ÇOMÜ Zir. Fak. Derg. (COMU J. Agric. Fac.)
2024: 12(1): 11-19
ISSN: 2147–8384 / e-ISSN: 2564–6826
doi: 10.33202/comuagri.1387580

Research Article

Identification of Some Sunflower Diseases Using Deep Convolutional Neural Networks

Hakkı Fırat Altınbilek[1*] **Ünal Kızıl[2]**

[1]Ministry of Agriculture and Forestry, İpsala District Directorate, 22400, İpsala/Edirne,
[2]Çanakkale Onsekiz Mart University, Faculty of Agriculture, Department of Agricultural Structures and Irrigation, Digital Agriculture Laboratory, 17020, Çanakkale
*Corresponding author: hfirataltinbilek@gmail.com

Received Date: 07.11.2023 Accepted Date: 29.02.2024

Abstract

Among the oilseed plants cultivated in Türkiye, sunflower ranks first in terms of cultivation area and production. Therefore, short time detection of sunflower diseases will help producers to take necessary actions on time. Computer-based deep learning techniques have made it possible to predict these diseases with high accuracy. This study aims at the effectiveness of image processing and modeling in predicting 3 different sunflower diseases. A total of 760 images were obtained and examined in the 2022-2023 production seasons in İpsala district of Edirne province. A series of data pre-processing techniques were applied to the developed Convolutional Neural Network (CNN) model and 3 different sunflower disease prediction systems were created. It has been revealed that the model can classify with an accuracy of 0.92. This study shows that the proposed CNN model demonsrated an effective classification performance and could help both producers and researchers for the early detection of sunflower diseases in Türkiye.
Keywords: Image classification, Deep learning, Machine learning, Sunflower diseases

Derin Evrişimli Sinir Ağları Kullanılarak Bazı Ayçiçeği Hastalıklarının Belirlenmesi

Öz

Türkiye'de yetiştirilen yağ bitkileri arasında ayçiçeği ekim alanı ve verim açısından ilk sırada yer almaktadır. Dolayısıyla ayçiçek hastalıklarının hızlı tespiti üreticilerin kısa sürede önlem almalarına yarayacaktır. Bilgisayar tabanlı derin öğrenme teknikleri bu hastalıkların yüksek doğruluk ile tahmin edilebilmesini mümkün kılmıştır. Bu çalışma, 3 farklı ayçiçek hastalıklarının tahmininde görüntü işleme ve modellemenin etkinliğini amaçlamaktadır. Toplamda 760 görüntü Edirne ili İpsala ilçesinde 2022-2023 üretim sezonlarında elde edilerek incelenmiştir. Geliştirilen Convolutional Neural Network (CNN) modeline bir dizi veri ön işleme teknikleri uygulanmış ve 3 farklı ayçiçek hastalığı tahmin sistemi yaratılmıştır. Modelin 0.92 doğrulukla sınıflandırma yapabildiği ortaya konmuştur. Bu çalışma, önerilen CNN modelinin etkili bir sınıflandırma performansı sergilediğini ve Türkiye'deki ayçiçeği hastalıklarının erken teşhisinde hem üreticilere hem de araştırmacılara yardımcı olabileceğini göstermektedir.
Anahtar Kelimeler: Görüntü sınıflandırma, Derin öğrenme, Makine öğrenmesi, Ayçiçek hastalıkları

Introduction

With the increase in the world population, the need for oil seed products has become an important issue. Ecological factors and agricultural practices affect yield of sunflower, one of the most consumed oil seed products (Sahin et al., 2021). While this is the case, the use of new technological approaches against the factors affecting quality in sunflower is becoming popular (Yunus Khan et al., 2014; Singh et al., 2019). One of these technologies is image processing and pattern recognition techniques (Khirade and Patil, 2015; Lati et al., 2019; Sethy et al., 2020).

Image processing and pattern recognition models are artificial intelligence products and have high skills in finding solutions, understanding and extracting meaningful relationships. It has an important role with its detection features after decision making. With increased calculation capabilities and automatic classification options, its use in predicting plant diseases in agriculture is becoming widespread (Camargo and Smith, 2009; Patrício and Rieder, 2018; Devaraj et al., 2019).

Wicaksono et al., (2020) used the convolutional neural network (CNN) method in the detection of apple leaf disease. In their study, they used a total of 3151 images consisting of images from 4 different leaf classes. They achieved an average accuracy of 94.9% on the test set. Singh and Misra (2017) classified rose and bean leaves with bacterial disease, lemon leaves with sunburn disease, and banana leaves with early blight disease using a genetic algorithm with image segmentation. In the study, when Support Vector Machines (SVM) was used with the proposed algorithm, the overall accuracy was found to be 95.71%.

Mohanty et al., (2016) achieved 99.35% classification accuracy in their study by training AlexNet and GoogleNet deep learning models to detect 14 different plant species and 26 plant diseases from a dataset consisting of 54306 images of infected and healthy plant leaves. Liu et al., (2020) proposed a deep learning-based model to detect 7 different grape diseases. By improving the images of the data set containing 107336 grape leaf images with image enhancement techniques, they achieved a 97.22% accuracy rate in detecting grape diseases.

Ensari et al., (2020) proposed the CNN method for the detection of diseases in grapes and corn. They used 1600 healthy and infected images, and achieved 97.03% accuracy. In the study of Wallelign et al., (2018) a model based on deep neural networks was proposed to classify and detect soybean plant diseases. This model consists of convolution, pooling and relu layers. The developed model was trained using images taken from real-natural environments and a classification accuracy of 99.32% was achieved.

Lu et al., (2017) proposed a new disease identification approach based on deep convolutional neural networks for paddy rice diseases. An average classification performance of 95.48% was achieved in experimental studies based on a data set containing different rice disease images, including diseased and healthy. Altınbilek and Kızıl (2022) used CNN to detect rice blight, brown spot diseases and healthy leaf images in paddy rice. As a result of their study with 1569 images, they stated that the model detect two different paddy rice diseases with 91.70% accuracy. In another study with paddy rice leaf dataset, a better accuracy rate for leaf disease detection was achieved using a machine learning approach based on transfer learning (Sharma et al., 2022).

Although there are studies on monitoring sunflower diseases, both laboratory and field-based, convolutional neural networks studies on sunflower disease prediction are very limited in Türkiye. Therefore, approaches that can make decisions and solve the problem in a short time are needed. In this study it was aimed to classify major sunflower diseases using RGB images and CNN techniques. With the developed model, a technological approach has been put forward in the Thrace region, where sunflower production is intense. The CNN model that can be used in devices based on the parcel-based early diagnosis system principle has been developed for small family businesses.

Material and Method

Study Area and Used Data

The images used to detect diseases on sunflower leaves were acquired from sunflower fields in İpsala district of Edirne province (Figure 1). Edirne ranks 3rd in oil sunflower production in Türkiye with 12.7% (TÜİK, 2022).

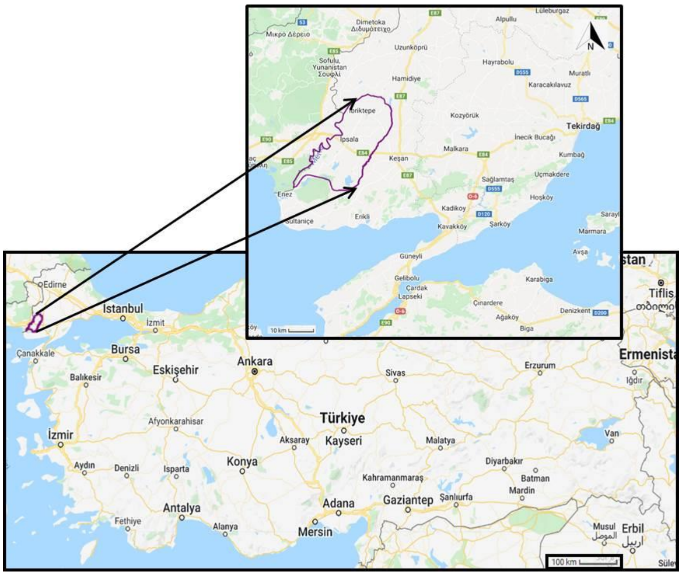

Figure 1. Location of study area

Image acquisition started from the second week of May (the first appearance of the diseases) in 2022 and 2023 and continued until the end of September in both years. Healthy and infected images were obtained from various sunflower fields to identify leaf scar, phoma blight, and gray mold diseases. A total of 760 RGB images were obtained from above, below and from the side during sunny and cloudy hours of the day using the Redmi Note 9 Pro mobile device with 64 megapixel image quality.

Leaf Scar

Leaf scar is one of the diseases that causes important yield losses in sunflower cultivation. It can be seen anytime during the early and developmental stages of the plant when conditions are suitable. It is recommended to use resistant varieties, clean seeds and implementation of rotation measures in its control (Lagopodi and Thanassoulopoulos, 1998).

Gray Mold

The gray mold on the head and stem is caused by this pathogen. The leaves start to dry off in the interim. When the head grains are ripening, these symptoms show up. There are brown spots on the back. These parts appear dusty as a result of fungal spores and mycelium covering them. When the weather is wet, spores can spread. It is advised to try natural control methods since chemical control is challenging owing to the pathogen's resistance (Mukhtar, 2009).

Phoma Blight

The fungus enters the plant through natural openings such as stomata or damages caused by physical means. Conidia are spread by rainwater. Rainfall, humidity, and temperature all play a significant role in the onset and spread of the disease. Rains after flowering increase the severity of infection. In cultural precautions against the disease, it is recommended to avoid frequent planting, choosing resistant varieties and preventing physical damages on the stems of the plants (Mukhtar, 2009; Deb et al., 2020).

Pre-processing of Images

A range of preprocessing procedures were applied to sunflower leaf images in order to shorten model runtime and minimize noise disruption. The images were first classified as having leaf scars, phoma blight, gray mold and healthy (Figure 2). The overall image data was compressed by downsizing the image from 4640 × 3472 to 256 × 256. Since the backgrounds of all leaf images were gray and the images themselves were all in color, they were all compressed into a range of 0 to 1 to ensure stability and reduce noise. The photos were then inverted and their horizontal and vertical ratios were chosen to be 0.2 so that the CNN model could process the data set more thoroughly. The rotation ratio, which determines the range in which the images will be randomly rotated, is set to 40. The width_shift and height_shift ratios, which determine the intervals at which the images will be randomly rotated vertically or horizontally, are taken as 0.2. The zoom range ratio used to zoom in the images was determined as 0.2. Then, in order to extend the image at a certain angle, known as the shear angle, the shear range ratio was set to 0.3 and the data augmentation process was completed.

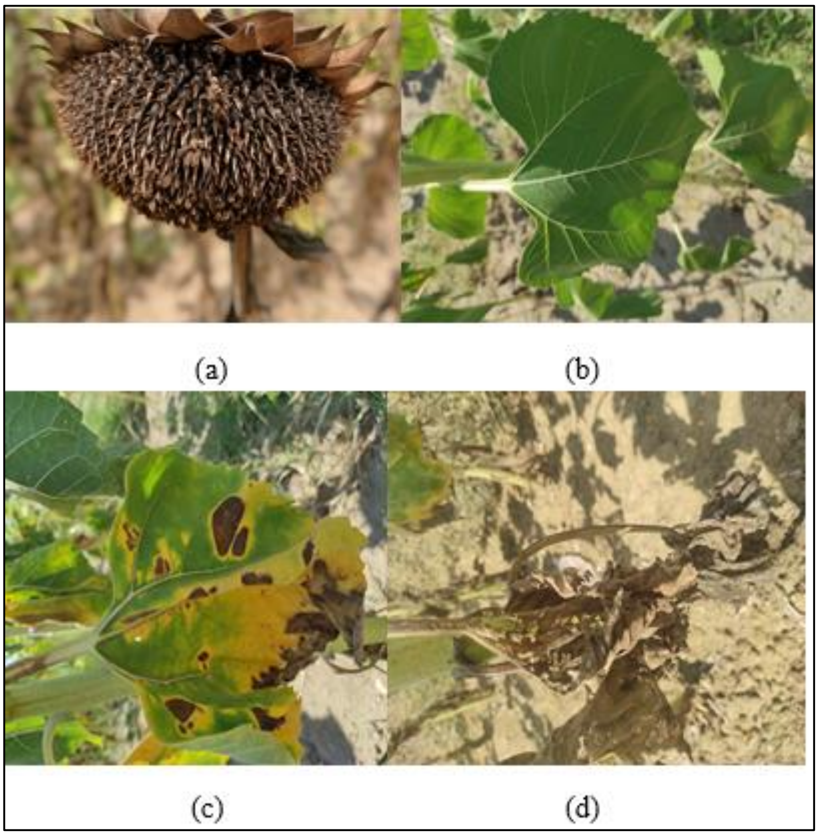

Figure 2. Sample dataset. (a) gray mold, (b) healthy, (c) leaf scars, (d) phoma blight

Modelling

Deep learning is a machine learning model with sequential layers. Each layer uses the output of the previous layer as input (Şeker et al., 2017). CNN architecture; It stands out with its popularity among various algorithms of deep learning. CNN is used to classify the image, cluster similarities and perform object recognition (Radovic et al., 2017). The CNN model for detecting sunflower diseases was created in the Google Colaboratory (Colab). In classification, tensorflow, matplotlib.pyplot, IPython.display, gpu, Sequential, compile, model.fit, and sklearn.metrics libraries were used. The dataset for the proposed CNN model, consisting of 760 photos, was uploaded to Google Drive and then moved to Google Colab. The training, validation, and test percentages were 70, 20, and 10%, respectively. The aim for selecting these ratios is to improve the model's learning speed and accuracy score. The epoch number was set at 100 to assure that the model had no overfit and accurately represented the validation accuracy. Figure 3 shows the model's flowchart.

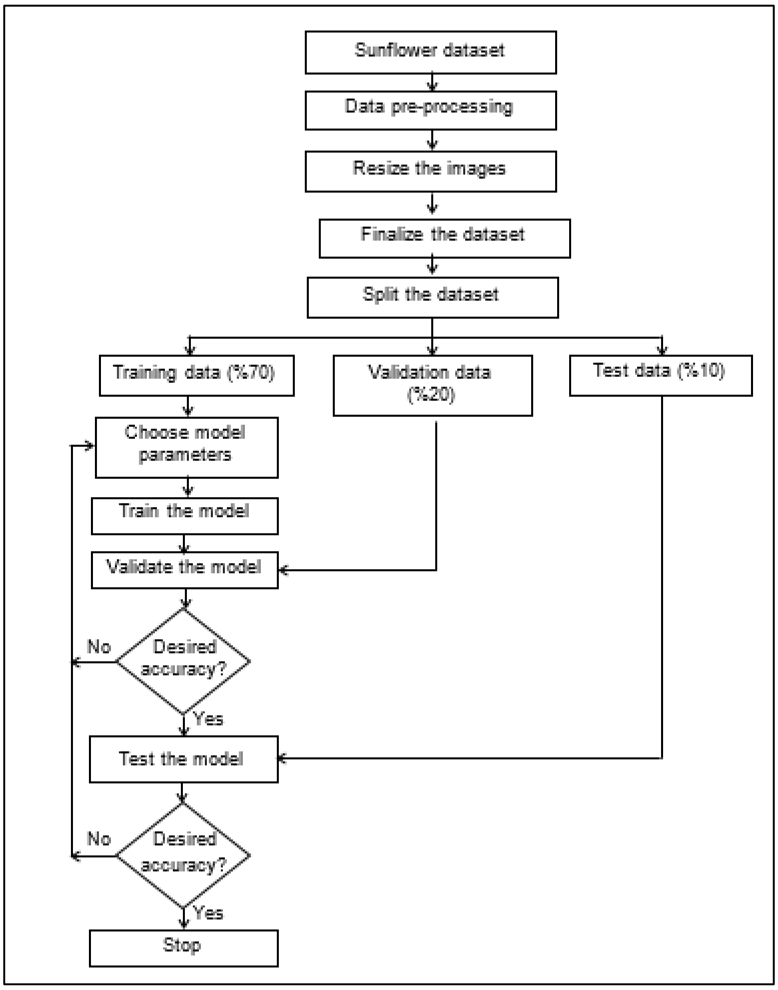

Figure 3. Flowchart of the model

The convolution layer is the first layer from which the input image features are extracted. Large-dimensional input allows lower-dimensional features containing image features to be obtained by moving filters of different sizes on the image (Prabhu, 2018). Low-level features of an image are extracted using the convolution process in the first layer, and more complex features are eliminated in the subsequent convolution layers. (Demir et al., 2020). The image features obtained as a result of the activation layer convolution process are eliminated by a non-linear function called Rectified Linear Unit (RELU). In the pooling layer, the number of channels of the features matrix obtained as a result of the activation process is kept continuously stable and the maximum or average value method lowers the widths and heights of this matrix. The fully connected layer enables the conversion of features of different dimensions obtained as a result of convolution, activation and pooling processes into one-dimensional features (Aslan, 2022). Softmax enables the classification of features obtained in previous layers. In the classification process, probabilistic values are used to assign the class of interest (Tumen et al. 2018). The number of filters used in each convolution layer in the model is listed as 32 × 64 × 64 × 64 × 64 × 64. 6 convolution + correction layers made up the total of eight layers that were utilized (Figure 4). 32 and 64 nodes were found to be provide in each layer. Through trial and error, this quantity can be changed to be greater or smaller based on the accuracy score of the model. The modeling process used the 2 × 2 maximum pooling method. Two dense layers were added at the end of the CNN to extract features from the convolution and pooling layers. The categorization was produced by the second dense layer. The probability of which class each entry belongs to was determined with the softmax function.

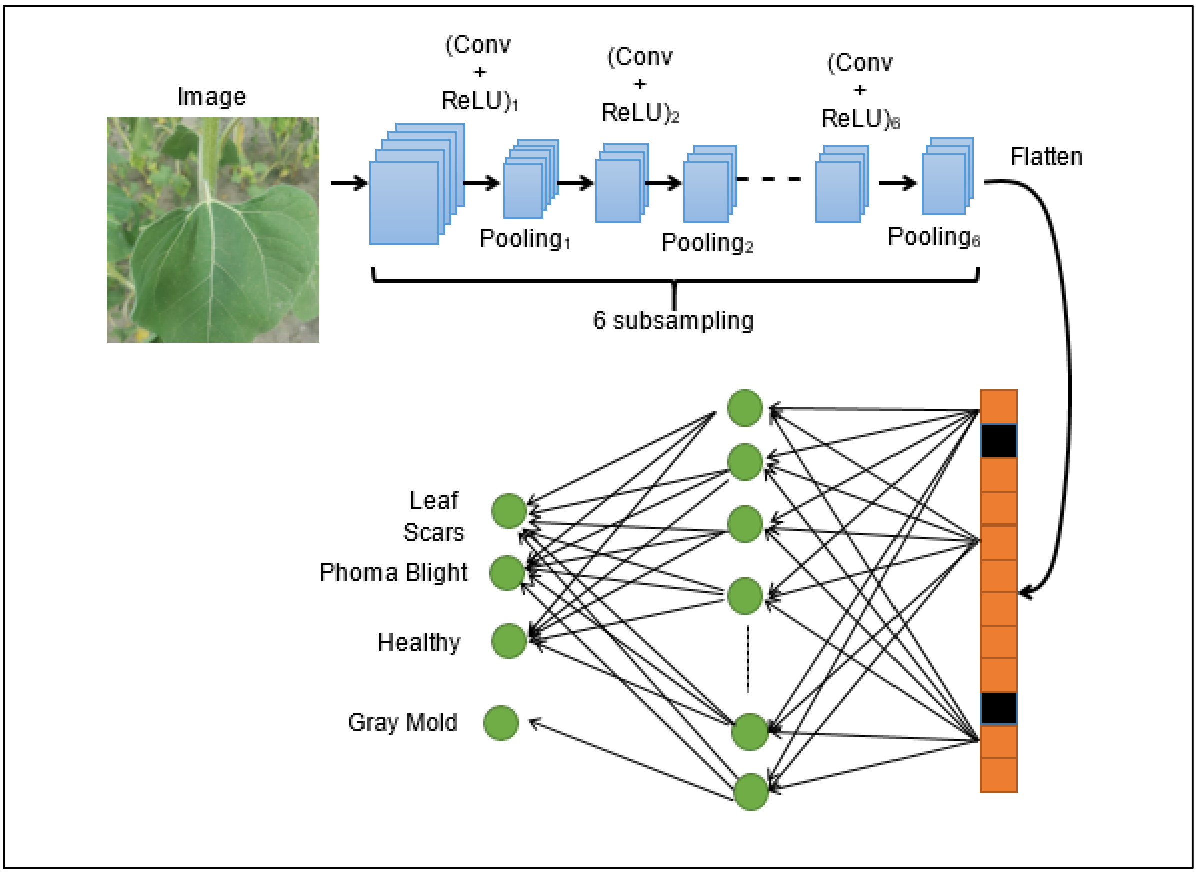

Figure 4. CNN model structure

Results and Discussion

Three portions, representing 70%, 10%, and 20% of the total dataset images, were randomly selected for training, validation, and testing. The validation data set was used to develop the model, and the training data set was utilized as a portion of the training set in the proposed model's learning process. The model parameters were also modified using the validation data set. The performance evaluation was conducted using the test dataset. CNN parameters (learning rate, batch size, and number of epochs) were tuned during training to provide the optimal model. Numerous experiments were conducted in order to determine the proper values for various parameters. As a result, 0.0001, 32, and 100 were chosen as the starting learning rate, batch size, and maximum epochs, respectively. Figure 5 displays the model's accuracy and loss information graphs during the transfer learning process.

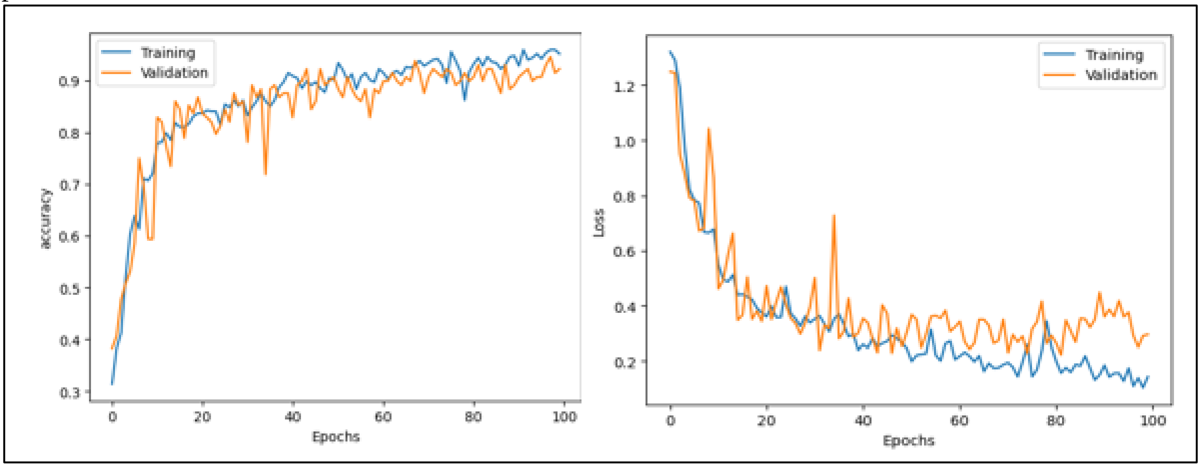

Figure 5. CNN training-validation accuracy and loss curve

As seen in Figure 5, there was no significant improvement in validation accuracy after approximately the 95th epoch, training process was successfully completed with 0.92. Additionally, as seen in Figure 4, the loss data rate during training and validation was well below 0.3 and 0.2, respectively. Thus, the training process was completed successfully without any significant data loss. Figure 6 shows a more comprehensive performance evaluation specific to classes with confusion matrix.The estimated number of data for each category is shown by the sum of each column in Figure 6's confusion matrix, while the actual values in the data set are displayed by each row. The model properly categorized 23 (90%) class data of gray mold, 12 (93%) class data of phoma blight, and 32 (97%) class data of leaf scars, as shown in Figure 6. However, it looks that 9 (16%) of the leaf scars categorized data were mistakenly identified as healthy data. Values of the confusion matrix demonstrate how frequently the data labels in the data set are predicted.

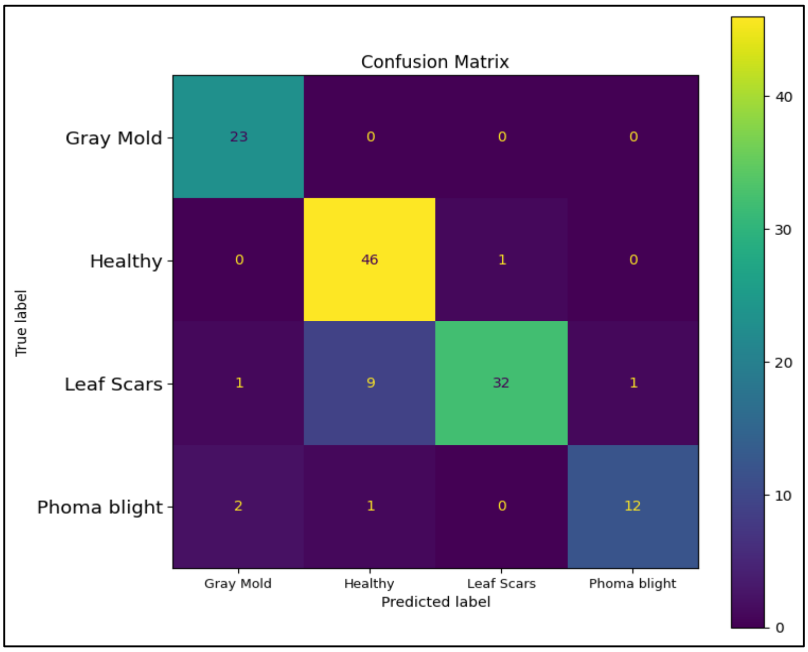

Figure 6. Confusion matrix of the model

In the study, the proposed method's performance was assessed using the performance evaluation criteria as a guide. The general classification results of the study are shown in Table 1. Accuracy, precision, recall and f1-score results are given in the table.

Table 1. General classification results

Accuracy	Precision	Recall	F-1 Score
0.92	0.90	0.89	0.90

As seen in Table 1, the sunflower disease prediction model can detect diseases with a 0.92 accuracy rate. Additionally, a 0.90 success rate was achieved in precision and f1-score values. Table 2 lists several research using the CNN approach to identify sunflower diseases. The research included in Table 2 generally agree with this study in terms of accuracy. Gulzar et al., (2023) achieved the best model accuracy using 1892 images and 300 epochs. Because there are only 100 epochs in the proposed study and only 760 photos in the dataset, the accuracy rate is lower than that of other studies. Using a larger dataset and epoch number can boost the model's accuracy rate, but doing so may lead to overfitting. The model's predictive score in the test dataset is decreased as a result of this issue.

Table 2. Several research using the CNN approach to identify sunflower diseases

Author	Dataset	Diseases	Overall accuracy
Ghosh et al., 2023	467	Downy Mildew, Gray Mold, Leaf Scars, Healthy	0.93
Dawod and Dobre, 2022	858	Downy Mildew, Powdery Mildew, Alternaria Leaf Blight, Rust, Healthy	0.92
Malik et al., 2022	329	Alternaria Leaf Blight , Downy Mildew, Phoma Blight, Verticillium Wilt, Healthy	0.89
Gulzar et al.,, 2023	1892	Downy Mildew, Gray Mold, Leaf Scars, Healthy	0.97
Sirohi and Malik, 2021	Not specified	Alternaria Leaf Blight , Downy Mildew, Verticillium Wilt, Phoma Blight, Healthy	0.89

Conclusion

Traditional machine learning algorithms for plant disease detection involve feature extraction from disease images without the need for any processing. This is a difficult and time-consuming procedure. In this case, it may delay the necessary steps to counteract identified plant diseases. In order to identify sunflower diseases more quickly and automatically, this paper proposes a practical deep learning-based approach. Using images of sunflower leafs, a few basic CNN model parameters were defined and allowing for the rapid and accurate identification of the disease sunflower leafs. Sunflower diseases were identified in experimental research with an accuracy rate of 0.92. Although the result is promising, it is not practical for use by farmers. Knowledge and awareness are needed for models running on computers. Therefore, it is planned to develop farmer-friendly, easily accessible and integrated mobile application software in future studies.

Authors' Contributions

The authors declare that they have contributed equally to the article.

Conflicts of Interest Statement

The authors declare no competing interests.

References

Altınbilek, H.F., Kızıl, Ü., 2022. Identification of some paddy rice diseases using deep convolutional neural networks. Yuzuncu Yil University Journal of Agricultural Sciences. 32(4): 705-713.

Aslan, M., 2022. CoviDetNet: A new -19 diagnostic system based on deep features of chest x-ray. International Journal of Imaging Systems and Technology. 32(5): 1447-1463.

Camargo, A., Smith, J.S., 2009. An image-processing based algorithm to automatically identify plant disease visual symptoms. Biosystems Engineering. 102: 9–21.

Dawod, R.G., Dobre, C., 2022. Automatic segmentation and classification system for foliar diseases in sunflower. Sustainability. 14: 11312.

Deb, D., Khan, A., Dey, N., 2020. Phoma diseases: Epidemiology and control. Plant Pathology. 69: 1203–1217.

Demir, F., Türkoglu, M., Aslan, M., Şengür, A., 2020. A new pyramidal concatenated cnn approach for environmental sound classification. Applied Acoustics. 170: 107520.

Devaraj, A., Rathan, K., Jaahnavi, S., Indira, K., 2019. Identification of plant disease using image processing technique. International Conference on Communication and Signal Processing. 4-6 April 2019, Chennai, India.

Ensari, T., Armah, D.C., Balsever, A.E., Dağtekin, M., 2020. Görüntü tabanlı dijital bitki fenotiplemesi için konvolüsyonel sinir ağları. European Journal of Science and Technology. (Special Issue): 338-342.

Ghosh, P., Mondal, A.K., Chatterjee, S., Masud, M., Meshref, H., Bairagi, A.K., 2023. Recognition of sunflower diseases using hybrid deep learning and its explainability with ai. Mathematics. 11: 2241.

Gülzar, Y., Ünal, Z., Aktaş, H., Mir, M.S., 2023. Harnessing the power of transfer learning in sunflower disease detection: A comparative study. Agriculture. 13: 1479.

Khirade, S.D., Patil, A. B., 2015. Plant disease detection using image processing. International Conference on Computing Communication Control and Automation. 26-27 February 2015, Pune, India.

Lagopodi, A.L., Thanassoulopoulos, C.C., 1998. Effect of a leaf spot disease caused by alternaria alternata on sunflower in Greece. Plant Dis. 82: 41–44.

Lati, R.N., Filin, S., Elnashef, B., Eizenberg, H., 2019. 3-D image-driven morphological crop analysis: a novel method for detection of sunflower broomrape initial subsoil parasitism. Sensors. 19(7): 1569.

Liu, B., Ding, Z., Tian, L., He, D., Li, S., Wang, H., 2020. Grape leaf disease identification using improved deep convolutional neural networks. Frontiers in Plant Science 11.

Lu, Y., Yi, S., Zeng, N., Liu, Y., Zhang, Y., 2017. Identification of rice diseases using deep convolutional neural networks. Neurocomputing. 267: 378-384.

Malik, A., Vaidya, G., Jagota, V., Eswaran, S., Sirohi, A., Batra, I., Rakhra, M., Asenso, E., 2022. Design and evaluation of a hybrid technique for detecting sunflower leaf disease using deep learning approach. Hindawi Journal of Food Quality. 2022: 9211700.

Mohanty, S.P., Hughes, D.P., Salathé, M., 2016. Using deep learning for image-based plant disease detection. Frontiers in Plant Science. 7: 1419.

Mukhtar, I., 2009. Sunflower disease and insect pests in Pakistan: a review. African Crop Science Journal. 17(2): 109 – 118.

Patrício, D. I., Rieder, R., 2018. Computer vision and artificial intelligence in precision agriculture for grain crops: A systematic review. Computers and Electronics in Agriculture. 153: 69-81.

Radovic, M., Adarkwa, O., Wang, Q., 2017. Object recognition in aerial images using convolutional neural networks. Journal of Imaging. 3(2):21.

Sethy, P.K., Barpanda, N.K., Rath, A.K., Behera, S.K., 2020. Image processing techniques for diagnosing rice plant disease: a survey. Procedia Computer Science. 167(220): 516–530.

Sharma, M., Kumar, C. J., Deka, A., 2022. Early diagnosis of rice plant disease using machine learning techniques. Archives of Phytopathology and Plant Protection, 55(3): 259-283.

Singh V, Misra, A.K., 2017. Detection of plant leaf diseases using image segmentation and soft computing techniques. Information processing in Agriculture. 2017; 4(1):41-49.

Tümen, V., Yıldırım, O., Ergen, B., 2018. Detection of driver drowsiness in driving environment using deep learning methods. Conference: The Scientific Meeting on Electrical-Electronics & Biomedical Engineering and Computer Science. 18-19 April 2018, İstanbul, Türkiye.

Sirohi, A., Malik, A., 2021. A hybrid model for the classification of sunflower diseases using deep learning. 2nd International Conference on Intelligent Engineering and Management (ICIEM). 28-30 April 2021, London, United Kingdom.

Şahin, S., Kurtulbaş, E., Toprakçı, İ., Pekel, A. G., 2021. Determination of lipid oxidation in sunflower oil treated with several additives. Biomass Conversion and Biorefinery. 13: 3953-3961.

Şeker, A., Diri, B., Balık, H.H., 2017. Derin öğrenme yöntemleri ve uygulamaları hakkında bir inceleme. Gazi Mühendislik Bilimleri Dergisi. 3(3): 47-64.

TÜİK, 2022. Ayçiçeği üretim miktarı, ekilen alan büyüklüğü, ithalat miktarı ve değeri, ihracat miktarı ve değeri gibi yıllar bazında sayısal veriler. https://biruni.tuik.gov.tr/medas/?locale=tr, Erişim Tarihi: 06.10.2023

Wallelign, S., Polceanu, M., Buche C., 2018. Soybean plant disease identification using convolutional neural network. In The Thirty-First International Flairs Conference. May 2018, Melbourne, United States. 146-151.

Wicaksono G. Andryana S., 2020. Aplikasi pendeteksi penyakit pada daun tanaman apel dengan metode convolutional neural network. Journal of Information Technology and Computer Science. 5(1): 9-16.

Yunus Khan, T.M., Atabani, A.E., Badruddin, I.A., Badarudin, A., Khayoon, M.S., Triwahyono, S., 2014. Recent scenario and technologies to utilize non-edible oils for biodiesel production. Renewable and Sustainable Energy Reviews. 37: 840-851.