MobileNetV2 model. After the fusion of the model an additional step of feeding batches of data through the network was done in order to compute the resulting distributions of the different activations. This information is used to determine how specifically the different activations should be quantized at inference time. The last step is actually quantization. Then the quantized feature extractor model is used to train the classifier. We manage to reduce the model size to over 7 Mb and increase the performance by 5 times with static quantization. Without quantization con-

secutive processing of 100 images take approximately 14.2 seconds. The quantized model consecutively processes 100 images in less than 2.6 seconds.

2. Results and discussion

We trained the model for plant disease detection 16 times each time with a new split into train and validation. The average accuracy was 97.8%, while the highest accuracy was 99.4%. The moss classification model in the same condition showed 97.65% average accuracy and 100% highest accuracy.

With TL based on ResNet50, we could not obtain accuracy more than 60% for both tasks. The accuracy of siamese networks with cross-entropy loss on plant disease detection task was over 89% and less than 81% for the moss classification task. With the triplet loss function, we can get the state-of-the-art results for both tasks. Fig. 6 shows the representation of the extracted feature vectors in 2D space for the moss spices model before training and after 300 epochs of training. For the plant disease detection model, the situation is similar but many classes are blocked off; however, we can see a clear separation on the others.

During the evaluation of [1] we collected a test dataset of images to compare our model with the well-known commercial AutoML-solution. The dataset had 60 images consisting of 15 classes. For each class, we had two images that were used for training and two new for the model images. We expanded the dataset with 40 images of ten new classes of cotton and cucumbers diseases. We have compared our old Siamese model with cross-entropy loss, the new model with triplet loss and few commercial AutoML platforms: Google Cloud Vision, Microsoft Custom Vision, and IBM Watson Visual Recognition. The results of the comparison on the test dataset of 100 images of 25 classes presented in table 1.

Table 1. Comparison of accuracies our models and AutoML solutions (%)

Old model	New model	Google	Microsoft	IBM
91	98	88	92	91

We are focused on the development of production-ready solution for plant disease detection and we have some tests for our model. One of the tests is the evaluation of the model on all our images. Tabl. 2 shows the confusion matrix for our current production model. There are 30 wrongly classified images of 935 images in total. One can see that most of the wrong recognition is for the same disease on the different crops or the diseases that look similar.

Crop losses due to diseases is a serious problem; however, the number of real-life applications allowing to detect the diseases is very limited. Owing to the lack of production field image datasets, it is necessary to improve One-shot and Few-shot learning methods. Our work shows that even with a limited dataset it is possible to train a good model. We have already deployed the new model into our PDD portals and mobile application thus it is available to the farmers’ community.

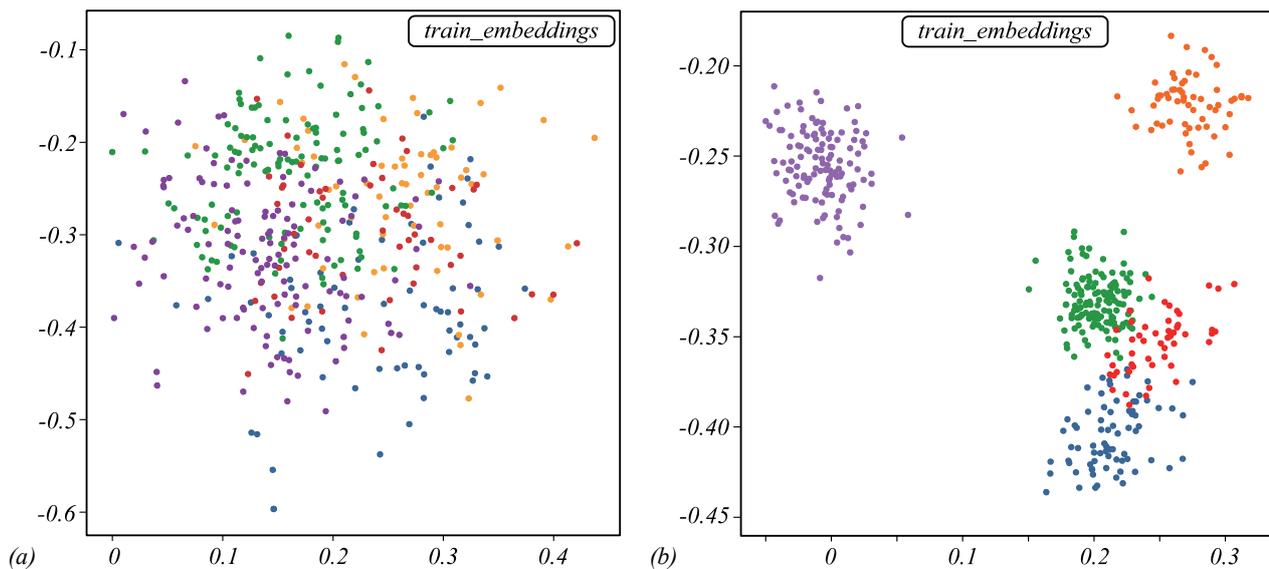


Fig. 6. Representation of the extracted feature vectors of 5 moss classes in 2D: (a) before training, (b) after 300 epochs of training

Air pollution has a significant negative impact on various components of ecosystems and human health. Moss as a bioaccumulators is a great data source for environmental monitoring. The identification of moss species is important for the quality of the analysis. We have deployed our new model into an application that allows filling in the information about sampling sites required by the UNECE ICP Vegetation manual. We hope this will help the contributors of the 2020–2022 survey to classify moss species right.

The accuracy of the resulting model shows that the combination of Siamese networks with the triplet loss function is highly promising in the case of vegetation classification tasks. We have shown effectiveness on two tasks; however, we believe that such an approach can be used in many other applications.

For us now, it is impossible to create a dataset for detector algorithms like Fast R-CNN, Faster R-CNN, YOLO, but we would like to have an ability to recognize disease on a part of an image. The processing time of our model decreased after quantization thus, we are going to build an R-CNN object detector based on our model and OpenCV abilities (selective search, confidence filtering, and non-maxima suppression).

We will continue to expand our databases and improve our models; event though, we already have good results for both of our tasks.

### Conclusion

We applied the Siamese networks along with the triplet loss function to solve two vegetation classification tasks. The model for plant disease detection trained on 25 classes of five crops shows 97.8% accuracy. The model for moss species classification on five classes shows 97.6% accuracy. This result is much better than the transfer learning approach itself with a base network like

ResNet50 and better than our previous approach with the Siamese network with two twins and with cross-entropy loss. The combination of the Siamese network with the triplet loss function has a great potential for classification tasks with a very small training dataset. We had only 935 images of diseased plants and 599 images of mosses; nevertheless, we manage to obtain good results. The dataset and the models are accessible via [pdd.jinr.ru](http://pdd.jinr.ru) and [moss.jinr.ru](http://moss.jinr.ru). We are going to expand our databases, improve our models, and build an R-CNN object detector based on our model and OpenCV.

### References

- [1] Uzhinskiy AV, OsoskovGA, Goncharov PV, Nechaevskiy AV. Multifunctional platform and mobile application for plant disease detection. CEUR Workshop Proc 2019; 2507: 110-114.
- [2] Goncharov P, Ososkov G, Nechaevskiy A, Uzhinskiy A, Nestsianenia I. Disease detection on the plant leaves by deep learning. In Book: Kryzhanovsky B, Dunin-Barkowski W, Redko V, Tiumentsev Y, eds. Advances in Neural Computation, Machine Learning, and Cognitive II. Cham, Switzerland: Springer Nature Switzerland AG; 2019:151-159.
- [3] Goncharov P, Uzhinskiy A, Ososkov G, Nechaevskiy A, Zudikhina J. Deep siamese networks for plant disease detection. EPJ Web of Conferences 2020; 226: 03010.
- [4] Schroff F, Kalenichenko D, Philbin J. FaceNet: A unified embedding for face recognition and clustering. 2015 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2015: 815-823. DOI: 10.1109/CVPR.2015.7298682.
- [5] Uzhinskiy A, Ososkov G, Frontasieva M. Management of environmental monitoring data: UNECE ICP Vegetation case. CEUR Workshop Proc 2019; 2507: 202-207.
- [6] Hughes D, Salathé M. An open access repository of images on plant health to enable the development of mobile disease diagnostics through machine learning and crowdsourcing. arXiv Preprint 2015. Source: <https://arxiv.org/abs/1511.08060>.
- [7] Mohanty S, Hughes D, Salathé M. Using deep learning for image-based plant disease detection. Front Plant Sci 2016; 7: 1419.

[8] Too EC, Yujian L, Njuki S, Yingchun L. A comparative study of fine-tuning deep learning models for plant disease identification. *Comput Electron Agric* 2019; 161: 272-279.

[9] Ferentinos KP. Deep learning models for plant disease detection and diagnosis. *Comput Electron Agric* 2018; 145: 311-318.

[10] Fuentes A, Yoon S, Kim S, Park D. A robust deep-learning-based detector for real-time tomato plant diseases and pests recognition. *Sensors* 2017; 17: 2022.

[11] Türkoğlu M, Hanbay D. Plant disease and pest detection using deep learning-based features. *Turk J Elec Eng & Comp Sci* 2019; 27: 1636-1651.

[12] Selvaraj M, Vergara A, Ruiz H, Safari N, Elayabalan S, Ocimati W, Blomme G. AI-powered banana diseases and pest detection. *Plant Methods* 2019; 15: 92.

[13] Saleem M, Potgieter J, Arif K. Plant disease detection and classification by deep learning. *Plants* 2019; 8: 468.

[14] Ise T, Minagawa M, Onishi M. Classifying 3 moss species by deep learning, using the “chopped picture” method. *Open J Ecol* 2018; 8: 166-173.

[15] Cheng D, Gong Y, Zhou S, Wang J, Zheng N. Person re-identification by multi-channel-parts-based cnn with improved triplet loss function. 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR) 2016: 1335-1344.

[16] Hermans A, Beyer L, Leibe B. In defense of the triplet loss for person re-identification. arXiv preprint. Source: <https://arxiv.org/abs/1703.07737>.

[17] Dong X, Shen J. Triplet loss in Siamese network for object tracking. *Proc European Conference on Computer Vision (ECCV)* 2018; 459-474.

[18] Puch S, Sánchez I, Rowe M. Few-shot learning with deep triplet networks for brain imaging modality recognition. In Book: Wang Q, Milletari F, Nguyen HV, Albarqouni S, Jorge Cardoso M, Rieke N, Xu Z, Kamnitsas K, Patel V, Roysam B, Jiang S, Zhou K, Luu K, Le N, eds. Domain adaptation and representation transfer and medical image learning with less labels and imperfect data. Springer; 2019.

[19] Anshul T, Daksh T, Padmanabhan R, Aditya N. Deep metric learning for bioacoustic classification: Overcoming training data scarcity using dynamic triplet loss. *J Acoust Soc Am* 2019; 146: 534-547.

[20] Zhang J, Lu C, Wang J, Yue X, Lim S, Al-Makhadmeh Z, Tolba A. Training convolutional neural networks with multi-size images and triplet loss for remote sensing scene classification. *Sensors* 2020; 20(4): 1188.

Table 2. Confusion matrix for 935 images from plant disease dataset

	Cotton Alternaria leaf blight	Cotton Healthy	Cotton Nutrient deficiency	Cotton Powdery mildew	Cotton Verticillium wilt	Grape Black rot	Grape Chlorosis	Grape Esca	Grape Healthy	Grape Powdery mildew	Corn Downy mildew	Corn Eyespot	Corn Healthy	Corn Northern Leaf	Corn Southern rust	Wheat Black chaff	Wheat Brown rust	Wheat Healthy	Wheat Powdery mildew	Wheat Yellow rust	Cucumber Anthracno	Cucumber Downy mildew	Cucumber Healthy	Cucumber Nutrient deficiency	Cucumber Powdery mildew
Cotton Alternaria leaf blight	33	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cotton Healthy	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cotton Nutrient deficiency	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Cotton Powdery mildew	1	0	0	28	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
Cotton Verticillium wilt	0	0	0	0	30	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Grape Black rot	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Grape Chlorosis	0	0	0	0	0	0	49	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Grape Esca	0	0	0	0	0	0	0	73	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Grape Healthy	0	0	0	0	0	0	0	0	121	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Grape Powdery mildew	0	0	0	0	0	0	0	0	0	22	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn Downy mildew	0	0	0	0	0	0	0	0	0	0	33	0	0	0	0	0	0	1	0	0	0	0	0	0	0
Corn Eyespot	0	0	0	0	0	0	0	0	0	0	0	31	0	0	0	0	0	0	0	0	0	0	0	0	0
Corn Healthy	0	0	0	0	0	0	0	0	0	0	0	0	34	0	0	0	0	0	0	0	0	0	0	0	0
Corn Northern Leaf	0	0	0	0	0	0	0	0	0	0	0	0	0	34	0	0	0	3	0	0	0	0	0	0	1
Corn Southern rust	0	0	0	0	0	0	0	0	0	0	0	0	0	0	32	0	0	0	0	1	0	0	0	0	0
Wheat Black chaff	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	21	1	0	0	0	0	0	0	0	0
Wheat Brown rust	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	30	0	0	1	0	0	0	0	0
Wheat Healthy	0	0	0	0	0	0	0	0	0	1	0	0	3	0	0	0	0	25	0	1	0	0	0	0	0
Wheat Powdery mildew	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	0	0	0	0	0	0
Wheat Yellow rust	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	1	1	0	0	29	0	0	0	0	0
Cucumber Anthracno	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	30	2	0	0	0
Cucumber Downy mildew	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	33	0	1	2
Cucumber Healthy	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	36	0	0
Cucumber Nutrient deficiency	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	32	0
Cucumber Powdery mildew	0	0	0	1	0	0	0	0	0	1	0	0	0	0	0	0	0	0	0	0	2	0	0	0	30

---

**Authors' information**

**Alexander Vladimirovich Uzhinskiy**, (b. 1983) Ph.D. in Technology, graduated from International University for Nature, Society, and Man “Dubna” in 2006. Works as a lead programmer at Joint Institute for Nuclear Research. Research interests: neural network, computer vision, software engineering, remote sensing. E-mail: [auzhinskiy@jinr.ru](mailto:auzhinskiy@jinr.ru).

**Gennady Alexeevich Ososkov**, (b. 1931), Dr. Sc., Professor, graduated from Moscow State University in 1953 with honor. Works as Principal Researcher of Joint Institute for Nuclear Research. Research interests: neural network, image processing, pattern recognition. E-mail: [ososkov@jinr.ru](mailto:ososkov@jinr.ru).

**Pavel Vladimirovich Goncharov**, (b. 1996), graduated from Pavel Sukhoi State Technical University in 2017 with honor. Works as Data Scientist at IHS Markit, also works as software engineer at Joint Institute for Nuclear Research. Research interests: deep learning, computer vision, natural language processing, few-shot learning, optimization. E-mail: [kaliostrogoblin3@gmail.com](mailto:kaliostrogoblin3@gmail.com).

**Andrey Vasilevich Nechaevskiy**, (b. 1982), graduated from International University for Nature, Society, and Man “Dubna” in 2006, majoring in System Analysis. Currently he works as the leading programmer at Joint Institute for Nuclear Research. Research interests are neural networks, simulation, system analysis. E-mail: [nechav@jinr.ru](mailto:nechav@jinr.ru).

**Artem Alekseevich Smetanin** (b. 1998) graduated from Dubna State University in 2020 with a degree in Computer Science and Engineering. Currently, he holds a master's degree at ITMO University. His research interests include pattern recognition, deep neural networks, and computer vision. E-mail: [artem\\_smetanin@niuitmo.ru](mailto:artem_smetanin@niuitmo.ru).

---

*Code of State Categories Scientific and Technical Information (in Russian – GRNTI): 28.23.15*  
*Received December 28, 2020. The final version – February 24, 2021.*

---