Journal of Environmental Sciences Studies

Journal home page: www.jess.ir

Non-destructive Estimation of Chemical Properties in Braeburn Apple using Convolutional Neural Network

Razieh Pourdarbani[a,*], Sajad Sabzi[b], Farzad AzadShahraki[c]

*a. Associate Prof., Dept. of Biosystem engineering, University of Mohaghegh Ardabili, Ardabil 56199-11367, Iran

b. Postdoc researcher, Dept. of Computer, Sharif University of Technology, Tehran 11365-11155, Iran

c. Agricultural Engineering Research Institute, Agricultural Research, Education and Extension Organization (AREEO), Karaj, Iran

*Email Address: r_pourdarbani@uma.ac.ir

Received: (November-12-2021)
Accepted: (January- 4 -2022)

Abstract

Determination of appropriate time to harvest apples prevents its waste and also affects the quality of the fruit. This time depends on the variety, climate during the growing season and also the purpose of harvesting the fruit. The aim of this study is to predict non-destructive chemical properties related to harvesting such as starch, soluble solids and acidity using spectral data and implementation of deep learning algorithm. First, images of Braeburn apples were taken by hyperspectral camera in 4 different stages of maturity. Next, the spectral information was extracted. Then the chemical properties of starch, titrtable acidity (TA) and total soluble solids (TSS) were measured using destructive methods in laboratory. Eventually, the prediction model was created by convolutional neural network (CNN). The results illustrated that the coefficient of determination and the mean squared error for the properties of starch, TSS and TA were 95.4%, 4.8, 91.6%, 0.284, 84.2% and 0.424, respectively.

Keywords

Apple, Prediction, Starch, TA, TSS, CNN.

Introduction

Apples are one of those fruits that have a higher supply to demand ratio at fresh market in Iran. Therefore, for price adjustment in market, the availability of apple in all seasons and the balance between supply and demand, the product must be stored in the correct way to increase the shelf life of apples. The shelf life of apples in the cooling store depends on many factors. One of the most important factor is the time of harvest. In general, over- ripe fruit have shorter shelf life. Thus, the question that arises is how to determine the appropriate time to harvest the apple in orchard. Various methods including destructive and non-destructive have been used to determine the level of fruit maturity. Among them, near-infrared spectroscopy is one of the most widely used methods in the field of quality assessment of agricultural products [1-3].Agriculture can always be useful to help farmers by using artificial intelligence and digitizing various processes. Deep learning is a subset of machine learning that falls into the category of artificial intelligence. Machine learning is the process of training an algorithm to do a task, not planning how to do it step by step. These predictions can include many things such as diagnosing

whether the fruit in a photo is a banana or an apple, detecting moving objects in automat system, diagnosing plant disease, and more.
Traditional non-destructive recognition technology requires complex extraction of features in image, while deep learning technology can automatically train and extract features through nonlinear multilayers. Deep learning applications in agriculture are in various fields such as crop management [4], water management [5], weed identification [6], identification of disease in plants [7-9], estimating crop yields [10-11] and counting fruit on a tree [12].Zhang et al. [13] studied on a data set containing images of seven types of agricultural machinery and six types of undesired categories. A network called AMTNet was designed and trained to automatically detect images of agricultural machinery. Under the same test conditions, AMTNet achieved 97.83% and 100% accuracy in the Top_1 and Top_5 validation sets, respectively. The obtained results were better than the classic ResNet_50 and Inception_v3 networks. Perugachi-Diaz et al. [14] classified images of cabbage seedlings with convolutional neural networks. Seedling growth was tracked over 14 days, so that the data set included 13,200 individual seedlings with the corresponding tags. Classification accuracy was 94%. This research has been deployed as a quick warning tool to assist experts in making important decisions. In addition, this model can be further developed to automate the process. Esmail Karar et al. [15] identified insect pests based on cloud computing using a convolutional neural network. This study was successfully performed for aphid pests, flaxseed worm, flea beetle and red spider. The highest accurate detection rate was 0.99% for all pest images. Franczyk et al. [16] identified grapes using color images and deep learning. They proposed the ExtResnet classification and the accuracy was 89%. Yuan et al. [17] focused on plant disease diagnosis using deep learning. They stated that in deep learning and transmission learning strategies, there is a constant need for better databases including images of real farm-grown crops, which could be a cornerstone of future improvements in view point of quantity and quality. The increasing use of smart mobile terminals also suggests that representing simple models should be an important part of future research considerations. Ismail and Malik [18] proposed a machine learning system based on deep learning to provide a non-destructive and cost-effective solution for automating visual inspection of fruit freshness. The performance of different deep learning structures including ResNet, DenseNet, MobileNetV2, NASNet and EfficNet were compared. The proposed system also provided a real-time visual inspection using a low-cost Raspberry Pi module with a camera and a touch screen for user interaction. The real-time system efficiently segmented image and graded the fruit accurately. This system was tested on two data sets (apple and banana) and the average accuracy was 99.2 and 99.2% for apple and banana, respectively using EfficNet model. During real-time testing, the accuracy was 96.7% for apples and 93.8% for bananas, indicating the efficiency of the developed system. Tetila et al. [19] categorized soybean pest using five deep learning structures. The performance of Inception-v3, Resnet-50, VGG-16, VGG-19 and Xception was evaluated for different transmission learning strategies through a 5,000 image in real-time condition. Results showed that the structure of deep learning trained with fine-tuning can lead to a higher classification rate compared to other approaches and reach a maximum accuracy of 93.82%.Since determination of the most appropriate harvest time reduces waste and increases storage ability in apples, prediction of chemical properties e.g. starch, titratable acidity and total soluble solids provides very useful information to beneficiaries so that it can be properly managed in the storage and marketing of apples. Thus, present study attempt to non-destructive prediction of starch, TSS and TA using 1D convolutional neural network in regard to prediction of appropriate ripening time.

2. Materials and Methods

The different stages of non-destructive estimation of the chemical properties of apples can be seen in Figure 1.

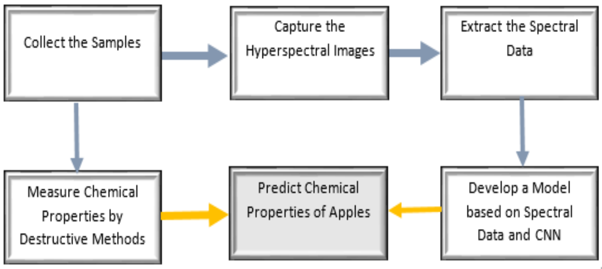

Figure 1. Different stages of non-destructive estimation of apple chemical properties

2.1. Collect the samples to extract the spectral data

According to the purpose of this study, which is to predict the chemical properties of apples in different stages of ripening in Braeburn apples, the first step was determination of harvest time. Hereupon, local gardeners were polled. Then, a total of 187 apples were harvested at 4 stage including 20 days before maturity, 10 days before maturity, on the maturity time and 10 days after maturity.

2.2. Extraction of spectral properties of hyperspectral images

A hyperspectral camera (made in Iran - Noor Iman TajhizCo.; www.hyperspectralimaging.ir) with a spectral range of 400 to 1100 nm was used to capture images. The camera was placed inside a chamber and illuminated by two 10-watt tungsten halogen lamps (SLI-CAL (StellarNet, USA) to block out disturbing ambient light. In this research, a laptop with features (Intel Corei5, 2430M at 2.40GHz, 4GB of RAM, Windows 10) was used to process the data. The raw spectral data was corrected by multiplicative scatter correction (MSC) algorithm in ParLeS, and finally the smoothing operation was performed by wavelet filter algorithm [20].

2.3. Measurement of chemical properties by laboratory destructive method

2.3.1. Starch

The amount of starch was measured based on the method described by [21]. The different steps to measure are as follows:

(1): Separation of a piece of apple in 0.5 g and extract the apple juice, (2): Preparation of buffer using phosphate buffer solution, (3): Mixing the sediment from step (1), with 1.5 ml of buffer solution, (4): Centrifuge at 12000 rpm for 20 minutes, to completely separate the sediment from the mixture, (5): After complete separation of the sediment from the solution, it was mixed by dimethyl sulfoxide / Hydrochloric acid (4: 1) and centrifuged at 12,000 rpm for 20 minutes. (6): Mixing the solution obtained from step (5) with iodine-hydrochloric acid reagent in a ratio of 1 to 5 and recording the number of samples taken using a spectrophotometer (Optizen 2120 UV plus, Company: Mecasys Co., Ltd., Korea). It was read at 600 nm and the results were presented in mg/g. (7): Preparation of standard starch solution using specified concentrations of starch (0 to 100 mg/l). (8): Finally, with the help of data obtained from adsorption numbers of standard starch solutions, a point graph was plotted in Excel and the linear fitting curve of the graph as well as the conversion of adsorption number to starch value was obtained in gr/ml.

2.3.2. Total soluble Solid (TSS)

One of the most common uses of Brix is to determine soluble solids (sugar levels) in fresh produce. Brix number is a measure of the mass ratio of sugar dissolved in a liquid. This amount can be found simply by reading a refractometer. For example, if the brix of a fruit is 32, which means that the fruit contains 32 gr of sugar (any type) in 100 gr of water.

2.3.3. Titratable Acidity (TA)

First, 10 ml of fruit juice is poured into beaker and diluted with 10 ml of distilled water. The beaker is placed on the shaker and under burette, while 4 to 5 drops of phenolphthalein reagent are added. The initial pH is measured before adding NaOH. Then the valve of burette is opened gradually to add the titrant into the juice. When more drops of NaOH fall into the juice, the color of the juice changes but is not stable. Whenever a pale pink color is created in the juice in a stable way, it means that we have reached the end point of the titration. In fact, when the pH is higher than 8.2, this indicates that we are close to the end of the titration. Eventually, the valve is closed and the amount of NaOH consumed is recorded and the following formula is used to calculate the acid in the juice.

$TA = (V_{NaOH} * 180 * 100 * N_{NaOH}) / (1000 * W)$

Here;

V_{NaOH}*: Volume of NaOH consumed

N_{NaOH} : Normality of NaOH

W: Weight of samples

2.4. Non-destructive estimation of chemical properties by CNN in apple

Deep learning is a category of machine learning and a set of algorithms that have a high ability to classified data due to their hierarchical structure. Deep structures can also provide a more comprehensive representation of functions than MLP structures. Deep learning based on artificial neural networks mimics the behavior of the brain when learning a set of samples. Deep learning algorithms have been further developed in the context of artificial neural networks. In common neural networks, the number of hidden layers is usually no more than two. In contrast, when the number of hidden layers increases, these networks are called deep networks. Convolutional neural network is known as one of the most popular types of deep neural networks (ConvNet). It does not require manual feature extraction, so automatic feature extraction makes deep learning models very popular for computer vision tasks such as object classification. CNNs learn to recognize different features of an image using tens or hundreds of hidden layers. Each hidden layer increases the complexity and features of the trained image.The structure of the CNN used in this study is shown in Table 1. The input vector of the algorithm contains spectral data and the output is the actual amount of chemical properties of starch, soluble solids and titratable acidity. It should be noted that after extracting the spectral properties, 70% were randomly used as train data and 30% as test data. In other words, 187 apples were picked in 4 different stages of ripening. 30% of the data (57 apples) were randomly excluded for testing. Due to the fact that the remaining number (130 apples) is not enough to train the deep learning model and on the other hand, the cost of measurement the properties are very expensive, so the artificial data generation method was used and a total of 2500 train data **were produced.**

Table 1.Convolutional neural network structure proposed in this paper

Layer	OutputShap	Param#
conv1d	(None, 1940, 32)	256
Maxpolling	(None, 970, 32)	0
conv1d	(None, 966, 64)	10304
Maxpolling	(None, 483, 64)	0
Dropout	(None, 483, 64)	0
conv1d	(None, 479, 128)	41088
Maxpolling	(None, 239, 128)	0
conv1d	(None, 237, 256)	98560
Maxpolling	(None, 118, 256)	0
conv1d	(None, 116, 512)	393728
Flatten	(None, 59392)	0
Dense	(None, 1)	59393

2.5. Evaluation of the performance of model for estimation of chemical properties in apple

Some statistical parameters were used to evaluate the performance of the prediction model such as Max error [22], Mean absolute error [23], coefficient of determination [24-27], score-variance [28], Mean squared error [29], Median absolute error [30] and Mean squared logarithmic error [31].

3. Results & Discussion

3.1. Starch

The results of performance evaluation of the prediction model for the chemical properties of starch can be seen in Table 2. Considering that the values of the values of variance score (Var) and coefficient of determination (R^2) are close to one, it states that the proposed model for pre-Starch nose has been successful.

Table 2. Performance evaluation of model for prediction of starch

Var	0.954
R2	0.944
MaxE	16.58
MedE	5.02
MSLE	0.12
MSE	37.86
MAE	4.8
Min	2
Max	89

Figure 2 shows the scatter plot of the actual and estimated starch content for the 57 test data.

The proximity of these two values indicates good ability of the proposed model for predicting the starch in apples. Figure 3 shows the train loss and validation loss diagrams for the starch content. In the proposed model, the number of epochs was considered to be 200 for data training, and during the execution of the program, the number of epochs was executed until over-fit occurred. As soon as Over-fit happened, the training was stopped. According to Figure 3, it can be seen that after 15 epochs, training was stopped and the loss rate was 0%.

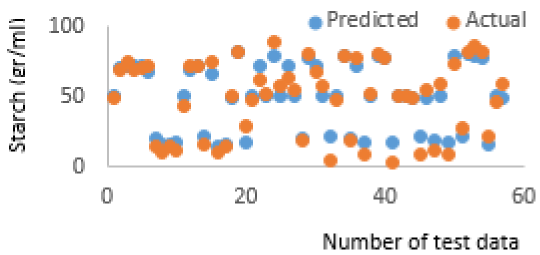

Figure 2. Scatter plot of the measured (actual) and predicted amount of starch for test data

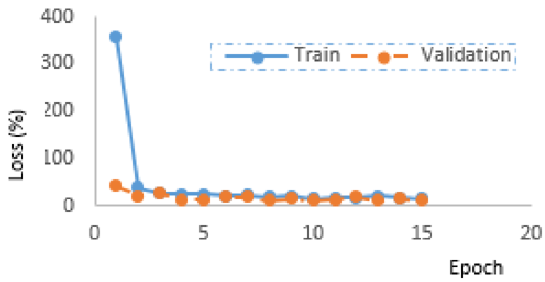

Figure 3. Train loss and validation loss diagram for predicting starch in apple

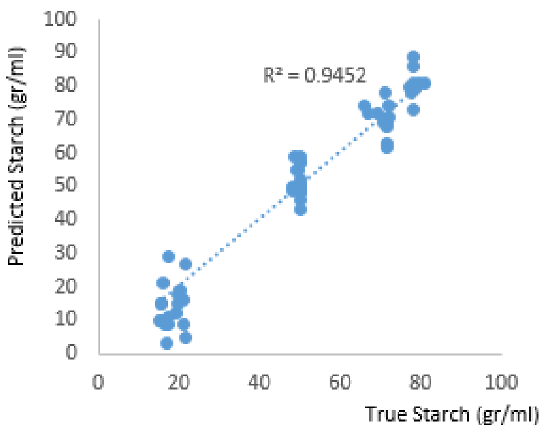

Figure 4. Regression plot between actual and estimated amount of starch for test data

3.2. Total Soluble Solid (TSS)

Table 3 shows the performance evaluation results of the prediction model for soluble solids. Considering that the values of Max error, Mean absolute error, Mean squared error, Median absolute error and Mean squared logarithmic error are close to zero and the values of variance score and coefficient of determination are close to one, it states that the proposed model for prediction of TSS has been successful.

Table 3. Evaluation of Performance of model for predicting TSS in apple

Var	0.916
R2	0.916
MaxE	0.907
MedE	0.239
MSLE	0.0008
MSE	0.122
MAE	0.284
Min	9.1
Max	13.8

Figure 4 shows the scatter plot of the actual and estimated TSS values for the 57 test data. In most samples, the two actual and estimated nitrogen values are close together that means the performance of proposed model for predicting TSS was promising. Figure 3 shows the train loss and validation loss diagrams for TSS. In the proposed model, the number of epochs for data training was considered to be 200. According to Figure 5, it can be seen that after 13 epochs, training has stopped and the loss rate is 0%.

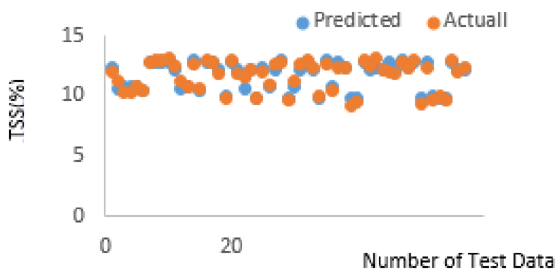

Figure 5. Scatter plot of measured and predicted TSS for test data in apple

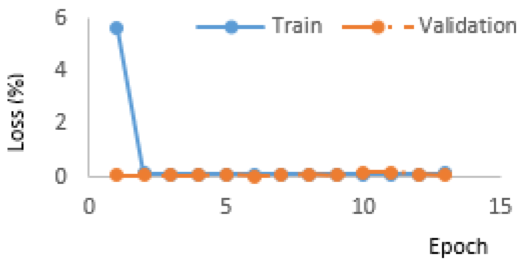

Figure 6: Train loss and validation loss diagram for predicting of TSS

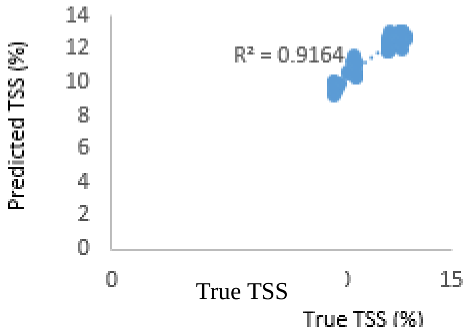

Figure7. Regression plot between of the actual and estimated value of the TSS for test data

3.3. Titratable Acidity (TA)

Table 4 shows the performance evaluation results of the predictive model for acidity. Considering that the values of Max error, Mean absolute error, Mean squared error, Median absolute error and Mean squared logarithmic error are close to zero and the values of variance score and coefficient of determination are close to one, it states that the proposed model for pre- Excess nitrogen nose has been 90% successful.

Table 6. Performance evaluation of model to predict TA

Var	0.842
R2	0.842
MaxE	1.264
MedE	0.421
MSLE	0.0074
MSE	0.280
MAE	0.424
Min	3.35
Max	8.33

Figure 7 shows the scatter plot of the actual and estimated TA for the test data. In most samples, the two actual and estimated values of TA are close to each other, indicating the ability of the proposed model. Figure 3 shows the train loss and validation loss diagrams for TA. In the proposed model, the number of epochs was considered to be 200 for data training, which according to Figure 8, it can be seen that after 15 epochs, training has stopped and the loss rate is 0%.

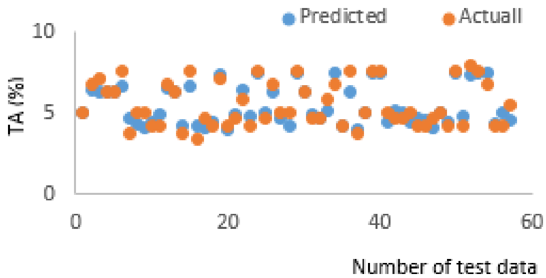

Figure 8: Scatter plot of measured and predicted TA for test data in apple

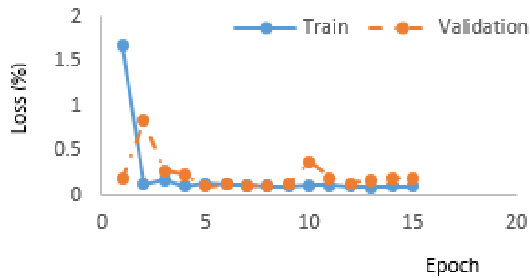

Figure 9. Train loss and validation loss diagram for predicting TA

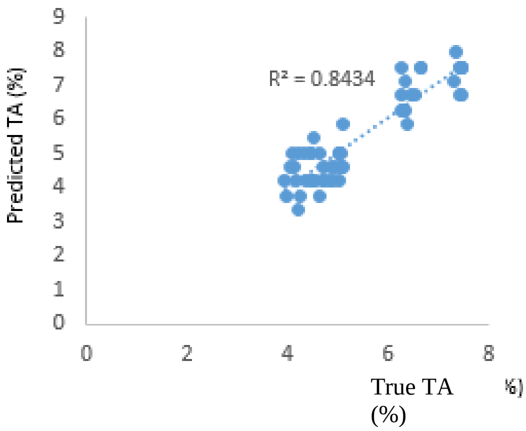

Figure 10. Regression plot between true value and estimated TA for test data

4. Conclusion

If apples are harvested earlier than the ripening stage, the fruits will be small and unfavorable in taste and color, they will soon wrinkle in storage and suffer severe weight loss. Such apples are very sensitive to storage anomalies such as bitterness, scale, etc. Also, in very late harvest, even under the best storage conditions, the fruits soften and flour sooner and their sensitivity to the complication of water soacking and tissue rupture increases. Therefore, the need for timely harvesting of apples should be considered.

Because CNN does not require manual extraction, it is one of the most popular types of neural networks. Automatic feature extraction makes work easy and timely especially in computer vision applications. Therefore, in this study, the chemical properties of starch, titratable acidity and total soluble solids as parameters involved in determining apple ripening, were predicted using the deep learning model that the results were as follows.

- The coefficient of determination and the mean squares error for the starch properties were 95.4% and 4.8.
- The coefficient of determination and the mean squares error for the properties of soluble solids was 91.6% and 0.284.
- The coefficient of determination and the mean squared error for the acidity characteristic were 84.2% and 0.424.

References

1. Pourdarbani, P., S., Sabzi, D. Kalantari, J.I. Arribas, Non-destructive visible and short-wave near-infrared spectroscopic data estimation of various physicochemical properties of Fuji apple (Malus pumila) fruits at different maturation stages, Chemometrics and Intelligent Laboratory Systems, **2020**. 206, 104147, https://doi.org/10.1016/j.chemolab.2020.104147.

2. Sabzi, S., R. Pourdarbani, D. Kalantari, T. Panagopoulos. Designing a Fruit Identification Algorithm in Orchard Conditions to Develop Robots Using Video Processing and Majority Voting Based on Hybrid Artificial Neural Network. Appl. Sci, **2020**. 10, 383.

3. Sharabiani, V.R., S. Sabzi, R. Pourdarbani, E. Solis-Carmona, M. Hernández-Hernández, J.L. Hernández-Hernández. Non-Destructive Prediction of Titratable Acidity and Taste Index Properties of Gala Apple Using Combination of Different Hybrids ANN and PLSR-Model Based Spectral Data. Plants**, 2020**, 9, 1718. https://doi.org/10.3390/plants9121718

4. Garg, D., S. Khan, A. Mansaf. Integrative Use of IoT and Deep Learning for Agricultural Applications. Springer, 2020. 12: p. 521– 531.

5. Jha, K., D. Aalap, P. Poojan. Intelligent Irrigation System Using Artificial Intelligence And Machine Learning: A Comprehensive Review. International Journal of Advanced Research, **2018**. 6 (10): p. 1493-1502.

6. Dhayabarani, R., K. Aravinth, M. Gowtham, D. Gowtham, F. Balakrishnan, Detection of Weed using Neural Networks, International. Journal of Engineering Research & Technology, 2018. 6 (8):p.1-5.

7. Patil, N., A. Rajab, W. Vaibhav, N. Deepali. Crop Disease Detection using Deep Convolutional Neural Networks. International. Journal of Engineering Research & Technology, 2019. 8(3): P. 3-7.

8. Ale L., S. Alaa, L. Longzhuang, W. Ye, Z. Ning, Deep Learning based Plant Disease Detection for Smart Agriculture. IEEE, 2019.

9. Mohanty S.P., D.P. Hughes, M. Salathe. Using deep learning for image-based plant disease detection. Frontiers in Plant Science, **2016**. 7:P. 36-47.

10. Manjula J., B. Ruth Ramya, K., Rama Rao, K.V.S.N. Swarna Kuchibhotla, P. Venkata Bala Kishore, S. Rahamathulla. Crop Yield Prediction Using Machine Learning. International Journal of Scientific & Technology Research, **2020**. 9(2.

11. Klompenburg, T., A. Kassahun, C. Cagatay. Crop yield prediction using machine learning: A systematic literature review. Computers and Electronics in Agriculture, 2020.

12. Rahnemoonfar M. and C. Sheppard. Deep Count: Fruit Counting Based on Deep Simulated Learning. Sensors, 2017. 17 (905): P. 58-64.
13. Zhang, Z., H. Liu, Z. Meng, J. Chen. Deep learning-based automatic recognition network of agricultural machinery images, Computers and Electronics in Agriculture, **2020**. 174, 105471
14. Perugachi-Diaz, Y., J.M. Tomczak, S. Bhulai. Deep learning for white cabbage seedling prediction, Computers and Electronics in Agriculture, **2021**. 184, 106059.
15. Esmail Karar M., F. Alsunaydi, S. Albusaymi, S. Alotaibi. A new mobile application of agricultural pests recognition using deep learning in cloud computing system, Alexandria Engineering Journal, 2021. 60 (5)): P. 4423-4432.
16. Franczyk, B., M. Hernes, A. Kozierkiewicz, A. Kozina, M. Pietranik, M., I. Roemer, M. Schieck. Deep learning for grape variety recognition, Procedia Computer Science, **2020**. 176: P. 1211-1220.
17. Yuan, Y., L. Chen, H. Wu, L. Lin. Advanced agricultural disease image recognition technologies: A review, Information Processing in Agriculture, **2021**,https://doi.org/10.1016/j.inpa.2021.01.003.
18. Ismail, A., A. Malik. Real-time visual inspection system for grading fruits using computer vision and deep learning techniques. Information Processing in Agriculture, **2021**. 1:P.36-47.
19. Tetila, E., B. Brandoli Machado, G. Astolfi, N. Alessandro de Souza Belete, W. Paraguassu Amorim, A. Railda Roel, H. Pistori. Detection and classification of soybean pests using deep learning with UAV images, Computers and Electronics in Agriculture, 2020. 179, 105836.
20. Rossel, R. A., A.B. McBratney. Diffuse reflectance spectroscopy as a tool for digital soil mapping. Digital Soil Mapping with Limited Data, **2008**.
21. Martínez-Valdivieso, D., Font, R., Blanco-Díaz, M. T., Moreno-Rojas, J. M., Gómez, P., Alonso-Moraga, Á., & Río-Celestino, M. D. Application of near-infrared reflectance spectroscopy for predicting carotenoid content in summer squash fruit. Computers and Electronics in Agriculture. 2014, 108 71–79. https://doi.org/10.1016/j.compag.2014.07.003
22. Douglas, F., C. Yang, S. Malcolm. The Maximum-Error Test. The Journal of the Astronautical Sciences, 2010. 55,
23. Sabzi, S, R. Pourdarbani, J.I. Arribas. A Computer Vision System for the Automatic Classification of Five Varieties of Tree Leaf Images. Computers 2020. 9 (1): https://doi.org/10.3390/computers9010006
24. M. Alibaba, R. Pourdarbani, M. H.Khoshgoftar Manesh, G. V. Ochoa, J. D. Forero, Thermodynamic, exergo-economic and exergo-environmental analysis of hybrid geothermal-solar power plant based on ORC cycle using emergy concept. Heliyon, 2020. 6 (4), e03758.
25. Mosavi A., S., Ardabili, A.R. Várkonyi-Kóczy. List of Deep Learning Models. Networks and Systems, 2020. 101. https://doi.org/10.1007/978-3-030-36841-8_20
26. Ardabili S., A. Mosavi, A. Varkonyi-Koczy. Advances in Machine Learning Modeling Reviewing Hybrid and Ensemble Methods. Engineering for Sustainable Future, 2020: p.215-227.10.1007/978-3-030-36841-8_21.
27. Ardabili, S., A. Mosavi, M. Dehghani, A. Varkonyi-Koczy, Annamaria. Deep Learning and Machine Learning in Hydrological Processes, Climate Change and Earth Systems: A Systematic Review. 2019. 10.20944/preprints201908.0166.v1.
28. Kristof, W. Estimation of true score and error variance for tests under various equivalence assumptions. Psychometrika, 1969. 34, 489–507.
29. Sabzi, S., R. Pourdarbani, M.H. Rohban, G. García-Mateos, J. Paliwal, J.M. Molina-Martínez. Early Detection of Excess Nitrogen Consumption in Cucumber Plants Using Hyperspectral Imaging Based on Hybrid Neural Networks and the

ImperialistCompetitiveAlgorithm. Agronomy 2021. 11, 575.
30. Wang, W. & Y. Lu. Analysis of the Mean Absolute Error (MAE) and the Root Mean Square Error (RMSE) in Assessing Rounding Model. IOP Conference Series. Materials Science and Engineering. 324. 012049. 10.1088/1757-899X/324/1/012049
31. Jeong, J. H., J.G. Woo, L. Park. Machine Learning Methodology for Management of Shipbuilding Master Data, International Journal of Naval Architecture and Ocean Engineering, 2020. 12: P. 428-439