BALKAN JOURNAL OF ELECTRICAL & COMPUTER ENGINEERING, Vol. 13, No. 1, March 2025
Research Article

Deep Learning based Disease Detection from Potato Leaf Images

Abdulkerim Oztekin and Kenan Almas

Abstract—**This study aims to detect diseases from potato leaf images using deep learning methods. In the proposed work, a large and comprehensive image dataset of healthy and various potato diseases was used to detect common diseases (late blight, early blight) seen in potato plants. Models were developed to detect potato diseases using different Convolutional Neural Network (CNN) architectures and hybrid models. The developed models were trained under various hyperparameters and evaluated using performance metrics such as accuracy, precision, recall, and F1-score. The study shows that deep learning methods can be effectively used in the detection of potato diseases and can contribute to previous studies in this field. The dataset was tested with four different ResNet models and evaluated with various performance metrics. It is thought that the obtained test results can provide a significant information for disease management and productivity increase in potato cultivation. And artificial intelligence (AI)-based disease detection can lead to innovations in the field of agriculture, and can also contribute to machine-human interaction. Our work also highlights the success and importance of ResNet deep learning models in the field of image extraction.**

Index Terms—**Convolutional Neural Network (CNN), Deep Learning, Potato Disease Detection, ResNet.**

I. Introduction

AGRICULTURE IS one of the oldest and most fundamental activities of humanity, and has made significant contributions to the emergence and development of civilizations. Today, the increasing world population, climate change and limited resources make agricultural production even more important and necessitate the need to increase efficiency. In order to cope with these challenges and ensure sustainable food production, technology has increasingly begun to be used in the agricultural sector. Especially in recent years, rapid developments in the field of artificial intelligence have the potential to revolutionize the agricultural sector. Today, the rapidly increasing world population and changing consumption habits have made food supply and security a global priority. In this context, the efficiency and sustainability of the agricultural sector is of critical importance for the future of humanity. Plant production, which forms the basis of the agricultural sector, is currently faced with human-induced factors such as incorrect planting, incorrect irrigation and incorrect fertilization, as well as environmental challenges such as climate change, water scarcity and soil erosion. These challenges negatively affect the productivity and quality of agricultural products, threatening food security.

Technological developments offer new opportunities to overcome these challenges and increase efficiency in the agricultural sector. Image processing and deep learning techniques, which have developed rapidly in recent years, have the potential to revolutionize agricultural applications. Many different machine learning methods have been successfully used in the literature for early diagnosis, quality control and classification of plant diseases [1-6]. These techniques provide fast, accurate and objective quality assessment thanks to the ability to objectively determine the visual features of food items such as color intensity, color distribution, visual defects, size and shape [7]. One of the biggest disadvantages of machine learning algorithms is the need to pre-process the data and manually determine the features. Deep learning eliminates this laborious process and automatically extracts meaningful features from raw data, thus making it possible to create more powerful and flexible models [8].

Potatoes are very important food sources and one of the most consumed agricultural products in the world. However, there are various difficulties in the potato production and processing process. Post-harvest classification and quality control processes are usually carried out manually, which leads to many problems such as substandard products, human errors, increased labor costs and production waste [9]. The activity of diagnosing and characterizing multiple aspects of potato diseases based on their characteristics, symptoms and appearances is called potato disease classification. Recent studies have proven the success of machine learning-based potato classification, disease detection and quality control systems [10-13].

Deep learning is a more advanced form of artificial neural networks and has the ability to recognize and classify complex patterns by training on large data sets. In this way, more

Abdulkerim Öztekin, is with Department of Electrical&Electronics Engineering, Batman University, Batman, Turkiye, (e-mail: abdulkerimoztekin@gmail.com).
https://orcid.org/0000-0002-0698-3525

Kenan Almas, is with Department of Information Technologies, Batman University, Batman, Turkiye, (e-mail: bajece@hotmail.com).
https://orcid.org/0000-0002-2282-2641

Manuscript received Feb 28, 2025; accepted March 24, 2025.
DOI: 10.17694/bajece.1649068

effective results can be obtained in agricultural applications such as plant disease detection, product classification and quality control [14]. Deep learning models used in potato quality control are generally based on the Convolutional Neural Networks (CNN) architecture, and there are various studies in the literature that uses CNN models in classification, disease and defect detection, and quality control [15-21]. CNNs have the ability to automatically learn and hierarchically represent features in images. In this way, various defects, diseases and quality traits in potatoes can be detected with high accuracy. These studies show that deep learning-based image processing systems can surpass human performance in potato quality control and increase productivity in the agricultural sector. In particular, the use of new imaging technologies such as hyperspectral imaging, thermal imaging, and 3D imaging can provide a more comprehensive analysis of the internal and external quality characteristics of potatoes [22-25]. In addition, studies on the explainability and interpretability of deep learning models will provide a better understanding of the decision-making processes of these models and increase their reliability.

The main purpose of this study is to detect diseases from potato leaf images using deep learning methods and to contribute to previous studies in this field. In this research, an approximation model from deep learning models will be applied to determine healthy and various known potato diseases. For this purpose, models will be developed to detect potato diseases using CNN architectures with the help of a large and comprehensive image dataset of potato diseases and the developed models will be optimized and trained with different parameters. Thus, the effectiveness and potential of deep learning methods in the detection of potato diseases will be evaluated.

This paper is organized as follows: Section 2 explains the dataset and the conducted Deep Learning algorithms for the proposed study. The results are presented in Section 3 and finally, concluding remarks and a comprehensive discussion are made in Section 4.

II. Material And Methods

2.1. Properties of the Dataset

In this study, the analyses on plant diseases were based on the PlantVillage dataset (www.kaggle.com/code/redpen12/cnn-disease-detection/input). This dataset was obtained from the Kaggle platform and consists of three channel 256x256 pixel color images which are rescaled and resized to 224x224 pixels to be used in our models. This dataset comprises high-resolution images of potato plants affected by various diseases, including early blight, late blight, and healthy leaves. It is designed to support the development and evaluation of image recognition models for precise disease detection and classification, contributing to advancements in agricultural diagnostics. Some augmentation techniques such as flipping, rotating, shifting, shearing and brightening has been applied the training data to improve the model's generalization capability and prevent overfitting.

The dataset is a benchmark dataset used to recognize disease and health status from potato leaves. The dataset consists of 3 different image groups as early blight, late blight and healthy. The dataset consists of 2152 images that are visually close to each other and not homogeneous. This situation has a complicating effect on classification tasks. Of these images, there are 1000 images with early blight disease, 1000 images with late blight disease and 152 healthy images. The dataset has been split into training, validation, and test sets with specified ratios (70%-15%-15%). Randomly selected images from the dataset are given in Fig. 1.

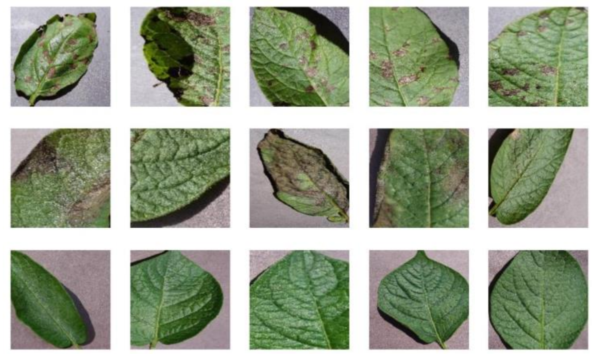

Fig.1. Some images from healthy and diseased potato leaves

2.2. Methods

2.2.1. Transfer learning methods

Transfer learning is a powerful technique that transfers previously learned information to new tasks in machine learning, accelerating the learning process and improving performance [26]. It is based on the principle of using knowledge acquired in a field in a similar field. For example, a model trained for image classification can be used for a related task such as object detection by fine-tuning. This is a great advantage, especially in cases where data sets are limited or training time is limited. The reasons why transfer learning is effective include data efficiency, fast learning, and improved performance. Since a pre-trained model has already learned general features, it needs less data specific to the new task and is trained faster. In addition, it can achieve better results in the target task by generalizing the information it has learned from the source task. Fig. 2 shows the differences between transfer learning and machine learning.

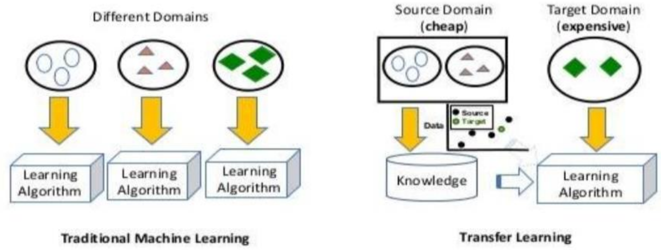

Fig.2. Illustration of process differences between transfer learning and machine learning [27]

Deep learning models benefit greatly from transfer learning, especially in areas such as image processing, natural language processing, and speech recognition. For instance, a CNN trained on a large dataset such as ImageNet can be fine-tuned and used for a different task such as medical image classification. Factors such as the similarity between the source and target tasks, the size and quality of the source dataset, and the model architecture are important for the success of transfer learning. In the transfer learning process, it is important to choose a source dataset with similar characteristics to the target dataset, select a suitable pre-trained model, and fine-tune the model to adapt it to the target task.

As a result, transfer learning leads to a significant paradigm shift in machine learning, enabling faster, more efficient, and more effective learning. With rapid developments in deep learning and increasing data availability, the importance of transfer learning is increasing.

2.2.2. Convolutional Neural Networks

Convolutional neural networks (CNN) are powerful deep learning models that have shown great success in image processing and data classification [28]. Although their foundations were laid in the 1980s, they began to be processed and trained after the 1990s and began to be used in image processing applications in 1995. CNNs provide superior performance by effectively identifying features in visual data thanks to their unique structures [29-30]. Convolutional Neural Networks (CNNs) are deep learning models widely used in processing visual data. They are designed to learn patterns and features in complex images. CNNs perform a shifting operation on the input image using small filters called kernels and thus perform the convolution operation. CNNs, which consist of successive layers such as convolution, pooling, full connection and output layers, extract important features in the image and achieve high success in classification, object recognition and similar tasks [28]. CNNs are a deep learning model that has revolutionized the field of image processing. These models analyze the relationships between pixels, the basic units that make up images, and extract higher-level meanings. CNNs consist of many layers built on top of each other. Each layer represents a different aspect of the image. For example, the first layers detect simple edges, while later layers identify more complex shapes and objects. The learning process of CNNs is optimized through the backpropagation algorithm. This algorithm allows the network to learn from its mistakes and make more accurate predictions. The general structure of the CNN network is shown in Fig. 3.

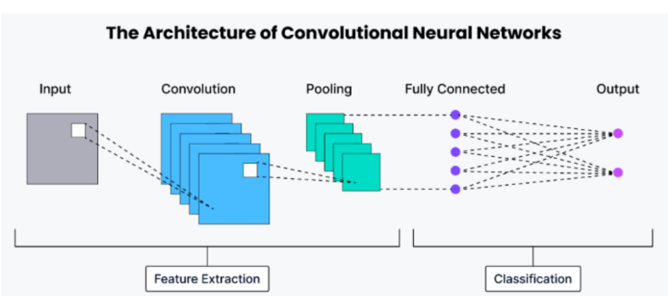

Fig.3. General Structure of CNN Network

CNNs are a powerful tool that allows computers to understand images like humans. These networks identify important features in images, just as a human focuses on specific features (for example, the eyes, nose, and mouth of a face) when examining a picture. Through a process called convolution, CNNs learn patterns, textures, and shapes in images. In this way, they are used in tasks such as recognizing, classifying, and locating objects. CNNs have achieved great success in many areas, such as diagnosing diseases in medical imaging and perceiving the environment in autonomous vehicles. CNNs generally consist of four main layers: convolution layer, pooling layer, correction layer, and fully connected layer [28]. Activation functions are critical components that organize the input signals of neurons, which are the basic building blocks of neural networks, and determine the learning ability of the network. These functions compress the outputs of neurons to certain ranges, allowing the network to work more efficiently. Various activation functions, such as Sigmoid, ReLU, and Hyperbolic Tangent, are selected according to the nature of the problem and control whether neurons are activated or not, directing the network's information processing process. It is important to use a sensitive activation function, especially in areas where matrix operations are intensive, such as image processing. Otherwise, the learning speed may decrease.

The pooling layer is an important component that reduces the number of weights and computational costs by reducing the data size in CNNs. This layer can speed up the learning process while also increasing the generalization ability of the model. No learning takes place in the pooling layer; the purpose of this layer is to summarize the feature maps coming from the convolution layer and reduce their size while preserving important information. In this way, the training time is shortened, thus saving time and resources [31]. These methods are widely used in many areas such as data processing and signal analysis. For example, a maximum pooling layer of 2x2 size takes the maximum value of each 2x2 block in the image matrices, reducing the size and reducing the processing cost. During the pooling process, the "shift" parameter determines how many units the filter will shift on the input data. The size depending on the pooling layer is also very important in controlling the process here. In summary, the pooling layer increases the processing speed of deep learning networks, emphasizes the important features of the data by not performing learning at the same time and increases the generalization ability of the model.

Fully connected layers come after convolution and pooling layers in CNNs. Convolution layers extract important features in the input image. Pooling layers reduce the size of these features and allow the network to do less computation. Fully connected layers take flattened feature maps and use them for classification. Each neuron is connected to each feature in the input. These connections have weights that are adjusted during the learning process of the network.

Dropout layer is a regularization technique used to prevent overfitting in deep learning models. Overfitting is when the model memorizes patterns in the training data and cannot generalize to new, unseen data. The most important requirement of the dropout layer is to prevent overfitting. Dropout works by temporarily removing randomly selected neurons from the network at each iteration during training. This prevents neurons from becoming too dependent on each other and memorizing the noise in the training data. Since a different subset of neurons will be active in each iteration, the model becomes more robust and has a higher ability to generalize.

In CNNs, visual data is processed by convolution and pooling operations. The resulting feature maps are combined and flattened and given as input to the fully connected layer. In this way, the high-dimensional information in the image is represented in a lower-dimensional vector and can be used for tasks such as classification.

ResNet is a deep learning model developed by Microsoft Research in 2015 and is based on the Convolutional Neural Network (CNN) architecture. It has achieved significant success in image classification and object detection tasks, especially on the ImageNet dataset. ResNet is based on the principle of "residual learning". In traditional deep neural networks, each layer receives and processes the output of the previous layer. This structure can lead to the "vanishing gradient" problem in deep networks. Gradient vanishing makes it difficult to update the weights in the lower layers of the network and slows down the learning process. ResNet uses "skip connections" or "residual connections" to solve this problem. These connections directly transmit the output of one layer to several layers later. Thus, gradients can be propagated more effectively to all layers of the network, making it easier to train deep networks.

ResNet-18, as its name suggests, consists of 18 layers. These layers consist of convolutional layers, batch normalization layers, and ReLU activation functions. Each residual block contains two or three convolutional layers. ResNet-18 can be used in various tasks such as image classification, object detection, and image segmentation. It is also often preferred for transfer learning. ResNet-18's fast trainability and good performance have made it a popular model for many applications. It has also formed the basis for the development of larger and deeper ResNet variants such as ResNet-50, ResNet-101, and ResNet-152.

ResNeXt is a convolutional neural network (CNN) model developed by Facebook AI Research in 2016 that has had a significant impact in the field of deep learning, especially in image recognition tasks. An extension of the ResNet architecture, ResNeXt improves the learning capacity and performance of the network by using a technique called "grouped convolutions." The basic idea of ResNeXt is to divide a convolution layer into multiple "groups" of smaller filters. Each group works on a subset of the input channels and learns its own filters. The outputs of the groups are then combined to obtain the result.

2.3. Performance Measures

In the context of deep learning, different metrics are used to measure and compare the performance of models used especially in classification problems. These metrics measure the agreement between the predicted values of the model and the real values and play an important role in evaluating the success of the model. Each metric has its own advantages and disadvantages, and the performance of a model may vary according to different metrics. For example, a model may achieve a high value in the accuracy metric, but may perform lower in the precision or recall metrics. Therefore, it is important to consider more than one metric when evaluating the performance of the model. In addition, the success of the model depends not only on the metrics used, but also on how the training and test data are divided and the distribution of classes in the classification problem. In imbalanced datasets, some classes are more represented than others, and this may affect the performance of the model and the evaluation metrics. In this study, a number of metrics were used to evaluate the performance of deep learning models.

2.3.1. Confusion Matrix

Confusion matrix is an important tool used to evaluate the performance of machine learning and deep learning models [33]. It is used to understand the accuracy of the model's predictions and the types of errors, especially in classification problems. The confusion matrix is created by comparing the true class labels with the labels predicted by the model. These values are used to calculate different metrics to evaluate the performance of the model. For example, metrics such as accuracy, precision, recall, and F1-score can be calculated from the values in the confusion matrix. The confusion matrix can also be used to understand the strengths and weaknesses of the model. For example, a high false positive rate indicates that the model is incorrectly classifying too many examples as positive. This information can help determine which areas to focus on for model improvement. The confusion matrix is a standard method used to measure the performance of a classification model on a test dataset where the true values are known [34]. This matrix allows you to analyze the overall accuracy of the model as well as different types of errors. Table 3.1 shows the confusion matrix. In summary, the confusion matrix is a powerful tool used to evaluate and understand the performance of deep learning and machine learning models. By analyzing the accuracy of the model's predictions and the types of errors, it guides the model to improve and make more accurate predictions. The metrics are calculated using the formulas below:

$$Accuracy = \frac{TP + TN}{TP + FP + FN + TN} \tag{1}$$

$$Recall = \frac{TP}{TP + FN} \tag{2}$$

$$Precision = \frac{TP}{TP + FP} \tag{3}$$

$$Specificity = \frac{TN}{TN + FP} \tag{4}$$

$$F1 - score = \frac{2 * recall * precison}{recall + precison} \tag{5}$$

2.3.2. ROC Curve

The ROC curve (Receiver Operating Characteristic) is a graphical tool used to evaluate the performance of a classification model in binary classification problems [35]. This curve shows the relationship between the true positive rate (sensitivity) and the false positive rate (1 - specificity) at different threshold values. The true positive rate is the rate at which samples that are actually positive are correctly classified as positive. The false positive rate is the rate at which samples that are actually negative are incorrectly classified as positive. In the ROC curve, the x-axis shows the false positive rate, and the y-axis shows the true positive rate. The curve consists of the union of points connecting the sensitivity and specificity values obtained at different threshold values. The area under the ROC curve (AUC) is a value that summarizes the overall performance of the model. If the AUC value is greater than 0.5, the model is better than random guessing. The performance of the model increases as the AUC value approaches 1. The advantages of the ROC curve include being independent of the threshold value, being robust to class imbalances, and being able to be used to compare the performances of different classification models [36]. The ROC curve is used in many fields, such as medicine, finance, and machine learning. In summary, the ROC curve is a powerful tool that visually shows how well a classification model does, i.e. how well it distinguishes positive examples from negative examples. Although the ROC curve is a visual representation, it has mathematical calculations underlying it. These calculations are based on the concepts of true positive rate (TPR) and false positive rate (FPR). While TPR shows how accurately positive class examples are predicted, FPR shows the rate at which negative class examples are falsely classified as positive.

III. Results and Discussion

ResNet models trained with the same hyperparameter values are given in Table 1. These values are the ones that provide the best performance obtained as a result of various experiments. The parameters shown in Table 2 belong to these models.

TABLE I
GENERAL RESNET MODEL ARCHITECTURES

```
x=base_model.output (ResNet18, ResNet50,
   Resnet152, ResNeXt)
x =Global Average Pooling 2D()(x)
x =Dropout (0.3) (x)
x =Dense (128, activation='relu')(x)
   predictions = Dense(7, activation='softmax')(x)
```

TABLE II
RESNET MODELS HYPERPARAMETER VALUES

Hyperparameters	Value
batch_size	64
epoch	100
Metrics	accuracy', 'loss', 'precision', 'recall', 'AUC', 'f1_score'
optimizer	Adam
learning_rate	0.001
loss function	SparceCategoricalCrossentropy

In this study, four different ResNet models, namely, ResNet18, ResNet50, ResNet152 and ResNeXt, were used for the detection of potato diseases, and their performances were compared. The images were modeled for 3 classifications of potato leaves: early blight, late blight and healthy. The accuracy and performance metrics collected for the dataset are given in Table 3 and Table 4, respectively.

TABLE 3
ACCURACY RATES OF THE RESNET MODELS

Model	Resnet18	Resnet50	Resnet152	ResNeXt
Accuracy	0.96	0.97	0.98	0.98

When Table 3 is examined, it is seen that nearly all models generally showed high accuracy performance. ResNet152 and ResNeXt models achieved the highest success with an accuracy rate of 98%, and Resnet18 and Resnet50 with 96% and 97% accuracy, respectively.

TABLE 4
PERFORMANCE METRICS OF THE RESNET MODELS

Model	Early Blight		
	Precision	Recall	F1-score
Resnet18	0.96	0.99	0.98
Resnet50	0.97	0.99	0.98
Resnet152	0.99	0.99	0.99
ResNeXt	0.99	0.99	0.99
Model	Late Blight		
	Precision	Recall	F1-score
Resnet18	0.99	0.92	0.95
Resnet50	0.97	0.97	0.97
Resnet152	0.98	0.97	0.97
ResNeXt	0.99	0.97	0.98
Model	Healthy		
	Precision	Recall	F1-score
Resnet18	0.84	0.97	0.90
Resnet50	0.97	0.89	0.93
Resnet152	0.88	0.97	0.92
ResNeXt	0.85	0.92	0.88
Model	Model Metric Averages		
	Precision	Recall	F1-score
Resnet18	0.93	0.96	0.94
Resnet50	0.97	0.95	0.96
Resnet152	0.95	0.97	0.96
ResNeXt	0.94	0.96	0.95
Model	Model Weighted Averages		
	Precision	Recall	F1-score
Resnet18	0.96	0.96	0.96
Resnet50	0.97	0.97	0.97
Resnet152	0.98	0.98	0.98
ResNeXt	0.98	0.98	0.98

As can be seen from Table 4, it is noteworthy that ResNet152 and ResNeXt models have higher precision and F1-score values among all the models, especially in the "Late Blight" and "Healthy" classes. This shows that ResNet152 and ResNeXt can learn complex features better and classify potato diseases more accurately thanks to their deeper architectures. ResNet18 and ResNet50 models reached 96% and 97% accuracy rates, respectively. Since these models have fewer layers than ResNet152 and ResNeXt, they can be trained faster. However, their performance may be lower in some cases due to their less depth. When evaluated in terms of metrics, all models achieved very high precision values (96% and above) in the "Early Blight" class. This shows that the models are quite successful in detecting early blight, ResNet18 and ResNet50 exhibited excellent performance with a sensitivity rate of 99% in the "Early Blight" class. Sensitivity values for other classes are also generally high. All models achieved high F1-score values for different classes.

As a result, the conducted tests show that deep learning models can be used effectively for the detection of potato diseases. ResNet152 and ResNeXt models achieved the highest accuracy rates and also performed successfully in terms of other metrics. It can be seen that the Resnet18 model lags behind other models. These models can be valuable tools for disease detection and control in potato cultivation.

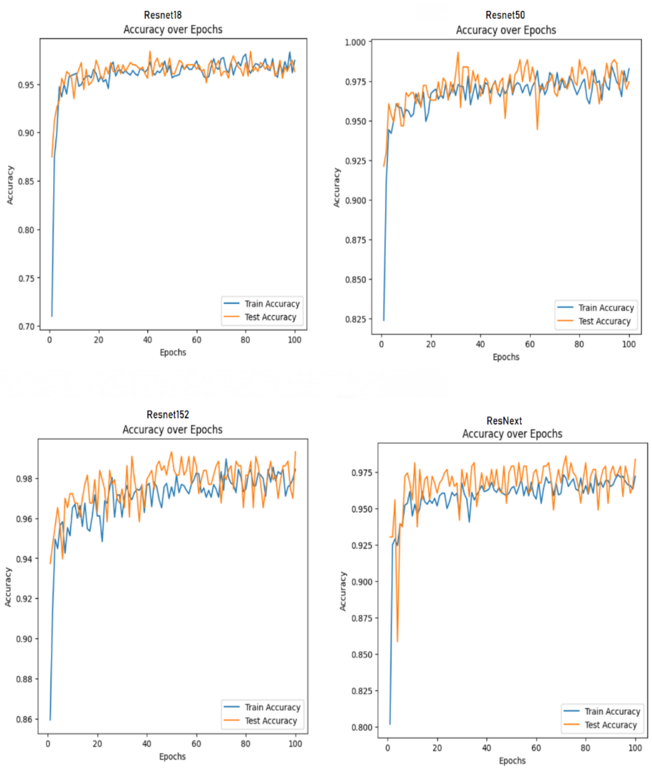

Fig.4. Accuracy of the ResNet models

The results obtained by the models during the training process are examined in detail with the accuracy graphs demonstrated in Fig. 4, and the loss graphs in Fig. 5. These graphs depict how well the models learned and how many errors is made in each iteration. The models were trained for 100 epochs, and early stopping was implemented to monitor validation loss with a patience of 5 epochs, preventing overfitting. The best model weights are saved based on validation loss improvements.

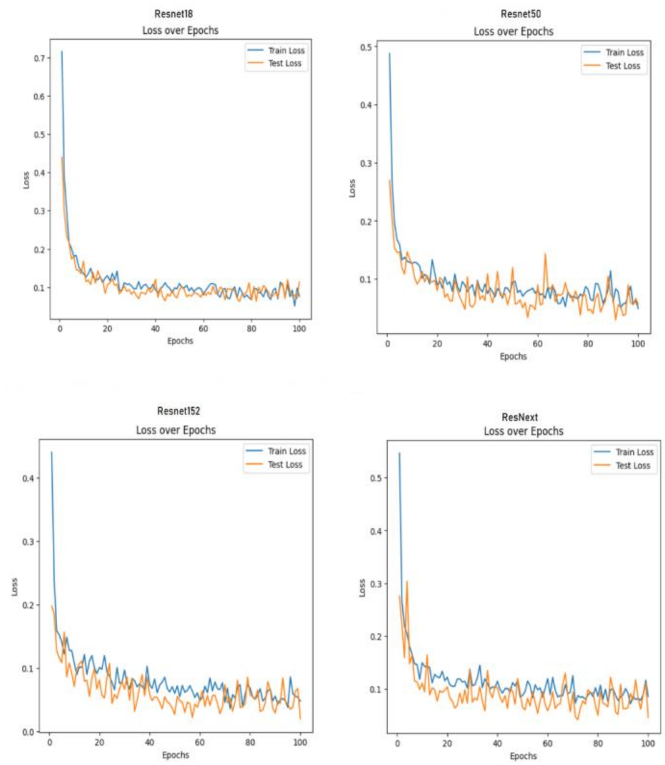

Fig.5. Loss graphs of the ResNet models

The ROC graphs of the models are demonstrated in Figure 6. According to the given graphs, the ROC curves being close to the upper left corner and the AUC values being 1.00 show that the models can distinguish diseases almost perfectly. This shows that the models have both high sensitivity (true positive rate) and high specificity (true negative rate). In other words, they can both correctly detect diseased potatoes and do not misclassify healthy potatoes as diseased. The ROC curves of the ResNet18, ResNet50 and ResNet152 models almost completely overlap and the AUC values are 1.00, indicating that the models are very close to each other in terms of performance. The ROC curve of the ResNeXt model is also very close to the upper left corner and the AUC value is 1.00. However, the curve looks a little "flatter" compared to the other three models. This may indicate that ResNeXt may make slightly more false positive predictions than the other models in some cases.

The confusion matrix can also be used to understand the strengths and weaknesses of the model. For instance, a high false positive rate indicates that the model is incorrectly classifying too many examples as positive. This information can help determine which areas to focus on for model improvement. Figure 7 shows the confusion matrix plots of the evaluated models.

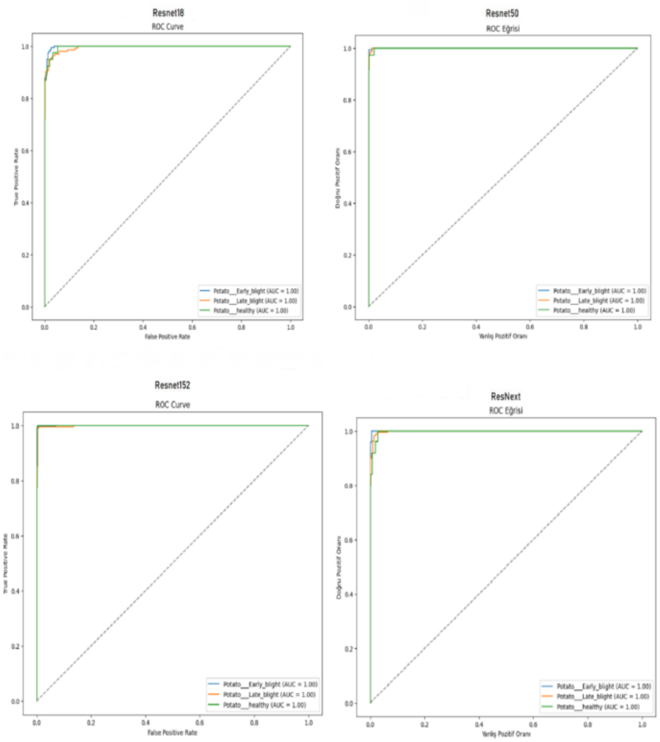

Fig.6. ROC curves of the ResNet Models

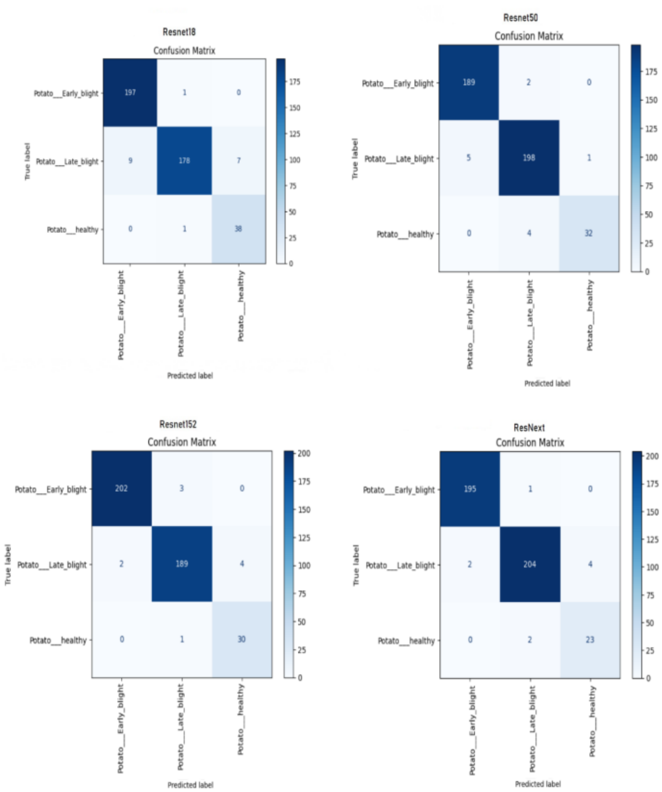

Fig.7. Confusion Matrices

ResNet18 shows high performance in the "Early Blight" class, while making some mistakes in the "Late Blight" and "Healthy" classes. ResNet50 shows a similar performance to ResNet18, but it is observed that it makes fewer mistakes in the "Healthy" class. ResNet152 achieves high accuracy rates for all classes, and it is especially successful in the "Late Blight" class. ResNeXt shows a similar performance to ResNet152, but it makes slightly more mistakes in the "Early Blight" class. In general, it can be said that all models are successful in detecting the "Early Blight" class, but have some difficulties in distinguishing the "Late Blight" and "Healthy" classes. ResNet152 and ResNeXt show higher performance than other models thanks to their deeper architectures.

IV. Conclusion

The aim of this study is to automatically diagnose and classify diseases in potato plants using visual data using deep learning methods. For this purpose, four different models based on ResNet architecture (ResNet18, ResNet50, ResNet152 and ResNeXt) were used and the performances of the models were evaluated with various metrics. The tests performed showed that deep learning models can be used effectively in the detection of potato diseases and that deep models such as ResNet152 and ResNeXt in particular can provide high accuracy rates. It is seen that the highest success rate for the training set belongs to ResNet152 and ResNeXt models. The success rate of both models achieved the highest success with an accuracy rate of 98%. Although the ResNet18 model has lower performance compared to other models, it can be preferred for some applications because it can be trained faster and requires less computational resources. These models show that deep learning methods can achieve high accuracy rates in the detection of potato diseases. In particular, it has been revealed that the deep structure of the ResNet architecture is an effective strategy in learning complex patterns and features. These results show that deep learning methods have high potential for the detection of potato diseases and can be valuable tools to increase efficiency and quality in agricultural production. Besides, running these models on a blended dataset which is prepared from many different datasets, is considered to be as a useful future study.

References

[1] Saleem, M. H., Potgieter, J., & Arif, K. M. (2019). Plant disease detection and classification by deep learning. *Plants*, *8*(11), 468.

[2] Ferentinos, K. P. (2018). Deep learning models for plant disease detection and diagnosis. *Computers and electronics in agriculture*, *145*, 311-318.

[3] Shoaib, M., Shah, B., Ei-Sappagh, S., Ali, A., Ullah, A., Alenezi, F., ... & Ali, F. (2023). An advanced deep learning models-based plant disease detection: A review of recent research. *Frontiers in Plant Science*, *14*, 1158933.

[4] Ramesh, S., Hebbar, R., Niveditha, M., Pooja, R., Shashank, N., & Vinod, P. V. (2018, April). Plant disease detection using machine learning. In *2018 International conference on design innovations for 3Cs compute communicate control (ICDI3C)* (pp. 41-45). IEEE.

[5] Shruthi, U., Nagaveni, V., & Raghavendra, B. K. (2019, March). A review on machine learning classification techniques for plant disease detection. In *2019 5th International conference on advanced computing & communication systems (ICACCS)* (pp. 281-284). IEEE.

[6] Ahmed, I., & Yadav, P. K. (2023). Plant disease detection using machine learning approaches. *Expert Systems*, *40*(5), e13136.

[7] Omid, M., Khojastehnazhand, M., & Tabatabaeefar, A. (2010). Estimating volume and mass of citrus fruits by image processing technique. *Journal of food Engineering*, *100*(2), 315-321.

[8] Chauhan, K., Jani, S., Thakkar, D., Dave, R., Bhatia, J., Tanwar, S., & Obaidat, M. S. (2020, March). Automated machine learning: The new wave of machine learning. In *2020 2nd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 205-212). IEEE.

[9] Blasco, J., Aleixos, N., & Moltó, E. (2003). Machine vision system for automatic quality grading of fruit. *Biosystems engineering*, *85*(4), 415-423.

[10] Gómez, D., Salvador, P., Sanz, J., & Casanova, J. L. (2019). Potato yield prediction using machine learning techniques and sentinel 2 data. *Remote Sensing*, *11*(15), 1745.

[11] Kurek, J., Niedbała, G., Wojciechowski, T., Świderski, B., Antoniuk, I., Piekutowska, M., ... & Bobran, K. (2023). Prediction of potato (solanum tuberosum l.) yield based on machine learning methods. *Agriculture*, *13*(12), 2259.

[12] Gold, K. M., Townsend, P. A., Herrmann, I., & Gevens, A. J. (2020). Investigating potato late blight physiological differences across potato cultivars with spectroscopy and machine learning. *Plant Science*, *295*, 110316.

[13] Kadam, S. U., Dhede, V. M., Khan, V. N., Raj, A., & Takale, D. G. (2022). Machine learning methode for automatic potato disease detection. *NeuroQuantology*, *20*(16), 2102-2106.

[14] Kamilaris, A., & Prenafeta-Boldú, F. X. (2018). Deep learning in agriculture: A survey. *Computers and electronics in agriculture*, *147*, 70-90.

[15] Lee, H. S., & Shin, B. S. (2020). Potato detection and segmentation based on mask R-CNN. *Journal of Biosystems Engineering*, *45*, 233-238.

[16] Bangari, S., Rachana, P., Gupta, N., Sudi, P. S., & Baniya, K. K. (2022, February). A survey on disease detection of a potato leaf using cnn. In *2022 Second International Conference on Artificial Intelligence and Smart Energy (ICAIS)* (pp. 144-149). IEEE.

[17] Rozaqi, A. J., & Sunyoto, A. (2020, November). Identification of disease in potato leaves using Convolutional Neural Network (CNN) algorithm. In *2020 3rd International Conference on Information and Communications Technology (ICOIACT)* (pp. 72-76). IEEE.

[18] Xi, R., Hou, J., & Lou, W. (2020). Potato bud detection with improved faster R-CNN. *Transactions of the ASABE*, *63*(3), 557-569.

[19] Gao, W., Xiao, Z., & Bao, T. (2023). Detection and identification of potato-typical diseases based on multidimensional fusion Atrous-CNN and hyperspectral data. *Applied Sciences*, *13*(8), 5023.

[20] Singh, M. K., Kumar, J., Vishwakarma, S., Gupta, V. K., Raghuwanshi, G., Joshi, K., ... & Mishra, K. (2024, March). Identification of Potato Leaves Diseases Using CNN with Transfer Learning. In *2024 2nd International Conference on Disruptive Technologies (ICDT)* (pp. 6-10). IEEE.

[21] Khobragade, P., Shriwas, A., Shinde, S., Mane, A., & Padole, A. (2022, December). Potato leaf disease detection using cnn. In *2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON)* (pp. 1-5). IEEE.

[22] Griffel, L. M., Delparte, D., & Edwards, J. (2018). Using Support Vector Machines classification to differentiate spectral signatures of potato plants infected with Potato Virus Y. *Computers and electronics in agriculture*, *153*, 318-324.

[23] Islam, M., Dinh, A., Wahid, K., & Bhowmik, P. (2017, April). Detection of potato diseases using image segmentation and multiclass support vector machine. In *2017 IEEE 30th canadian conference on electrical and computer engineering (CCECE)* (pp. 1-4). IEEE.

[24] Ji, Y., Sun, L., Li, Y., Li, J., Liu, S., Xie, X., & Xu, Y. (2019). Non-destructive classification of defective potatoes based on hyperspectral imaging and support vector machine. *Infrared Physics & Technology*, *99*, 71-79.

[25] Shen, D., Zhang, S., Ming, W., He, W., Zhang, G., & Xie, Z. (2022). Development of a new machine vision algorithm to estimate potato's shape and size based on support vector machine. *Journal of Food Process Engineering*, *45*(3), e13974.

[26] Weiss, K., Khoshgoftaar, T. M., & Wang, D. (2016). A survey of transfer learning. *Journal of Big data*, *3*, 1-40.

[27] Kulkarni, U., Meena, S. M., Gurlahosur, S. V., & Mudengudi, U. (2019, September). Classification of cultural heritage sites using transfer learning. In *2019 IEEE fifth international conference on multimedia big data (BigMM)* (pp. 391-397). IEEE.

[28] Venkatesan, R., & Li, B. (2017). Convolutional neural networks in visual computing: a concise guide. CRC Press.

[29] Zhang, A., Lipton, Z. C., Li, M., & Smola, A. J. (2023). *Dive into deep learning*. Cambridge University Press.

[30] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *nature*, *521*(7553), 436-444.

[31] Scherer, D., Müller, A., & Behnke, S. (2010, September). Evaluation of pooling operations in convolutional architectures for object recognition. In *International conference on artificial neural networks* (pp. 92-101). Berlin, Heidelberg: Springer Berlin Heidelberg.

[32] He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778).

[33] Narkhede, S. (2018). Understanding confusion matrix. *Towards Data Science*, *180*(1), 1-12.

[34] Powers, D. M. (2020). Evaluation: from precision, recall and F-measure to ROC, informedness, markedness and correlation. *arXiv preprint arXiv:2010.16061*.

[35] Bento, C. (2022). ROC analysis and the AUC—Area under the curve. Towards Data Science.

[36] Cook, N. R. (2007). Use and misuse of the receiver operating characteristic curve in risk prediction. Circulation, 115(7), 928-935.

BIOGRAPHIES

Abdulkerim Öztekin was born in Berlin, Germany in 1978. He received the B.S. degree in Electrical and Electronics Engineering from Hacettepe University, Ankara, Turkiye, in 2001, and the M.S. and the Ph.D. degrees in Electrical and Electronics Engineering from Gaziantep University, Gaziantep, Turkiye, in 2018.

He is currently working as an Assistant Professor at Batman University and also serves as the General Manager of Batman University Technology Development Zone. He has several publications in the field and his research interests include signal processing, artificial intelligence, optimization, video coding, and communications.

Kenan Almas received the B.S. degree in Department of Mathematics from Dicle University, Diyarbakır, Turkiye, in 2005 and the M.S. degree in Department of Information Technologies from Batman University, Batman, Turkiye, in 2025. He is currently working as a primary school mathematics teacher at Batman Science and Art Center affiliated with the Ministry of National Education.