Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi
Sayı 7, S. 65-70, 6, 2023

Araştırma Makalesi

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN: 2980-0811

International Journal of Advanced Natural Sciences and Engineering Researches
Volume 7, pp. 65-70, 6, 2023

Research Article

Classification of Plant Diseases with Machine Learning

Elif Akarsu*, Tevhit Karacali[2], İ. Yücel Özbek[3]

[1]*Electrical-Electronics Engineering /Faculty of Engineering, Ataturk University, Turkey*
[2] *Electrical-Electronics Engineering /Faculty of Engineering, Ataturk University, Turkey*
[3]*Electrical-Electronics Engineering /Faculty of Engineering, Ataturk University, Turkey*

[*]*(elif.akarsu@atauni.edu.tr) Email of the corresponding author*

(Received: 03 July 2023, Accepted: 24 July 2023)

(5th International Conference on Applied Engineering and Natural Sciences ICAENS 2023, July 10 - 12, 2023)

ATIF/REFERENCE: Akarsu, E., Karacali, T. & Özbek, İ. Y. (2023). Classification of Plant Diseases with Machine Learning. *International Journal of Advanced Natural Sciences and Engineering Researches*, 7(6), 65-70.

Abstract – Deep learning applications are also of great importance in the field of agriculture. It is an issue that should be examined in terms of what the diagnosis of the disease is and its treatment in the disease of plants. A study on the detection of plant diseases was desired. This study is intended to be done using CNN and Alexnet methods. It was aimed to observe the success rates of two different methods. It is aimed to examine the effect of classification with and without feature extraction. In other words, it is aimed to examine the effect of deep layers that alexnet has. In this study, the solution of a 38-class problem was dealt with. 38-class test and 38-class train data are available. Each of the 38 classes contains approximately 1500 to 2000 pictures. And as a result of the study, 68500 of approximately 70000 pictures consist of healthy leaves and 1500 of them are pictures of sick leaves. this whole data set is separated and classified as 80% and 20%. While the accuracy obtained with CNN was 98.27%, the accuracy obtained with Alexnet was 98.49%. In conclusion it was seen that the use of Alexnet increased the accuracy rate.

Keywords – Artificial Intelligence, Machine Learning, Plant Disease, Agriculture, Alexnet

I. INTRODUCTION

The use of machine learning (ML) models is becoming more common in industrial IoT applications, especially for use cases such as anomaly detection and predictive maintenance. Machine learning models are often trained in the cloud using iterative processing on large stored datasets and powerful servers. When it comes to running the model on new live data, called inference, the cloud may not be the best place. In many use cases, the most natural place to run machine learning models is at the endpoint, close to the data source [1]. Artificial intelligence, hardware related to intelligence development, examining computer models, after hardware formal, artificial intelligence system is a research area. Since artificial intelligence has different application areas, there are many definitions depending on the application area.

A machine learning algorithm is a software-implemented, statistics-based computational method that can discover hidden, non-obvious patterns in a dataset. [5] And moreover to make reliable statistical predictions about similar new data. The ability of machine learning to automatically identify patterns in data is particularly important when the expert knowledge is incomplete or inaccurate.[2] when the amount of available data is too large to be handled manually, or when there are exceptions to the general cases. [4] Deep learning techniques can be applied to help the farmer detect pests, diseases and weeds and are an

early stage and corrective measure that can help prevent loss of production [9]. Agriculture, one of the application areas of AI, hosts my column today. By replacing human power in production; New technologies that cultivate, sow, irrigate, harvest, analyze the soil and plant status, add new ones to traditional agricultural activities every day. Artificial intelligence, which adapts quickly to agricultural processes, especially in developed countries, attracts attention from most producers. In general, the main goal is to classify images as identification of plant pests, diseases and weeds, improving classification accuracy [10].

Determining the future of agriculture depends on the solutions to be produced against the problems in this field. Artificial intelligence applications used to support agricultural production enable more efficient production by detecting problems in production. With drones, conditions such as the development status of plants, the amount of pesticides to be used, and disease threats are determined; methods of dealing with any possible negativity can be determined. Thus, efficiency-oriented production is supported, saving time and labor. Considering the fact that the tissue culture optimization period for any plant in plant tissue culture laboratories today sometimes exceeds years, it cannot be denied that it is a very beneficial application in terms of time saving, labor, resource saving and cost. Within the framework of traditional agricultural activities, the factors that cast a shadow on plant production are illuminated with artificial intelligence technology and the sustainability of agriculture is shed light on. When you suspect something is wrong with the growth of your plants, the first step should be to characterize the symptoms you see. Symptoms can be local and limited, such as leaf spots, or general and systematic, such as poor growth or wilted leaves.[1] For example, growers have experienced poor growth as a general plant response to a weakened root system that is not working at full capacity. In such a case, examination of the root system may reveal that the roots are brown in color rather than white and plump, and are about to die. Thus, it will be understood that the plant has a disease localized in the roots and limited but affecting the whole plant. Another thing to consider about plant diseases is whether all or most plants in the same area show the same symptoms. If not all plants show the same symptoms, it is nutritional, climate and environmental differences between healthy and sick plants that need to be considered. If the symptoms you observe are on young sprouts that begin to grow but later die, or on seeds that have never germinated from the start, it indicates an infection attacking the seeds or young sprouts. If the plant withered suddenly and died after growing for a while, it is possible to talk about pale or rotten root diseases. The green parts of the plant, such as leaves, may exhibit color changes or deformations from time to time. This could mean that the pathogen is involved in the plant's metabolism and the disease involved is caused by a virus.[1] The symptoms of viral diseases are often confused with the symptoms of diseases caused by nutritional deficiencies. If the individual plant is suddenly sick and bent its neck despite the same and equal irrigation of the soil, it is possible to say that the plant has a disease originating from the soil. This type of disease, which is caused by the weaknesses caused by fungi or bacteria in the soil, can grow and spread within the vascular system that provides water and nutrient delivery after entering the roots. The disease, which starts to spread in the vascular system, prevents the transfer of water and nutrients from the soil to the plant after a point and weakens the plant. For an infectious plant disease to occur, three essential factors are required: the plant susceptible to the disease, a pathogen capable of causing the disease, and an environment conducive to the disease. If one of these three basic factors is missing, a contagious plant disease will not occur. On the other hand, if the plant shows symptoms such as poor growth or yellow and rotting leaves without the presence of any pathogen, the presence of a non-infectious plant disease is highly likely. [1]

II. Materials and Method

2.1. Multilayer sensor

One of the most basic deep learning components, the multilayer perceptron, creates digital neural networks that mimic the way the human brain works. Multilayer perceptrons have 3 basic layers: input, hidden and output to contribute to deep learning results. According to experts, multilayer perceptrons can be trained to implement any nonlinear input-output mapping. They also highlight their interpolation capabilities by generalizing on sparse data areas. deep webs of belief.Deep Belief Networks is a model that consists

of multiple layers stacked on top of each other. In this model, each layer captures increasingly complex features from the previous layer. In deep belief networks, each layer of the network can be trained independently. Deep belief networks are an important type of artificial intelligence that can be used in many different applications because they have high-capacity learning skills that focus on the decision-making process.

2.2. Convolutional Neural Networks

Convolutional Neural Networks, which are deep learning models used for image recognition and identification, have a special structure that enables the detection of independent relationships between features. Technically moving from the pixel arrangement of the image, convolutional neural networks use it to capture the style of similar images and match image components. Although not very interesting in theory, convolutional neural networks can give dizzying results in practice.

2.3. Iterative Neural Networks

Recurrent Neural Networks are distinguished by making accurate predictions. For example, it can generate text by guessing what should come after a word or letter in a sentence. Such deep learning models can write poetry or deliver a text output that can successfully pass the Turing test. Iterative neural network models may have several caveats to avoid confusing users. ChatGPT, for example, reminds us almost routinely that it is an artificial intelligence, not a human, so it cannot understand human feelings such as emotions and value judgments. Multi-concept perceptron, deep belief networks, evolutionary neural networks and iterative neural networks are the most preferred deep learning models today. Each of these models is emerging with different artificial intelligence applications. The close version, whether your learning is an acquisition or an image we see is man-made, we will definitely have a hard time deciding. Moreover, these maturing deep learning models are only the first step positions of the point that will exist. Near release deep learning will be used more and more for generative design tools. Image search, photo real processing, file enhancement, and various other deep learning applications will continue to evolve to achieve the best results. There will be significant changes in the market for the development of deep learning models, with expenses such as NVIDIA, which has invested heavily in artificial intelligence hardware and is remarkable for its innovative research.

Due to global climate changes, there are some changes in the population and spread areas of diseases and pests seen in sugar beet. Especially in our country, there has been an increase in disease rates in recent years. Therefore, it is necessary to prevent these diseases. Plant protection; means the methods applied to reduce their density below the economic damage level in order to prevent and reduce the damage caused by disease factors, pests, weeds and other animal organisms that may cause damage to plants in agricultural areas and their products. It is possible to obtain higher quality and higher products from the unit area and to increase the level of economic income with plant protection. So, a study on the detection of plant diseases was desired. This study is intended to be done using CNN and Alexnet methods. In the first method, no features were extracted in any way, the images were processed and sent for classification. In the second, features were obtained using the Alexnet algorithm and sent for classification. Alexnet has 25 layers and 8 deep layers. [7] and used in many applications. and very good results are obtained. It was aimed to observe the success rates of two different methods. For this study, the dataset named "New Plant Dataset" available on Kaggle was used. This dataset consists of approximately 87,000 RGB images of healthy and diseased plant leaves from 38 different classes. [6] At this point, regardless of the size of the images, we can re-analyze them as we want. The initial size of the images is 256*256. In this study, the initial dimensions for both methods were arranged as 227*227. [8] The first main reason for this is that the Alexnet algorithm does not accept input other than this measure. another is based on keeping the measure constant in order to compare with these measures.

Fig. 1 Images of leafs

There are also about 2000 pictures of trains in each class. 80% of these pictures are reserved in the code as train and 20% as test. Approximately 14000 of the resulting images are reserved as test data. 10% of this is reserved as validation data to test accuracy. [3]. 38-class test 38-class train and up to 10% of test data is reserved for validation to check accuracy.

Table 1. Number of leaves by health status

Class	Number of İmages
sick leaf pictures	68500 images
healthy leaf images	1500 images

III. RESULTS

Matrices are a visual representation of how accurately the classification has been made. that is, evidence for classification. These matrices are analyzed and many parameters such as accuracy and error rates are obtained. We made this diagnosis with a convolutional neural network. that is, we focused on solving a 38-class problem. While the accuracy obtained with CNN was 98.27%, the accuracy obtained with Alexnet was 98.49%. The higher the number of data in the diagonal sections of the pictures, the more accurate the definition is. When the data in other sections were examined, a homogeneous distribution was observed. homogeneous distribution in this way is a condition that should happen. Otherwise we can talk about memorizing or not getting the exact attributes.

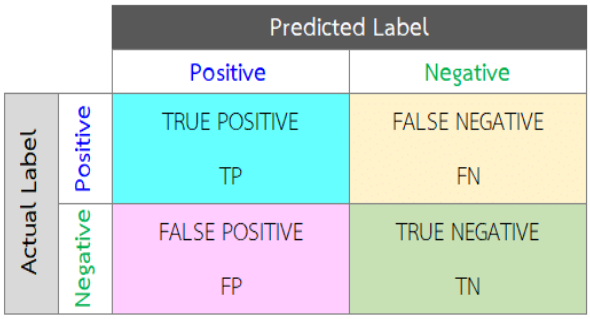

Fig. 2 content description of the confusion matrix

The accuracy rate is obtained by using the true positive and true negative value in the diagonal.[11]

Accuracy= True Positive/ True Positive + True Negative

We reach the error rate by using false positive and false negative values outside the diagonals.[12]

Error Rate= False Positive+ False Negative/ All of Samples

Presicion; It shows how many of the values we estimated as positive are actually positive. [13]

Presicion= True Positive/ True Positive + False Positive

Sensitivity (Recall) is a metric that shows how much of the operations we need to estimate as Positive, we estimate as Positive. [14]

Recall =True Positive/ True Positive + False Negative [15]

Table 2. Accuracy of Methods

Method	Accuracy	False	presicion	recall
Alexnet	98.49%	%1.51	% 96.4	%98.99
CNN	98.27%	%1.73	% 95.2	%98.76

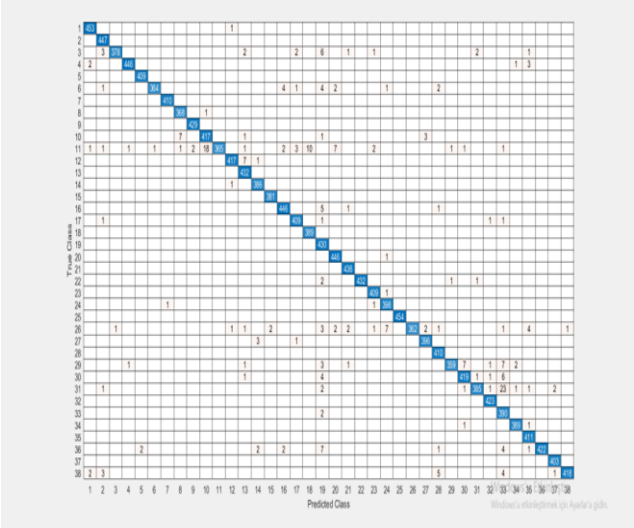

Fig. 3 CNN confusion matrix

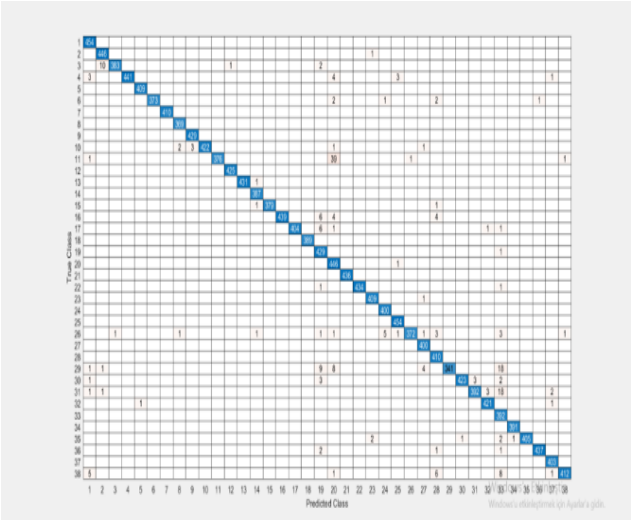

Fig.4 Alexnet confusion matrix

IV. DISCUSSION

The disease occurs due to excess moisture of the plant and is mostly seen in roses. Black spots appear first on the leaves of the infected plant. If not intervened, the leaves fall off partially or completely in a short time. The plant weakens and starts not to flower in the next period. Treatment method: First of all, you need to find out how much water the plant needs and follow a watering schedule. The point to be considered here is not to water the plant too much without drying it too much. Clean leaves where disease is seen on the plant. Well ventilating the room where the flower is located will also work, so you can lower the humidity in the room as well. However, be careful not to leave your flower directly in the air stream. Obtain the pesticide sprays specially developed for this disease from a greenhouse authorized to sell pesticides and apply it to your flower in the recommended amount. Plant diseases are of great importance for agriculture. and artificial intelligence and machine learning provide significant convenience at the point of detection of these diseases. Classification of a plant file with 38 classes was made and two different results were obtained. A total of 70000 images have been classified. In the Alexnet algorithm, the features are taken from the fully connected layer. 1000 of the received features for each image from the last fully connected layer are sent to classification. feature matrices of the classified features are obtained. And the confusion matrix of these classes was obtained. The more data in the diagonals of the obtained confusion matrix, the more accurate the classification is. and at this point it is of great importance. This work is of great importance in the field of artificial intelligence and machine learning. It has been investigated how effective the Alexnet toolbox is on the convolutional neural network algorithm. It has been observed that Alexnet toolbox increases the accuracy rate in disease detection.

V. CONCLUSION

Plants can be affected by various diseases and pests. Here are some of the common plant diseases and pests: Plant Diseases: Fungal Diseases: For example, fungal diseases such as powdery mildew, mildew, black spot, rust are common in plants. These diseases usually manifest as spots, rot or mold on leaves. Bacterial Diseases: For example, bacterial diseases such as febrile illness, blight can cause damage to plants. Bacterial diseases are usually manifested by lesions, bruises, or tissue deterioration. Viral Diseases: There are several types of viruses that affect plants. Viral diseases can manifest themselves in plants with symptoms such as color changes, stripes, mosaic patterns or growth disorders on leaves. Classification was done in Matlab. and adding only the alexnet command to the same parameters increased the success rate in the classification. This study, which includes 38 classes of plant dataset, has a very rich content. A wide variety of images of both healthy and diseased leaves were processed. Having a

detailed datset means collecting more features, which results in more accurate, clear and stable structures. On the other hand, results were obtained with only image processing without removing the features. It was desired to observe the effect of impact both along the feature and without subtraction. Positive results of exclusion after feature extraction were observed in this study.

REFERENCES

[1] Capizzi, A., Distefano, S., Mazzara, M. (2020). From DevOps to DevDataOps: Data Management in DevOps Processes. In: Bruel, JM., Mazzara, M., Meyer, B. (eds) Software Engineering Aspects of Continuous Development and New Paradigms of Software Production and Deployment. DEVOPS 2019. Lecture Notes in Computer Science, vol 12055. Springer, Cham. https://doi.org/10.1007/978-3-030-39306-9_4

[2] Yip KY, Cheng C, Gerstein M. "Machine learning and genome annotation: a match meant to be?" Genome Biol. 2013;14(5):205.

[3] (2018) The Kaggle website. [Online]. Available: https://www.kaggle.com/datasets/vipoooool/new-plant-diseases-dataset

[4] Yip, K.Y., Cheng, C. & Gerstein, M. Machine learning and genome annotation: a match meant to be?. Genome Biol 14, 205 (2013). https://doi.org/10.1186/gb-2013-14-5-205

[5] Chicco, D. Ten quick tips for machine learning in computational biology. BioData Mining 10, 35 (2017). https://doi.org/10.1186/s13040-017-0155-3

[6] Kewat, K. (2023). Plant Disease Classification Using Alex Net.. https://doi.org/10.21203/rs.3.rs-2612739/v1

[7] Krizhevsky, Alex, Ilya Sutskever, and Geoffrey E. Hinton. "ImageNet Classification with Deep Convolutional Neural Networks." Advances in neural information processing systems. 2012.

[8] Bungărdean RM, Şerbănescu MS, Streba CT, Crişan M. Deep learning with transfer learning in pathology. Case study: classification of basal cell carcinoma. Rom J Morphol Embryol. 2021 Oct-Dec;62(4):1017-1028. doi: 10.47162/RJME.62.4.14. PMID: 35673821; PMCID: PMC9289702.

[9] F.LiuandZ.Xiao,"Disease Spots Identification of Potato Leaves in Hyperspectral Based on Locally Adaptive 1D-CNN," 2020 IEEE International Conference on Artificial Intelligence and Computer Applications (ICAICA),2020, pp.355-358, doi:10.1109/ ICAICA 50127.2020.9182577

[10] M.A.Khanetal.,"An Optimized Method for Segmentation and Classification of Apple Diseases Based on Strong Correlation and Genetic Algorithm Based Feature Selection, "inIEEEAccess,vol.7,pp.4626146277,2019,doi:10.1109/ACCESS.2019.2908040

[11] M. U. Mojumdar, S. R. Haider Noori and F. Faisal, "Malabar Nightshade Disease Detection and Recognition Using Transfer Learning," 2021 IEEE 6th International Conference on Computing, Communication and Automation (ICCCA), Arad, Romania, 2021, pp. 637-641, doi: 10.1109/ICCCA52192.2021.9666348.

[12] Saeed Abu-Nimeh and Dario Nappa and Xinlei Wang and Suku Nair, Hardening Email Security via Bayesian Additive Regression Trees, Machine Learning, IntechOpen, Rijeka, 2009, https://doi.org/10.5772/6554

[13] Özkaraca O, Bağrıaçık Oİ, Gürüler H, Khan F, Hussain J, Khan J, Laila UE. Multiple Brain Tumor Classification with Dense CNN Architecture Using Brain MRI Images. Life (Basel). 2023 Jan 28;13(2):349. doi: 10.3390/life13020349. PMID: 36836705; PMCID: PMC9964555.

[14] Sinan Keskin and Adnan Yazıcı, FSOLAP: A fuzzy logic-based spatial OLAP framework for effective predictive analytics, Expert Systems with Applications, 2023, https://doi.org/10.1016/j.eswa.2022.118961

[15] Poux, Florent, Romain Neuville, Gilles-Antoine Nys, and Roland Billen. 2018. "3D Point Cloud Semantic Modelling: Integrated Framework for Indoor Spaces and Furniture" *Remote Sensing* 10, no. 9: 1412. https://doi.org/10.3390/rs10091412