Research Journal of Agricultural Science, 52 (3), 2020

MOBILE APPLICATION FOR DETECTING SOME WHEAT PATHOGENS USING AI

Raluca INCICAȘ*, Otilia COTUNA, F. SALA*
Banat University of Agricultural Sciences and Veterinary Medicine "King Michael I of Romania" from Timisoara, Timisoara, 300645, Romania
Corresponding author: ralu.incicas96@gmail.com; florin_sala@usab-tm.ro

Abstract. *The present study aimed to develop an application for the recognition of five pathogens in wheat culture based on artificial intelligence (AI). Machine learning (ML), an important branch of AI, was the basis for the application of the pathogens recognition in wheat culture. Five pathogens in wheat cultivation were studied, Blumeria graminis, Pyrenophora tritici repentis, Puccinia recondita, Puccinia striiformis and Puccinia graminis. A data set of 323 images with pathogens studied in wheat culture was used. The images have been processed and transformed so that the model receives the same size for each image. The first step in building the data set to train the ML model was data augmentation, in order to increase the number of data through known changes. A training set and a validation set were used. Google Colaboratory was used to build the ML model. The React Native framework was used to have an application available on both iOS and Android. Heroku and Flask were used to integrate the systems. In order to evaluate the model the ”Class Activation Map visualization” (CAM) was used. One of the techniques that CAM uses is to produce heatmaps (areas of interest) on the parts of the image that correspond to the different classes, over the input images. The class activation heatmap is a 2D network of scores associated with a specific output class, calculated for each area of the input image, indicating how important each area is relative to the output class. To view these heatmaps on the images taken in the study, the ClassificationInterpretation class was used. Another method used in evaluating the model was the confusion matrix which will show for each label how many times it was correctly predicted. The model correctly predicted in a percentage of 88.4% for Puccinia striiformis, 72.03% for Puccinia recondita and 94.67% for Blumeria graminis.*

***Keywords**: artificial intelligence, identification, machine learning, pathogens, wheat*

INTRODUCTION

Artificial intelligence is an area of computer science that highlights the creation of intelligent machines that work and react like humans (TECHOPEDIA). For 60 years, experts have said that the future of AI is approaching. This period called "AI winter" has ended due to several aspects: access to cheaper processing, Big Data ("the university where AI studies"), better algorithms (neural networks).

Several types of artificial intelligence are known, such as: Weak AI, Strong AI, Reactive Machines, Limited Memory, Theory of Mind, Self – awareness, Artificial general intelligence (AGI), Artificial narrow intelligence (ANI) etc. (FJELLAND, 2020).

Among the most important branches of Artificial Intelligence, Machine Learning (ML), Neural Network (NN), Natural language processing (NLP), Expert system (ES), Fuzzi Logic, Vision, Speech, Planning, Robotics, are considered with applications in different fields (MISHRA et al., 2016; RAI and CHATTERJEE, 2020; SCOLTIS et al., 2020).

Machine learning (ML) has already shown a special interest in agriculture, providing opportunities for the operationalization of large databases, in the development of intelligent and high-performance agro-technologies (LIAKOS et al., 2018). Machine learning was used in the authentication of some parameters for crops and agricultural products, the control of some vegetation factors, decision systems in different agricultural systems. (MORELLOS et al., 2016; FRAGNI et al., 2018; MAIONE and BARBOSA, 2019; ZHOU et al., 2018)

Imaging analysis was used in agriculture to study areas, crops, optimize image analysis according to time and angle of image acquisition, foliar analysis in plants, being useful in a limited volume of data (HERBEI and SALA, 2014; GOVEDARICA et al., 2015; HERBEI et al., 2015; DRIENOVSKY et al., 2017a,b). Machine learning has advantages in terms of processing a large volume of images and data, and has been used in studies in various crops for the study and detection of plant pathogens (PINEDA et al., 2018; MISHRA et al., 2019; SALEEM et al., 2019; SPERSCHNEIDER, 2019).

Machine learning was also used to study stress tolerance in wheat (MOGHIMI et al., 2018), the localization of compounds in plants(ZHANG et al., 2018), molecular plant biology (SILVA et al., 2019), biotechnology and plant breeding (NIAZIAN and NIEDBAŁA, 2020) etc.

The present study aimed to develop an application for the recognition of five pathogens in wheat culture.

MATERIAL AND METHODS

In order to develop the application for the recognition of the symptoms of five foliar pathogens of cereals, a set of images taken in the experimental fields established on the territory of SDT of USAMVB Timișoara was used. For this, photos were taken for pathogens: *Blumeria graminis, Pyrenophora tritici repentis, Puccinia recondita, Puccinia striiformis and Puccinia graminis*. The photos were taken during spring, between 2018 and 2019, for four of the five pathogens. Images taken in an earlier period have been made available for *Puccinia graminis* by COTUNA (2017).

In figure 1 the architecture of the proposed system is illustrated. Google Colaboratory was used [towardsdatascience.com] to build the ML model. We decided to use the React Native [reactnative.dev] framework to have an application available on both iOS and Android. To integrate the systems we used Heroku [devcenter.heroku.com] and Flask [flask.palletsprojects.com].

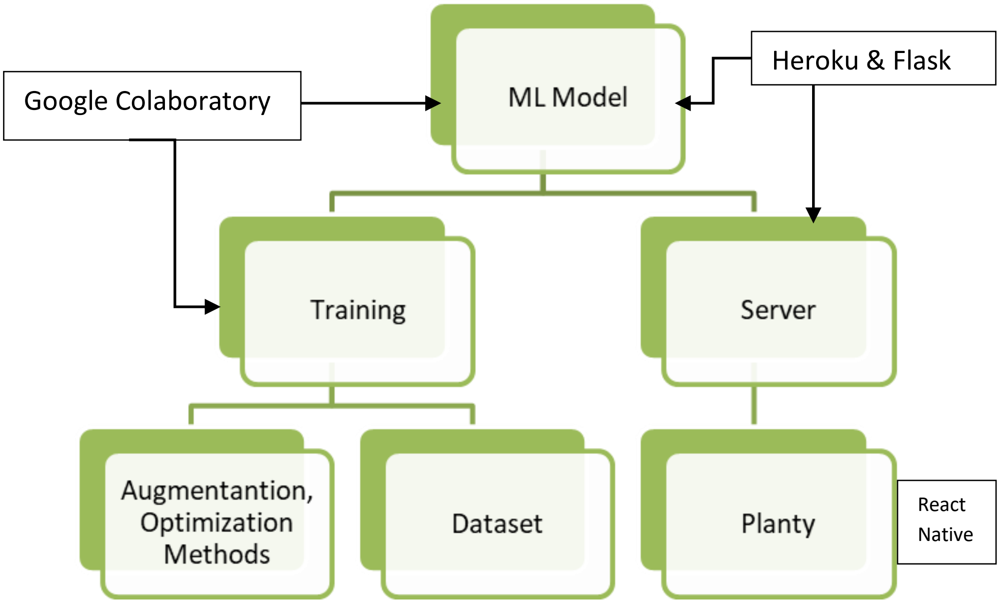


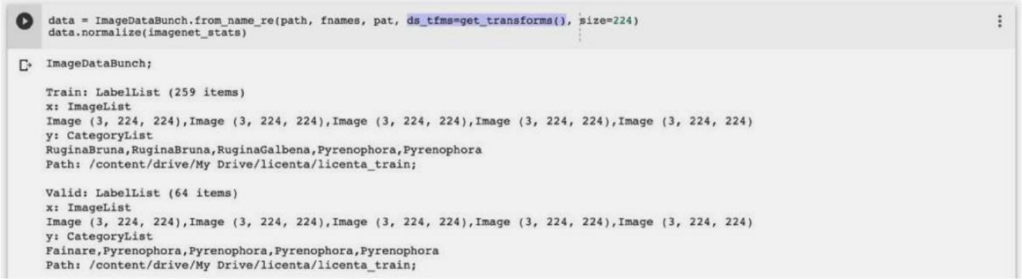
Fig. 1. The architecture of the proposed solution

The first step in building the data set to train the ML model was to augment the data. Data augmentation means increasing the number of data through known changes, which make sense in the context of training and help the model to be more robust.

The question is: But how and why? The answer to this question is simple. We do not have an infinite amount of data. The more data we have, the better the ML model will be. However, each data collection comes at a cost. This cost can mean money, human effort, computational resources and of course, time spent during the process.

Therefore, we may need to increase the data set we give to ML classifiers to offset the costs involved in future data collections. In practice, what happens is that geometric and color augmentations are applied, meaning the images change randomly, changes such as: turning, zooming, rotating, stretching, changing the lighting and padding parameters.

To apply the transformations to our dataset, when the Image-Data Bunch object is created, we send it the list of transformations. Fastai has a predefined list, obtained from intense team experiments (BHUTANY, 2019; towardsdatascience.com). This object returns a tuple of length 2, containing 2 lists: one for the training set and the other for the validation set as we can see in figure 2.



```
data = ImageDataBunch.from_name_re(path, fnames, pat, ds_tfms=get_transforms(), size=224)
data.normalize(imagenet_stats)
```

```
ImageDataBunch;

Train: LabelList (259 items)
x: ImageList
Image (3, 224, 224),Image (3, 224, 224),Image (3, 224, 224),Image (3, 224, 224),Image (3, 224, 224)
y: CategoryList
RuginaBruna,RuginaBruna,RuginaGalbena,Pyrenophora,Pyrenophora
Path: /content/drive/My Drive/licenta/licenta_train;

Valid: LabelList (64 items)
x: ImageList
Image (3, 224, 224),Image (3, 224, 224),Image (3, 224, 224),Image (3, 224, 224),Image (3, 224, 224)
y: CategoryList
Fainare,Pyrenophora,Pyrenophora,Pyrenophora,Pyrenophora
Path: /content/drive/My Drive/licenta/licenta_train;
```

Fig. 2. ImageDataBunch Object.

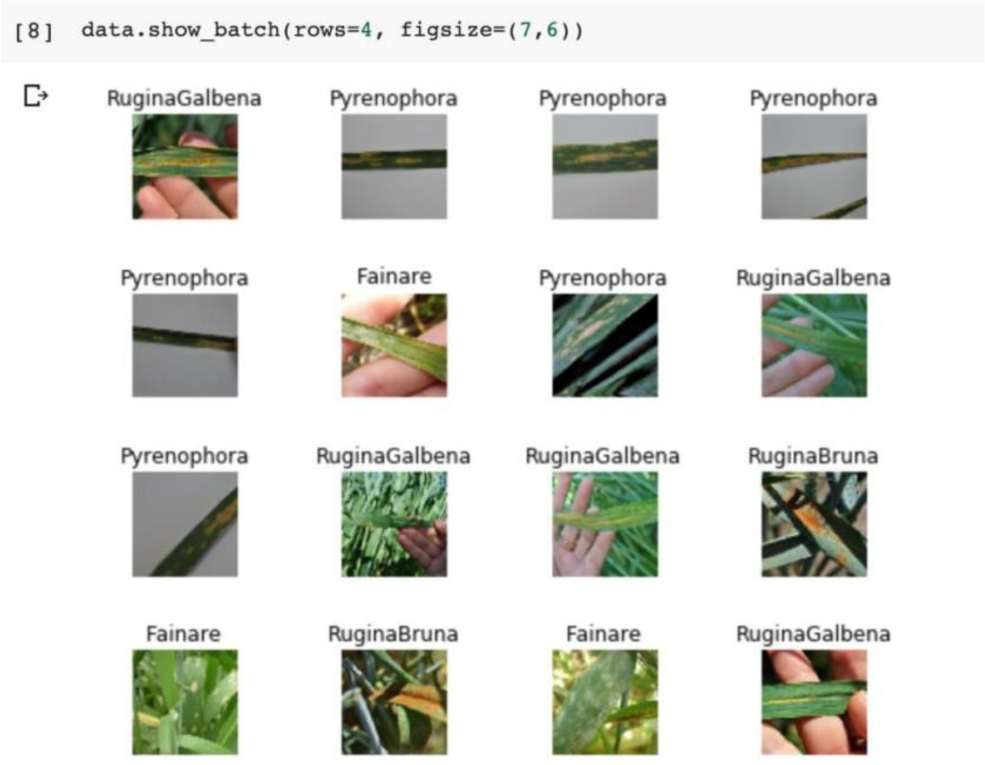
Machines learn through a loss function. It is a method of evaluating the way in which specific algorithms model the given data. If the predictions deviate too much from the actual results, the train loss function would show a very large number. Gradually, with the help of an optimization function, the network learns to reduce the loss function that is the error in prediction, changing its parameters.

There is no "one-size-fits-all" data loss function for algorithms in ML. There are different factors involved in choosing a loss function for a specific problem, such as the type of ML algorithm chosen, the ease of calculating the derivatives and, to some extent, the percentage of exaggerated values in the data set. The default loss function used is FlattenedLoss of CrossEntropyLoss.

Data is at the heart of any Machine Learning problem. In recent years, steps taken using AI would not have been possible without access to relevant data. The data used for this study were accessed from the database of the discipline of phytopathology (COTUNA, 2017).

Thus, the data set consists of 323 images comprising five pathologies specific to cereals. The classes of pathologies or pathogens used were: *Blumeria graminis, Pyrenophora tritici repentis, Puccinia recondita, Puccinia striiformis, Puccinia graminis*.

As described above, the images have been transformed so that the model receives the same size for each image, so after all the transformations, the dataset looks like in figure 3.



```
[8]  data.show_batch(rows=4, figsize=(7,6))
```

Fig. 3. Images in the data set prepared to be able to train the model (Fainare - *Blumeria graminis*; Pyrenophora - *Pyrenophora tritici repentis*; RuginaBruna - *Puccinia recondite*; RuginaGalbena - *Puccinia striiformis*; RuginaNeagra - *Puccinia graminis*)

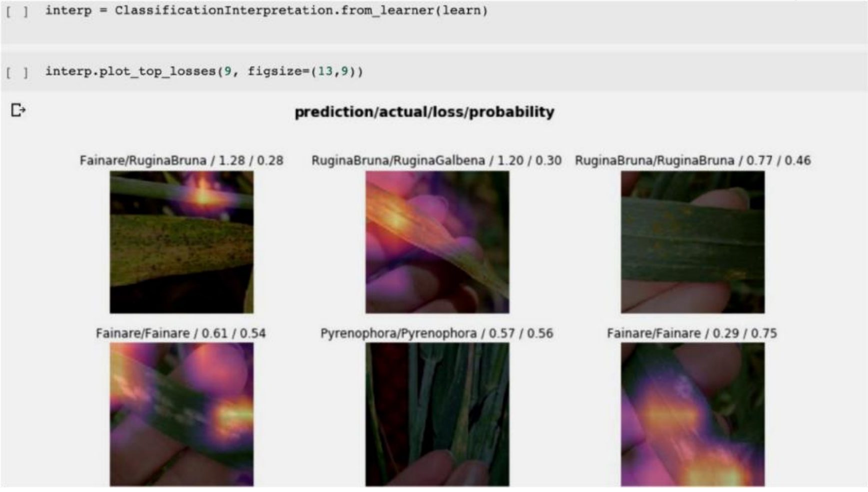
RESULTS AND DISCUSSIONS

The application for the recognition of some diseases from cereal crops was initially thought to be able to recognize five foliar diseases that occur in wheat crops. Of course, there is a possibility to develop this application in such a way that it can recognize all the diseases that occur in cereals and many more. The images used were made in the experimental fields of USAMVB Timișoara, Phytopathology discipline.

The results obtained are good, encouraging. In the following we will show the main steps taken in building the disease recognition application based on external symptoms, visible to the naked eye. As described in the Material and Method chapter, results were obtained based on the trained model, which are presented and discussed below.

To evaluate the model we used the method "Class Activation Map (CAM) visualization". One of the techniques that CAM uses is to produce heatmaps (areas of interest) on the parts of the image that correspond to the different classes, over the input images (Paliwal, 2018). The class activation heatmap is a 2D network of scores associated with a specific output class, calculated for each area of the input image, indicating how important each area is relative to the output class. To view these heatmaps on images taken in the studio, the ClassificationInterpretation class was used.

The CAM view of the data set after it has been trained is shown in the figure 4.

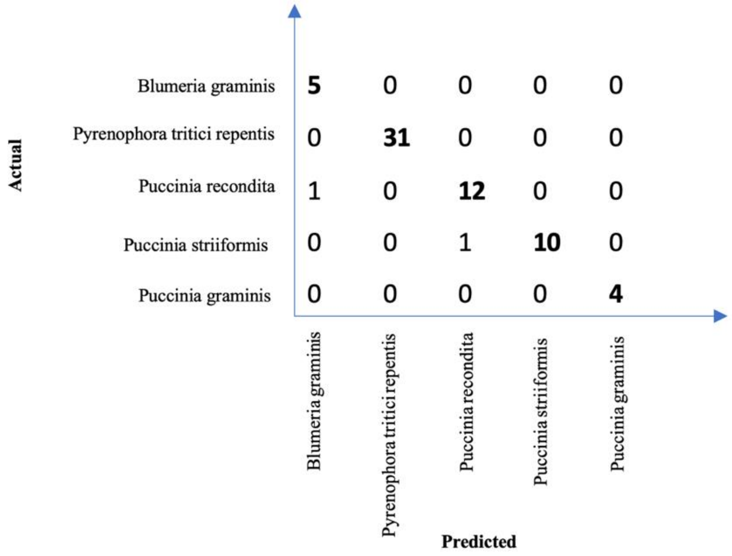


```
[ ]  interp = ClassificationInterpretation.from_learner(learn)
```

```
[ ]  interp.plot_top_losses(9, figsize=(13,9))
```

Fig. 4. CAM viewing of images after the model has been trained

Another method used in evaluating the model was the confusion matrix which will show for each label how many times it was correctly predicted. Confusion matrix is a good technique to summarize the performance of a classification algorithm, figure 5.



Actual \ Predicted	Blumeria graminis	Pyrenophora tritici repentis	Puccinia recondita	Puccinia striiformis	Puccinia graminis
Blumeria graminis	**5**	0	0	0	0
Pyrenophora tritici repentis	0	**31**	0	0	0
Puccinia recondita	1	0	**12**	0	0
Puccinia striiformis	0	0	1	**10**	0
Puccinia graminis	0	0	0	0	**4**

Fig. 5. Confusion matrix (shows for each model how many times it was predicted correctly)

Below are the results obtained when the model interprets new images. For the test images we will have as a result the predictions, namely in what percentage the given image belongs to one of the classes, but also the image over which a “heatmap” will appear to see which points in the image favored the respective class.

Test image for *Puccinia striiformis* is shown in the figure 6. As can be seen in the picture, the model predicts 88.4% that the disease present on the leaves is *Puccinia striiformis*.

```
[0.534239 0.03264  0.634466 0.884206 0.101735]
['Fainare', 'Pyrenophora', 'RuginaBruna', 'RuginaGalbena', 'RuginaNeagra']
```

Fig. 6. The predicted result of the model and the heatmap that shows the part of the image after which it predicted the pathogen

Test image for *Puccinia recondita* is shown in the figure 7. It can be seen in the image that the model predicts with an accuracy of 72.03% that the disease on the leaves is Puccinia recondita.

```
[0.642646 0.024403 0.720325 0.324196 0.143852]
['Fainare', 'Pyrenophora', 'RuginaBruna', 'RuginaGalbena', 'RuginaNeagra']
```

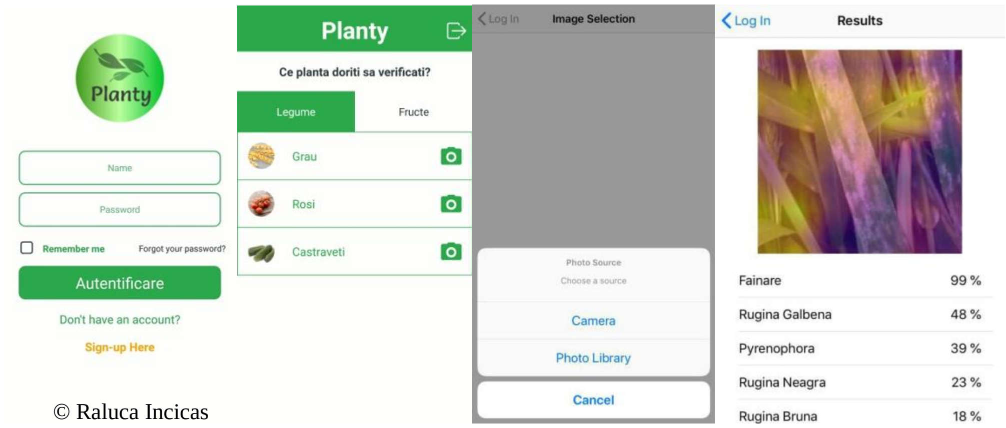
Fig. 7. The predicted result of the model and the heatmap that shows the part of the image after which it predicted the pathogen

Test image for *Blumeria graminis* is shown in the figure 8. As in previous situations, the model predicts with an accuracy of 94.67% that the disease on the leaves is *Blumeria graminis*.

```
[0.94672  0.003805 0.422032 0.716437 0.486702]
['Fainare', 'Pyrenophora', 'RuginaBruna', 'RuginaGalbena', 'RuginaNeagra']
```

Fig. 8. The predicted result of the model and the heatmap that shows the part of the image after which it predicted the pathogen

In order to be able to use this method by various users, we decided to develop a mobile application in which to use the model described above, figure 9.



© Raluca Incicas

Fig. 9. Application interface

CONCLUSIONS

We can say that the developed application works well, being able to recognize with a fairly high accuracy the pathogens for which it was created. On the other hand, however, its testing was not carried out in field conditions, as it was in the prototype stage.

We are certainly just starting out, the application will be developed, but we believe that the use of artificial intelligence for this kind of problem should be encouraged, by facilitating access to various information such as some databases.

The model correctly predicted in a percentage of 88.4% for *Puccinia striiformis*, 72.03% for *Puccinia recondita* and 94.67% for *Blumeria graminis*. For the remaining 2 pathogens we are still developing our model.

Also, the application can be adapted for other studies at the level of the leaf surface of the plants, or at the level of agricultural crops.

BIBLIOGRAPHY

BHUTANI, S., 2019 - Introduction to Image Augmentations using the fastai library, online, Apr 14,

COTUNA, O., 2017 - Image database, phytopathology studies (2016 – 2018 period).

DRIENOVSKY, R., NICOLIN, A.L., RUJESCU, C., SALA, F., 2017a - Scan LeafArea – A software application used in the determination of the foliar surface of plants. Research Journal of Agricultural Sciences, 49(4): 215-224.

DRIENOVSKY, R., NICOLIN, A.L., RUJESCU, C., SALA, F., 2017b - Scan Sick & Healthy Leaf – A software application for the determination of the degree of the leaves attack, Research Journal of Agricultural Sciences, 49(4): 225-233.

FJELLAND, R., 2020 - Why general artificial intelligence will not be realized. Humanities and Social Communications, 7: 10.

FRAGNI, R., TRIFIRÒ, A., NUCCI, A., SENO, A., ALLODI, A., DI, ROCCO, M., 2018 - Italian tomato-based products authentication by multi-element approach: A mineral elements database to distinguish the domestic provenance. Food Control, 93, 211-218.

GOVEDARICA, M., RISTIC, A., HERBEI, M.V., SALA, F., 2015 - Object oriented image analysis in remote sensing of forest and vineyard areas. BulletinUASVM Horticulture, 72(2): 362-370.

HERBEI, M.V., SALA, F., 2014 - Using GIS technology in processing and analyzing satellite images - case study Cheile Nerei Beusnița National Park, Romania. Journal of Horticulture, Forestry and Biotechnology, 18(4): 113-119.

HERBEI, M., SALA, F., BOLDEA, M., 2015 - Relation of Normalized Difference Vegetation Index with some spectral bands of satellite images. AIP Conference Proceedings, 1648:670003-1 – 670003-4.

LIAKOS, K., BUSATO, P., MOSHOU, D., PEARSON, S., BOCHTIS, D., 2018 - Machine Learning in Agriculture: A Review. Sensors, 18: 2674.

MAIONE, C., BARBOSA, R.M., 2019 - Recent applications of multivariate data analysis methods in the authentication of rice and the most analyzed parameters: A review. Critical Review in Food Science and Nutrition, 59(12): 1-12.

MISHRA, S., MISHRA, D., SANTRA, G.H., 2016 - Applications of Machine Learning Techniques in Agricultural Crop Production: A Review Paper. Indian Journal of Science and Technology, 9(38): 1-14.

MISHRA, B., KUMAR, N., MUKHTAR, M.S., 2019 - Systems Biology and Machine Learning in Plant-Pathogen Interactions. Molecular Plant-Microbe Interaction Journal, 32(1): 45-55.

MOGHIMI, A., YANG, C., MILLER, M.E., KIANIAN, S.F., MARCHETTO, P.M., 2018 - A novel approach to assess salt stress tolerance in wheat using hyperspectral imaging. Frontiers in Plant Science, 9: 1182.

MORELLOS, A., PANTAZI, X.-E., MOSHOU, D., ALEXANDRIDIS, T., WHETTON, R., TZIOTZIOS, G., WIEBENSOHN, J., BILL, R., MOUAZEN, A.M. 2016 - Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy. Biosystems Engineering, 152: 104-116.

NIAZIAN, M., NIEDBAŁA, G., 2020 - Machine Learning for Plant Breeding and Biotechnology. Agriculture, 10(10): 436.

PALIWAL, A., 2018 - Understanding your Convolution network with Visualizations. Online, Oct 1, 2018,https://towardsdatascience. com/understanding-your-convolution-network-with-visualizations-a4883441533b.

PINEDA, M., PÉREZ-BUENO, M.L., BARÓN, M., 2018 - Detection of bacterial infection in melon plants by classification methods based on imaging data. Frontiers in Plant Science, 9: 164.

RAI, H.M., CHATTERJEE, K., 2020 - Detection of brain abnormality by a novel Lu-Net deep neural CNN model from MR images. Machine Learning with Applications, 2: 100004.

SALEEM, M.H., POTGIETER, J., ARIF, K.M., 2019 - Plant disease detection and classification by deep learning. Plants, 8: 468.

SILVA, J.C.F., TEIXEIRA, R.M., SILVA, F.F., BROMMONSCHENKEL, S.H., FONTES, E.P.B., 2019 - Machine learning approaches and their current application in plant molecular biology: A systematic review. Plant Science, 284: 37-47.

SPERSCHNEIDER, J., 2019 - Machine learning in plant–pathogen interactions: empowering biological predictions from field scale to genome scale. New Phytologist, 228(1): 35-41.

SOLTIS, P.S., NELSON, G., ZARE, A., MEINEKE, E.K., 2020 - Plants meet machines: Prospects in machine learning for plant biology. Applications in Plant Science, 8(6): e11371.

ZHANG, N., RAO, R.S.P., SALVATO, F., HAVELUND, J.F., MØLLER, I.M., THELEN, J.J., XU, D., 2018 - MU-LOC: A machine-learning method for predicting mitochondrially localized proteins in plants. Frontiers in Plant Science, 9: 634.

ZHOU, C., LIN, K., XU, D., CHEN, L., GUO, Q., SUN, C., YANG, X., 2018 - Near infrared computer vision and neuro-fuzzy model-based feeding decision system for fish in aquaculture. Computers and Electronics in Agriculture, 146: 114-124.

https://www.techopedia.com/definition/190/artificial-intelligence-ai; Accesed on 10.04.2019

https://towardsdatascience.com/getting-started-with-google-colab-f2fff97f594c; Accesed on 05.02.2019

https://reactnative.dev/docs/getting-started; Accesed on 11.10.2018

https://devcenter.heroku.com/articles/getting-started-with-python; Accesed on 10.03.2019

https://flask.palletsprojects.com/en/1.1.x/; Accesed on 10.05.2019

https://towardsdatascience.com/introduction-to-image-augmentations-using-the-fastai-library-692dfaa2da42; Accesed on 26.03.2019