Hittite Journal of Science and Engineering, 2022, 9 (3) 159–166
ISSN NUMBER: 2148–4171
DOI: 10.17350/HJSE19030000267

Powdery Mildew Detection in Hazelnut with Deep Learning

Tulin Boyar[1] **Kazim Yildiz**[2]
[1]Marmara University, Institute of Pure and Applied Sciences, Department of Computer Engineering, Istanbul, Turkey
[2]Marmara University, Faculty of Technology, Department of Computer Engineering, Istanbul, Turkey

Article History:
Received: 2022/05/11
Accepted: 2022/09/16
Online: 2022/09/28

Correspondence to*: Tulin BOYAR, Marmara University, Computer Engineering, Istanbul, TURKEY*
E-Mail: tulinboyar@marun.edu.tr
Phone: +90 535 038 19 36

ABSTRACT

Hazelnut cultivation is widely practised in our country. One of the major problems in hazelnut cultivation is powdery mildew disease on hazelnut leaves. In this study, the early detection of powdery mildew disease with the YOLO model based on machine learning was tested on a unique data set. Object detection on the image, which is widely applied in the detection of plant diseases, has been applied for the detection of powdery mildew diseases. According to the results obtained, it has been seen that powdery mildew disease can be detected on the image. Using YOLOv5, diseased areas were detected with approximately 90% accuracy in diseased leaf images. Multiple leaves in one image were detected with approximately 85% accuracy in detecting healthy areas using images with complex backgrounds. The model, which has been used in different studies for the detection of disease in plant leaves, also gave effective results in the detection of powdery mildew disease in hazelnut leaves. Early detection of powdery mildew with a method based on machine learning will stop the possible spread of disease. It will increase the efficiency of hazelnut production by preventing the damage of hazelnut producers.

Keywords:
Powdery mildew disease in hazelnut; Powdery mildew disease detection; Object detection with Yolo; Leaf disease; Leaf disease detection.

INTRODUCTION

In this paper, we focus on the detection of diseases in hazelnut plants using a deep learning architecture such as YOLOv5. Detection of diseases in plants in large agricultural areas is a very challenging task. Manual identification is difficult as it would take a lot of time to examine the leaves of the field. Also, this method has low accuracy and takes a lot of time to detects diseased areas. So there must be a model to detect diseases with greater accuracy. Our paper explains a model that can support the farmers to identify the disease at a low cost. One of the most common diseases on hazelnut leaves is powdery mildew. It is a disease that mainly affects hazelnut leaves. For the purposes we mentioned in our article, we focused on an object detection model that can detect powdery mildew at a low cost and early. The experimental results show that farmers can see the affected area on the images. With the proposed deep learning architecture, powdery mildew disease in hazelnut can be identified as soon as it occurs. Thus, it will be a very beneficial process contribution for framers by taking appropriate measures to prevent the spread of the disease. The aim of this study is to find a method to detect powdery mildew disease in the plant from pictures obtained from hazelnut leaves.

In this study first a literature review part is given. In material method section, a deep learning based method for powdery mildew detection is introduced and the dataset, theoretical background and methodology is given. Experimental results are proposed in discussion section and finally conclusion part is given.

LITERATURE REVIEW

Turkey is the leading country in the world hazelnut production and export a production of 675 thousand tons. Approximately 70% of the world hazelnut production Turkey is followed by Italy, Caucasian Countries (Azerbaijan+Georgia), Iran, USA and Spain, respectively[1]. In Turkey, hazelnut is the product that provides the highest export income among agricultural products and it is predicted that it will maintain this position for a long time [2]. Negative weather conditions seen at intervals, losses due to diseases and pests adversely affect many producers and the country's economy. Powdery mildew, which is known as the most important of these diseases, is a long-known and common disease in our country and in the world.

Phyllactinia guttata and Erysiphe cause powdery mildew disease in hazelnut which is generally seen on the lower surface of the host leaves, and is partially seen on the upper surface in cases of high density infestations. Conidia are unicellular, club-shaped, or sometimes rhomboid. Cleystothecium is round and broad and have 3-15 attachment arms. There is swelling at the base of the retainer arms, which is quite distinguishable. The mycelium of Erysiphe sp. develops on both surfaces of the leaves. Conidia are unicellular, oval, ellipsoid or barrel-shaped. The clestothecium are rounded and the attachment arms show dichotomous branching at the ends[1] .

It is inevitable to experience significant losses if studies are not carried out to determine the diseases that harm crop plants in agriculture. If the disease-causing factors are in the form of an epidemic, all the plants in the area become sick and most of them die. In order to prevent losses, it is necessary to fight against disease factors. In order to determine the situation in the region, the presence and diagnosis of the factor to be combated should be determined [3].

Studies are carried out to reduce the damage caused by plant diseases to agricultural products. Machine learning algorithms are widely used for early detection and prevention of plant diseases. Mohammadpoor and friends proposed a method for detecting viruses on grape leaves. They identified the diseased parts on the leaves and classified them with support vector machine (SVM). The K-fold cross validation method was applied with k=3 and k=5 values are used to increase the accuracy of the system. According to the experimental results, the average accuracy of the proposed method was observed about 98.6% [4]. Liang and friends proposed a convolutional neural network(CNN) based system to detect blast disease in rice. Training process was provided with a dataset containing 2906 samples without disease and 5808 samples with disease. The experimental results show that the proposed model has higher discriminating ability and effective feature extraction potential than traditional approaches [5].

Zhang and friends proposed an algorithm to identify cucumber disease. A global pooling dilated convolutional neural network (GPDCNN)-AlexNet-based approach. The model is trained on cucumber leaf diseases and able to detect cucumber leaf diseases with an accuracy of 95.65% [6]. Wagh and friends have proposed a system which can automatically detect grape diseases. The system was able to detect five types of diseases: downy mildew, bacterial spots, powdery mildew and anthracnose. In the study, AlexNet architecture was used which had previously performed feature extraction and model training for leaf images. Experimental results show that the model could classify five diseases [7].

Petrellis carried out a study to detect plant diseases with image processing techniques, using a smartphone, with computing power and high resolution camera support [8]. Mohanty developed a deep learning and computer vision based approach. The dataset which has totally 54,306 diseased and healthy plant leaves presented publicly. A deep neural network was trained to detect 14 plant species and 26 diseases. When the experimental results were examined, it was seen that the training model achieved 99.35% accuracy[9]. Vijai and friends proposed an algorithm for image segmentation to the automatic detection and classification of plant leaf diseases. At the first stage, the accuracy is 85.56% in the classification made with Minimum Distance Criterion and K-Means Clustering. When these classification techniques are used with the proposed algorithm, the accuracy rate increased to 93.63%. The accuracy of the proposed algorithm is 95.71% when classification is made with the SVM technique [10].

The YOLOv5 model has been used to detect postnatal defects in kiwi fruit. A fast and highly accurate defect detection model has been developed to detect minor defects. The developed YOLOv5 model reached 94.7% accuracy[11]. The improved YOLOv5 model has been proposed for the detection of tomato plant diseases. The dataset was created by collecting disease images with a mobile phone. A squeeze and excitation(SE) module has been added to the YOLOv5 model to extract key features. The result was compared with R-CNN and Faster R-CNN models and SE-YOLOv5 gave better results with 94.10% accuracy [13]. The YOLOv5 model was proposed by Midhun and Yamuna for the detection of bacterial disease in bell pepper. It is aimed to detect the disease as soon as it occurs and to prevent its spread. The model was trained with images taken randomly from various parts of the farm and tested with mobile phone images. Small disease spots were detected quickly and with high accuracy [14].

MATERIAL AND METHODS

The operations performed in this study are shown in Fig. 1. During the data collection phase, powdery mildew disease images were collected from various internet sites. The appropriate images were selected from the images collected during the preprocessing stage of the dataset, all images were sized, labeling was done by selecting the

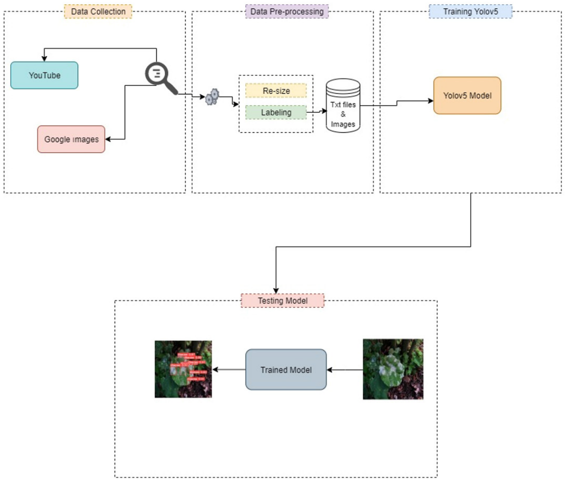
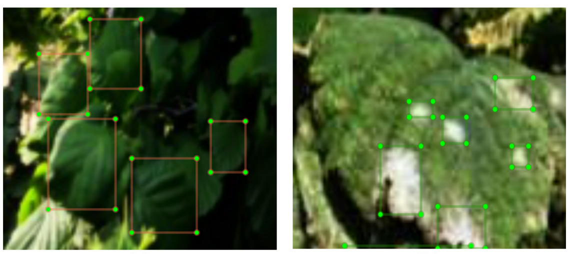


Figure 1. General diagram of the proposed study.

hazelnut leaf and disease section, and txt files were created. The trained model was tested in the testing phase and the results were obtained.

Dataset Preprocessing and Preparation

Since the aim of the study was to detect powdery mildew disease in hazelnuts through the image, images of healthy and diseased hazelnut leaves were needed. In this study, a data set was created with the researches made on the internet and the images collected[12]. In order to detect powdery mildew disease in hazelnut with machine learning, diseased and healthy leaf images of hazelnut leaves were used. Images were created with images collected from Google[13] and Youtube [14] videos. The results obtained with Google were analyzed one by one, the image quality, qualities and uniqueness in the data set were evaluated and the images that would form the basis for the data set were obtained. The dataset was augmented using the hazelnut fields videos on Youtube. Every frame of these downloaded videos has been converted into a picture and the number of images for the data set has been increased. Diseased leaf images obtained in the study of powdery mildew disease by Sezer and friends constituted another source for dataset [15]



(a) (b)

Figure 2. (a) Labeling of healthy leaves, (b) Labeling of diseased parts on the leaf can be created by selecting the desired parts in the squares one by one..

Generating the Data Set for the YOLOv5 Model

YOLO[16] is a neural network algorithm which is very fast. It gives successful results when used to detect real-time objects. It trains the network via images and a ".txt" file where the images are mapped with numerical values. In order to prepare the dataset in a format suitable for the YOLO model, all images were first rescaled to 224x224. The images and ".txt" file are created by labeling the scaled images one by one using the labeling program.

As seen in Fig. 2 (a) and (b), the height and width are marked to fully include the object to be labeled which is very important for the correct and complete learning of the network. While the YOLO model accesses the labeled regions, it reads from the file with the ".txt" extension containing the numerical values of the image.

Fig. 3 (a) and (b) shows the ".txt" files which is containing a series of numerical information are obtained as a result of tagging the images. Each file belongs to an image and each line contains information about the labeling done on the image. The first information is the label, that is, the class number, and indicates which object is labeled. The remaining information is the location and size information of the labeled object.

As seen in Table 1, the object represented by "0" is "disease" and the object represented by "1" is "leaf". The numerically represented values in the file are converted to the tag name during the object detection phase.

After the preliminary preparations, the data set has reached its final form. The dataset consists of a total of 424 images, 224 of which are "disease" and 200 are "leaf". As can be seen in Figure 3 (a) and (b), the images of the classes "disease" and "leaf" are balanced in the data set. The class distribution of the dataset is one of the important factors affecting the learning of the network.

Table 1. Numeric value and meaning of the object represented in the file created for the tagged images.

	Tag No (Numeric)	Tag Name (String)
1	0	disease
2	1	leaf

```
1 0.311719 0.532870 0.042188 0.056481
1 0.384375 0.264815 0.036458 0.051852
1 0.400781 0.327315 0.034896 0.060185
1 0.521875 0.469907 0.030208 0.030556
1 0.399740 0.183796 0.021354 0.056481
```

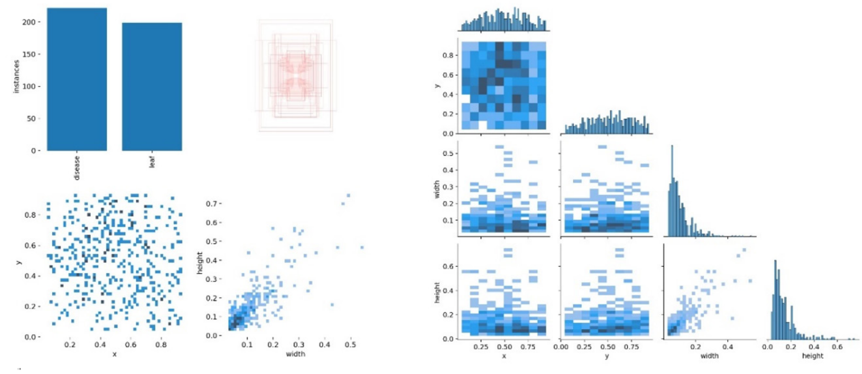
(a)

```
0 0.303571 0.287946 0.053571 0.075893
0 0.609375 0.794643 0.058036 0.089286
0 0.462054 0.667411 0.066964 0.066964
0 0.225446 0.430804 0.040179 0.093750
0 0.491071 0.533482 0.062500 0.058036
```

(b)

Figure 3. ".txt" file content obtained as a result of labeling healthy leaves, (b) ".txt" file content obtained as a result of labeling diseased parts on the leaf..

T. Boyar and K. Yildiz/ Hittite J Sci Eng, 2022, 9 (3) 159–166



(a) **(b)**

Figure 4. (a) Numerical information of "disease" and "leaf" image samples, (b) Correlogram graph of tagged samples.

Neural Network

The mathematical structure that can analyze the connections between input and output data (linear or non-linear) with various learning algorithms is defined as artificial neural networks(ANN) [17]. It consist of artificial neurons that imitate the biological neuron abilities that exist in the human brain. Artificial neurons are the smallest unit that can store and process information[18]. ANN consist of an input layer, one or more hidden layers, and an output layer. Each layer can contain a different number of artificial neurons[19].

Deep Learning

Deep learning, which is a sub-branch of machine learning, allows working on multidimensional data by using ANN algorithms in multi-layer artificial neural network architectures. With deep learning, many complex problems have been studied and successful results have been obtained. In the field of computer vision; face detection, object detection, license plate recognition, lane detection, etc. Solutions have been proposed to the problems with a deep learning approach and the proposed methods have been successful. It is a suitable area for object detection and image processing[20].

YOLO Model

It was aimed to detect the disease through the image, while choosing the method to be used for this purpose, similar studies were examined and effective methods were searched. Current algorithms used for image recognition include the R-CNN and YOLO frameworks. R-CNN is superior for high accuracy object detection but its speed is lower. For this reason, it cannot perform well in real time practical scenarios. The algorithms used in the YOLO series for real time object detection use a regression method that solves the speed problem and learns the determined properties of the object to be detected. The YOLO algorithm series uses a single-stage neural network to directly complete the positioning and classification of the object to be detected [21, 22]. It splits images with objects to be detected in a grid view in SxS ratio. Each grid square has a different detection task. The network structure consists of two full link layers and 24 convolution layers. After the full link layer, its tensor SxSx(Bx5+C) is the output. B refers to the number of predicted targets in each grid, C indicates the number of categories.

The final detection result is obtained by retrieving the detection box position and inferring the category probability

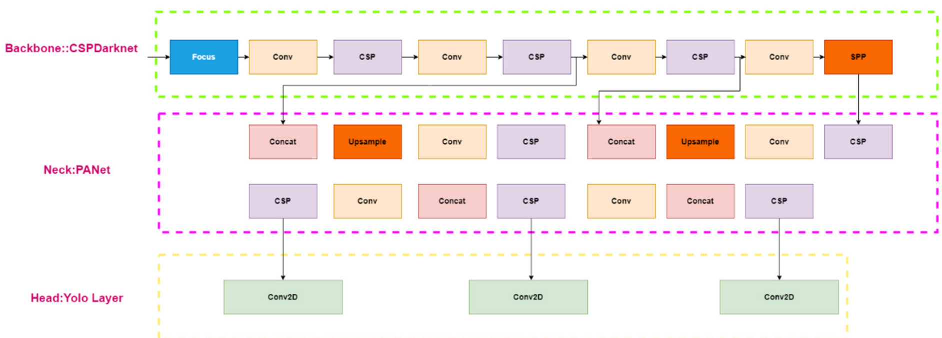


Figure 5. Yolov5 Model Architecture [27].

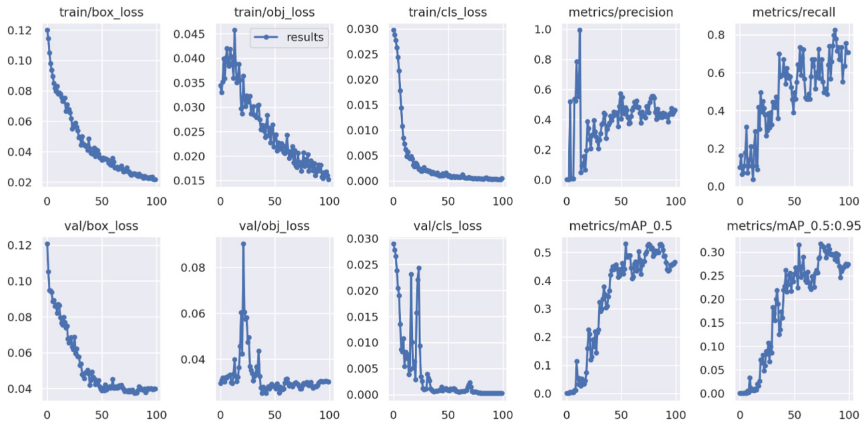


Figure 6. Graphs of loss values, recall values, and mean precision (mAP) during training periods for the training and validation set.

of the tensor data. It is fast at detecting objects however not successful in detecting small targets. This is because without the grid system, there are several targets on the same grid. This has been developed to address this deficiency with YOLOv5[23, 24]. It transmits the training data through the data loader, while also improving the training data. Data loader can apply three types of data enhancements which are color space, mosaic enhancement, and scaling. In image detection process, it uses R-CNN's anchor mechanism to improve the ability to detect small targets in the image with a multi-scale mechanism. And it can be adapted to different size images.[25, 26]. It consists of three main components basically [26]. Backbones, a convolutional neural network that generates and collects image features at different levels of image detail. The neck is the set of layers that mix and combine the collected image features and transmit these features to the prediction layer. Head can predict image features, create bounding boxes, and do category prediction. Fig. 5 shows the architecture of the YOLOv5 Model with its main components.

RESULTS AND DISCUSSION

The network is trained through Google Colab, which provides free access to powerful GPUs and requires no configuration. It is used a notebook computer using pretrained COCO weights based on YOLOv5. The dataset was added and adjusted the model's top layers to train disease and leaf classes to detect it, and set the period number to 300. Training of a 300-term model takes 30 minutes. Performance metrics for the training, development and validation clusters of our model are shown in Fig. 6. There are three types of loss. Box loss shows how well the algorithm is at finding the center of the object and the detected object box boundaries cover the object. Loss of Objectivity is a measure of the probability that the object is in the detected region. If the objectivity is high, it means that the viewport contains objects. Classification Loss determines how well the algorithm can detect which class the object belongs to.

The model developed rapidly in terms of precision, recall, and mean precision values after 50 periods. Loss values (loss of box, loss of objectivity and classification) in validation data decreased towards Period 100.

Training Results

The model was run on the same data set at different times and the results were compared. The mAP value in the obtained results is between 0.6-0.7. The sampled values

(a)

(b)

Figure 7. Collected images for the training dataset.

T. Boyar and K. Yildiz/ Hittite J Sci Eng, 2022, 9 (3) 159–166

Figure 8. Batch images for the validation dataset.

are described over the best recorded results. The dataset is not large enough, YOLOv5's data augmentation capability is important for object detection. During the training mosaic, copy paste, random affine(Rotation, Scale, Translation and Shear), MixUp, Albumentations, Augment HSV(Hue, Saturation, Value), Random horizontal flip data augmentation operations are performed. Fig. 7 shows the data augmentation and labeling operations performed on the training data.

Fig. 8 shows the object detection processes obtained for the validation dataset. The results of the determinations made for the "disease" and "leaf" classes are seen in the images taken in different periods. Classification and detection processes for the validation dataset yielded successful results.

After the model was trained with diseased and healthy leaf images, the images that were not included in the training dataset were applied to the network input. In Fig. 9 (a), the diseased parts are scattered and sparsely seen on the leaf. In this case, diseased sections could be detected. In Fig. 9 (b), the completely covered with disease was given to the entrance of the net upon detection. In this example it can be seen that almost all diseased sections have been detected. In Fig. 9 (c), the diseased parts are larger and more pronounced than in Fig. 9 (a), in this case the detection accuracy rate is about 90%. Fig. 9 (d) shows the detection of healthy leaves. It was detected about 82% accuracy in a dark image containing many leaves.

The study showed that early detection of powdery mildew disease on hazelnut leaves can be achieved with a deep

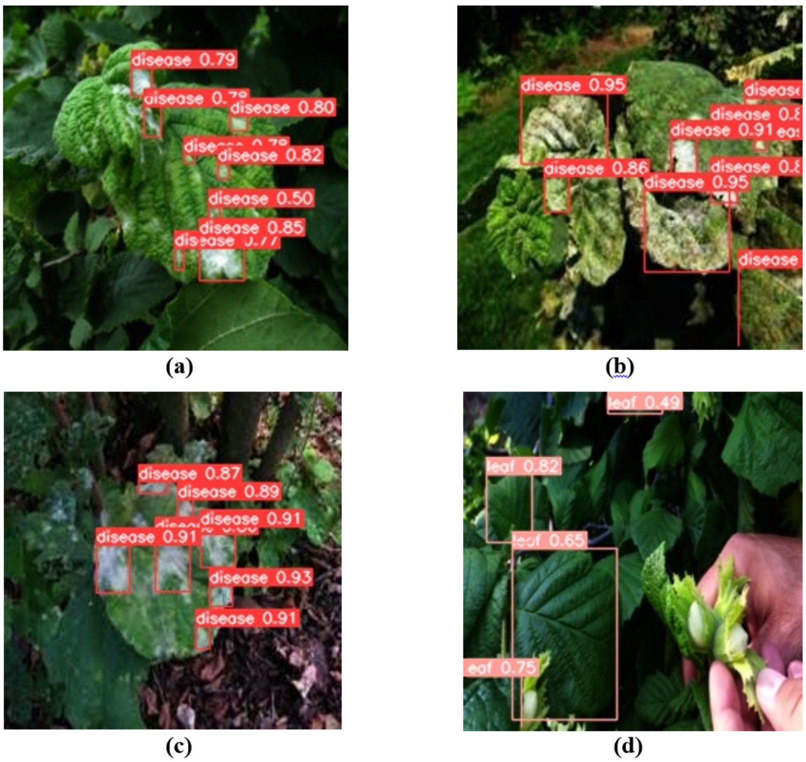


Figure 9. Detection process on images not included in the data set.

learning model. The proposed method for detecting powdery mildew disease on images in a short time is low-cost and saves time. The spraying method can be done on time with the help of the model's ability to detect small images. After the disease is detected on the images, the spread of the disease among the leaves can be prevented by applying agricultural pesticides. Thus, yield and quality loss in hazelnut will be prevented.

CONCLUSION

Our country is the world leader in hazelnut production. The increase in hazelnut production and the elimination of problems directly affect the country's economy. Powdery mildew disease encountered in hazelnut production reduces hazelnut production and effects the producers. Powdery mildew is a disease that can be transmitted from leaf to leaf and spreads rapidly. If powdery mildew disease on the leaves is detected early, the disease can be prevented from spreading by spraying the diseased area. For this reason, early detection of powdery mildew disease on hazelnut leaves is very important. Machine learning is widely used in the detection of plant leaf diseases.

In this study, YOLO was used to detect powdery mildew disease on images of hazelnut leaves. The images collected over the internet were processed. The detection process with external images was tried and the detection rate about 90%. It was shown that it is possible to detect powdery mildew disease on hazelnut leaves through images. However, the network could not reach equilibrium due to the lack of sufficient dataset. For future study during the periods when powdery mildew disease is seen in the orchards of hazelnuts, a large data set will be created by drone shots. With the large dataset, the model will be trained and the stability of the network will be ensured. An embedded system software capable of real-time detection that can be used by hazelnut producers will be developed.

CONFLICT OF INTEREST

There is no financial conflict of interest with any institution, organization, person related to our article named "Powdery Mildew Detection in Hazelnut with Deep Learning." and there is no conflict of interest between the authors.

AUTHOR CONTRIBUTION

The original research article titled "Powdery Mildew Detection in Hazelnut with Deep Learning" prepared by us has not been published in any journal before and is not in the publication stage in any other journal.

All authors contributed equally to the submission of our article for consideration in the Hittite Journal of Science and Engineering.

References

1. Anonim. Fındıkta Külleme. 2019 [cited 2022; Available from: https://arastirma.tarimorman.gov.tr/findik/Belgeler/Sol%20Men%C3%BC/E%C4%9Ftim%20ve%20Yay%C4%B1m/%C3%87ift%C3%A7i%20E%C4%9Fitim/K%C3%.
2. Erdoğan, V., Fındık: Yetiştiricilik, Sorunlar, Öneriler ve Yenilikler. Türktob Dergisi, 2018. 27: p. 4-10.
3. Kurt, Ş., Bitki fungal hastalıkları. Akademisyen Kitap Evi, 2013.
4. Mohammadpoor, M., M.G. Nooghabi, and Z. Ahmedi, An Intelligent Technique for Grape Fanleaf Virus Detection. International Journal of Interactive Multimedia and Artificial Intelligence, 2020. 6(1): p. 62-67.
5. Zhang, S., et al., Cucumber leaf disease identification with global pooling dilated convolutional neural network. Computers and Electronics in Agriculture, 2019. 162: p. 422-430.
6. Liang, W.-j., et al., Rice blast disease recognition using a deep convolutional neural network. Scientific reports, 2019. 9(1): p. 1-10.
7. Wagh, T.A., et al., Grapes leaf disease detection using convolutional neural network. International Journal of Computer Applications, 2019. 178(20).
8. Petrellis, N. A smart phone image processing application for plant disease diagnosis. in 2017 6th international conference on modern circuits and systems technologies (MOCAST). 2017. IEEE.
9. Mohanty, S.P., D.P. Hughes, and M. Salathé, Using deep learning for image-based plant disease detection. Frontiers in plant science, 2016. 7: p. 1419.
10. Singh, V. and A.K. Misra, Detection of plant leaf diseases using image segmentation and soft computing techniques. Information processing in Agriculture, 2017. 4(1): p. 41-49.
11. Yao, J., et al., A real-time detection algorithm for Kiwifruit defects based on YOLOv5. Electronics, 2021. 10(14): p. 1711.
12. Boyar, T., Powdery Mildew Disease on Hazelnut Leaves, in Leaf Disease on Hazelnut, T. BOYAR, Editor. 2022: https://drive.google.com/drive/folders/1iUNSnbPR9MFvehne1cHXcDE93pCJ2LdZ?usp=sharing.
13. Brin, L.P.S. Google. 1998; Available from: https://www.google.com.tr/.
14. Steve Chen, C.H., and Jawed Kari. Youtube. 2005; Available from: https://www.youtube.com.
15. SEZER, A., et al., Erysiphe corylacearum'un neden olduğu külleme hastalığına karşı Giresun ili fındık üretim alanlarında kimyasal mücadele olanaklarının belirlenmesi. Akademik Ziraat Dergisi. 8(Özel Sayı): p. 71-78.
16. Jocher, G. YOLOv5. 2020; Available from: https://ultralytics.com/yolov5.
17. Kubat, M., Neural networks: a comprehensive foundation by Simon Haykin, Macmillan, 1994, ISBN 0-02-352781-7. The Knowledge Engineering Review, 1999. 13(4): p. 409-412.
18. Akkoyun, S., N. Yildiz, and H. Kaya, Neural Network Estimation for Attenuation Coefficients for Gamma-Ray Angular Distribution. Physics of Particles and Nuclei Letters, 2019. 16(4): p. 397-401.
19. Demirbay, B. and A.B. Karakullukçu, Artificial neural network (ANN) approach for dynamic viscosity of aqueous gelatin solutions: a soft computing study. Avrupa Bilim ve Teknoloji Dergisi, 2020(18): p. 465-475.

20. LeCun, Y., et al., Gradient-based learning applied to document recognition. Proceedings of the IEEE, 1998. 86(11): p. 2278-2324.
21. Ruan, J., Design and Implementation of Target Detection Algorithm Based on YOLO. Beijing University of Posts and Telecommunications: Beijing, China, 2019.
22. Krizhevsky, A., I. Sutskever, and G.E. Hinton, Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 2012. 25.
23. Gorban, A.N., E.M. Mirkes, and I.Y. Tyukin, How deep should be the depth of convolutional neural networks: a backyard dog case study. Cognitive Computation, 2020. 12(2): p. 388-397.
24. Tang, Y., Exponential stability of pseudo almost periodic solutions for fuzzy cellular neural networks with time-varying delays. Neural Processing Letters, 2019. 49(2): p. 851-861.
25. Arif, I., W. Aslam, and Y. Hwang, Barriers in adoption of internet banking: A structural equation modeling-Neural network approach. Technology in Society, 2020. 61: p. 101231.
26. Wu, W., et al., Application of local fully Convolutional Neural Network combined with YOLO v5 algorithm in small target detection of remote sensing image. PloS one, 2021. 16(10): p. e0259283.
27. Fang, Y., et al., Accurate and Automated Detection of Surface Knots on Sawn Timbers Using YOLO-V5 Model. BioResources, 2021. 16(3).