International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056
Volume: 11 Issue: 05 | May 2024 www.irjet.net p-ISSN: 2395-0072

Classification of Plant Species from Microscopic Plant Cell Images Using Machine Learning Methods

Celalettin Arslan[1], Volkan KAYA[2]

[1] Master's Student, Graduate School of Natural and Applied Sciences, Department of Artificial Intelligence and Robotics, Erzincan Binali Yıldırım University, Erzincan, Türkiye
[2]Associate Professor, Faculty of Engineering and Architecture, Department of Computer Engineering, Erzincan Binali Yıldırım University, Erzincan, Türkiye

---***---

Abstract - *Understanding plant biology and classification of plant species stands out as an important issue in the field of biology. In recent years, with advancements in the field of artificial intelligence, the use of artificial intelligence for plant classification has increased. In this context, plant leaf images have begun to be examined with artificial intelligence. However, in the classification of plant species using artificial intelligence, the use of cell images may provide more accurate and reliable results compared to leaf images. Cell images allow for a closer focus on the genetic structure and fundamental characteristics of the plant, whereas leaf images may be more sensitive to environmental variability. Therefore, in plant classification using artificial intelligence, analyses based on cell images are preferred. In this study, microscopic cell images of four different plant species (Ficus Benjamin, Spathiphyllum, Ficus Elastica and Anthurium) were classified using machine learning methods such as KNN, SVM, Logistic Regression, Decision Trees and Random Forest. In order to classify plant species, a new data set consisting of microscopic cell images of four different plant species was created. Using this data set, plant species were classified with five different machine learning methods and their success accuracies were compared. As a result of the comparison, the best plant species classification was obtained by Random Forest with a success rate of 96.74%, and the worst plant species classification was obtained by the KNN method with a success rate of 86.05%. According to the results obtained, it was seen that microscopic plant cell images were successfully classified using machine learning methods.)*

Key Words: Machine Learning, KNN, SVM, Logistic Regression, Decision Trees, Random Forest, Plant Classification

1.INTRODUCTION

Nowadays, the use of artificial intelligence is increasing. One of the areas where artificial intelligence is most used is image classification studies [1][2][3][4]. During the image classification process, artificial intelligence applications can detect qualities in images that are difficult to detect by the human eye. Thanks to these detections, the machine trains itself and decides which class the new images should belong to, with the help of different features in different image classes. Image classification studies have gained significant momentum with the rapid development of artificial intelligence [5]. These studies aim to increase accuracy rates by replacing humans in various fields. However, in a complex field such as plant classification, the integration of artificial intelligence is of great importance in order to obtain more consistent results that are not user-dependent. Artificial intelligence-supported systems can provide more reliable and consistent results by minimizing human error. Artificial intelligence can be used as an analysis method to automatically measure plant traits to aid genetic discoveries [6]. When studies on plant species classification in the literature are examined, plant leaf images are generally used as the dataset. Many image classification studies carried out to classify plants for various purposes show that classification is achieved with high success rates if appropriate data is used.

2. LITERATURE SURVEY

Wu et al. present a neural network approach to recognize plant leaves in their study. The computer can automatically classify leaf images of 32 different plant species loaded from digital cameras or scanners. For this purpose, the PNN (Probabilistic Neural Network) method, which attracts attention with its fast training process and simple structure, was preferred. 12 features were extracted, which were processed by PCA(Principal Component Analysis) to generate the input vector. Experimental results show that the algorithm can work with more than 90% accuracy. Compared to other methods, this algorithm has been determined to be fast, effective and easy to implement [7].

In Doğan and Türkoğlu's research, deep learning methods were used to classify plant leaves and the performance of these methods was evaluated. They used a total of 7628 leaf images from 32 plant classes in their study. Using deep learning algorithms such as GoogleNet, AlexNet, ResNet50, Vgg16, Vgg19, they achieved successful results in the range of 97.77%-99.72% in classifying plant leaves [8].

In their study, Yaman and Tuncer aimed to detect leaf diseases using deep learning and feature selection methods using 726 images of walnut leaves. Images were divided into two classes: healthy and diseased, 17 different deep learning models were evaluated and DarkNet53 and ResNet101 were selected as the two models with the highest performance. By combining the features of these models, feature extraction was carried out with a hybrid approach and features were selected using the ReliefF algorithm. SVM algorithm, a machine learning method, was used to classify the selected features and 99.58% accuracy was achieved. This study provides an effective tool for early detection of plant diseases [9].

In their study by Shruthi et al., the general stages of machine learning classification techniques for plant disease detection were presented and a comparative study was conducted. In this study, machine learning methods such as KNN, SVM and deep learning methods were used, and it was observed that CNN (Convolutional Neural Networks) provided high accuracy and detected more diseases in more than one product [10].

Sabancı et al. in their study, they conducted a research on the classification of leaves using a data set containing 14 different physical properties of 12 different leaf species. The classification success of leaves was examined using various (KNN, MLP, NaiveBayes, RBFNetworks, KStar, J48) data mining, machine learning and deep learning algorithms. As a result of the classifications, success rates between 86.52% and 94.32% were obtained. The results of the study show that artificial neural networks can be used effectively in classifying plant leaves and that MLP performs better than other classification algorithms [11].

Priya et al. in their study, a research was conducted on plant classification processes by using a data set containing leaf images of 15 different plant classes and classifying these data with SVM and KNN machine learning methods. As a result of the research, it was determined that the results obtained from SVM provide more successful results compared to the KNN method [12].

Significant success rates have been achieved by using many artificial intelligence methods in the literature review regarding the classification of plants. In many plant classification studies in the literature, plant leaves have been used, but it seems that the classification process based on plant cells is insufficient. For this purpose, in the process of classifying plants, the characteristics of each plant species should be determined and it should be decided which features should be taken into consideration for classification. However, if the data examined are similar to each other, determining the characteristics of the species makes the classification process difficult. Therefore, the main aim of this study is to ensure successful classification of cell images of different plant species using machine learning methods. In this way, it will contribute to a better understanding of plant biology, accelerate the identification process of plant species and make a significant contribution to studies on the conservation of biological diversity. This research will make a significant contribution to the field of plant cell classification by adopting an interdisciplinary approach including biology, artificial intelligence, and data science. The achievements to be achieved will provide significant support to efforts to preserve biodiversity by allowing faster and more effective research in fields such as plant biology, agriculture and environmental sciences. For this reason, 4 different plant species were classified using a new dataset consisting of microscopic cell images with the KNN, SVM, Logistic Regression, Decision Trees and Random Forest methods discussed in the study. Using this dataset, the classification experimental results of five different machine learning methods were compared. The contributions of this study can be summarized as follows:

1. Faster, more precise and more effective results will be obtained in the classification of plant species than traditional methods.

2. Ensuring classification accuracy by using microscopic plant cell images in plant biology research will allow the discovery of new scientific information.

3. It will offer application potential in understanding plant biology and classifying plant species in many fields such as agriculture, biotechnology and natural resource management.

4. By using microscopic cell images instead of leaf images in the classification of plant species, it will be possible to classify with cell images obtained from a small part of the plant's leaf, even if we do not have the entire leaf.

5. It will be possible to obtain more precise results from two different sources with classification studies on both leaf images and microscopic cell images.

3. MATERIAL AND METHODS

3.1. Dataset and Image Pre-processing

Microscopic cell images of ficus benjamin, spathiphyllum, ficus elastica, and anthurium plants were used in the study. While creating the dataset, sections were taken from the leaves of each plant in the study and different images were obtained under the microscope using 40/0.65 - 160/0.17 objective zoom and 10x eyepiece size. After these images were taken with a high-resolution camera, sections with dimensions of 256×256 pixels were taken during the data pre-processing stage. The data obtained was used as 80% training data and 20% as test data to be used in machine learning methods. Images of the data set used are given in Figure 1.

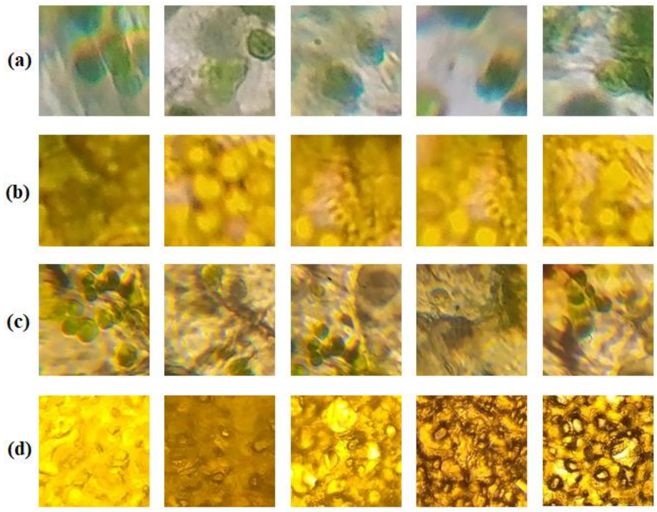

Figure 1: Sample data set image obtained according to plant species (a) ficus benjamin, (b) spathiphyllum, (c) ficus elastica, (d) anthurium

The obtained images were divided into training and test data in the numbers specified in Table 1 to be used with KNN, SVM, Logistic Regression, Decision Trees and Random Forest machine learning methods.

Table 1: Numbers of training and test images separated by plant species

Plant Type	Train (%80)	Test (%20)	Total
Ficus benjamin	1252	313	1565
Spathiphyllum	552	138	690
Ficus elastica	584	146	730
Anthurium	680	170	850
Total	**3068**	**767**	**3835**

3.2. Machine Learning Methods

In areas such as image classification, machine learning methods can quickly process large data sets and classify them successfully [2]. Well-trained models tend to provide more accurate results than humans and can be used in a variety of industries, as well as offering flexibility with their scalability and ability to adapt to different types of data [13]. Since they can be continuously improved, they increase classification accuracy and improve the performance of systems.

3.2.1 KNN (K-Nearest Neighbor)

KNN algorithm is a widely used machine learning method in the literature. In this method, a feature is selected and then it classifies the samples within the k neighbors closest to this feature. Various metrics are used in proximity calculation. When calculating the distance between new data and existing data, methods such as Manhattan or Euclidean are preferred [14]. According to the basic operation, in order to classify an object, its nearest neighbors are examined. Among these neighbors, it is inferred that the neighbors with the most common features and the examined object are in the same class. The number of neighbors (k) is determined by the user. The user evaluates the performance of the model by trying different numbers of neighbors. It tries to determine the most suitable number of neighbors by measuring the accuracy rate of the model. The number of neighbors is usually determined as an odd number, so that there is no equality between neighbors in the classification process [15]. The general structure of the KNN algorithm is given in Figure 2.

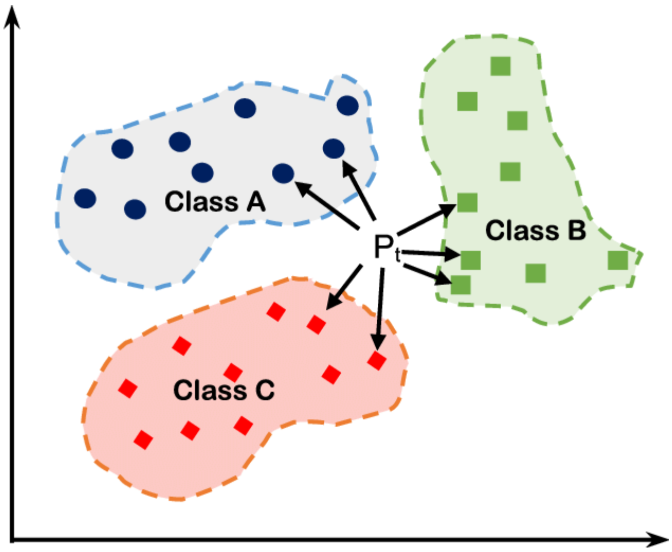

Figure 2: KNN classification example [16]

3.2.2. SVM (Support Vector Machine)

Vladimir N. Vapnik and Alexey Y. Chervonenkis began working in the fields of learning theory and statistical learning in the Soviet Union in the early 1960s. During this period, they laid the foundations of SVM by concentrating on basic concepts. Vapnik's book "The Nature of Statistical Learning Theory", published in 1995, is an important resource covering the modern formulation and foundations of SVM [17]. In 1995, Vapnik and Cortes published a paper called "Support-Vector Networks" [18]. This article described the modern formulation of SVM, demonstrating its effectiveness in classification problems and increasing the popularity of SVM. In this way, SVM has become an important machine learning algorithm over time. SVM is one of the developed machine learning algorithms and also provides solutions to classification problems. SVM is a machine learning algorithm widely used in the field of pattern recognition and classification problems [19][20]. The general structure of the SVM algorithm is given in Figure 3.

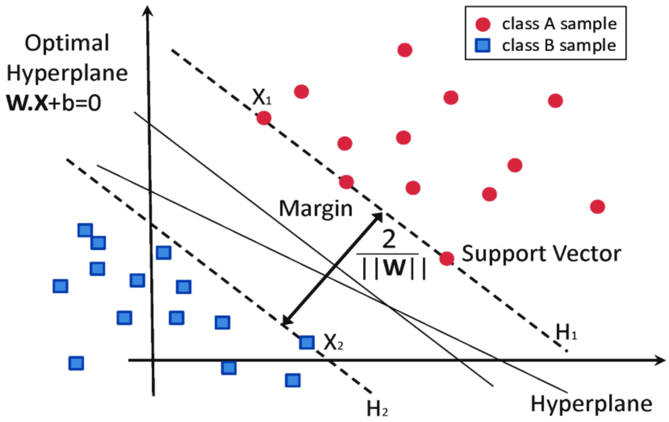

Figure 3: SVM classification example [21]

The basic principle of SVM is to determine the solution that can best linearly separate the data [22]. Non-linear data is analyzed by moving it to another dimension using the transformation technique. The number of samples used in SVMs is negligible. SVM can classify previously undetermined data during learning. This is possible with its generalization ability and is the most important feature that distinguishes it from other methods. Today, SVM plays an important role in data mining, gene analysis, and pattern recognition and classification problems, enabling accurate results to be obtained [23].

3.2.3. Logistic Regression

Logistic Regression is an algorithm used to determine the effect of independent variables on a set of dependent variables and to make predictions and classifications accordingly [24]. Dependent variables can structurally take continuous or categorical values. In classification problems, the dependent variable is generally used categorically and with two levels. Logistic regression analysis is preferred when the dependent variable is two-level [25].

In addition, the model to be created with this method aims to realize the relationship between dependent and independent variables with the least possible variables and the best fit [26].

3.2.4. Decision Trees

Decision trees are a machine learning method often used to solve classification and regression problems. Decision trees use the features in the dataset to create decision rules and classify or evaluate data points using this rule set [27]. Decision trees are used in many industries and applications, especially in classification and prediction problems. The basic structure of the decision tree consists of three main components: nodes, branches and leaves. Each feature is expressed by a node, and the parts between the root and the leaves form the branches [28]. An example of a decision tree is given in Figure 4.

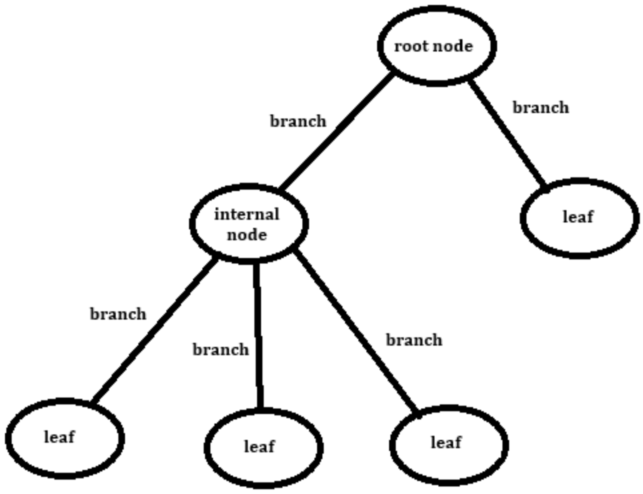

Figure 4: General structure of the decision trees

3.2.5. Random Forest

Random forest algorithm is one of the prominent ones among successful classification methods [29]. Random forest is an ensemble learning method frequently used for classification and regression problems. In this method, more than one model is brought together to create a more powerful and generalizable model. The random forest algorithm first analyzes the classes of training data and then makes a conclusion based on decision trees to predict the classes of test data [30]. Additionally, when splitting at each node, randomly selected features are used to create the decision tree. Randomly selecting these features ensures that each tree is different from each other. The main principle of Random Forest is to make a stronger prediction by combining the predictions of each tree, using the ability of each decision tree to make predictions alone [31]. For classification problems, these predictions are often combined by a voting method: each tree makes a prediction for a class and determines the class with the most votes as the final prediction.

Although the Random Forest method is specific to decision trees, it is a generally valid approach for all classifiers, and being fast and being able to handle many variables as input is an important advantage [32].

4. RESULTS AND DISCUSSION

Classification results of a total of 3835 plant cell images belonging to 4 plant classes were obtained using KNN, SVM, Logistic Regression, Decision Trees and Random Forest machine learning methods. These results were analyzed supported by performance tables and confusion matrices. In the performance tables, the model's accuracy, precision, recall, f1-score, support, macro average and weighted average results were calculated. The confusion matrix is a table in which the results obtained with the developed classification model are compared numerically with the actual values. All results in the performance table can be calculated using the data in this matrix.

According to the experimental results obtained, an accuracy value of 86.05% was obtained with the KNN method. In the KNN algorithm, the number of neighbors that gave the most successful results was determined to be 3 by trial and error method. Classification performance results are given in Table 2 and the confusion matrix obtained according to the prediction results is given in Figure 5.

Table 2: Performance results obtained with the KNN method

Plant Type	Precision	Recall	F1-Score	Support
ficus benjamin	1.00	1.00	1.00	313
spathiphyllum	0.57	0.98	0.72	138
ficus elastica	0.99	0.93	0.96	146
anthurium	0.96	0.45	0.61	170
accuracy			0.86	767
macro avg	0.88	0.84	0.82	767
weighted avg	0.91	0.86	0.86	767

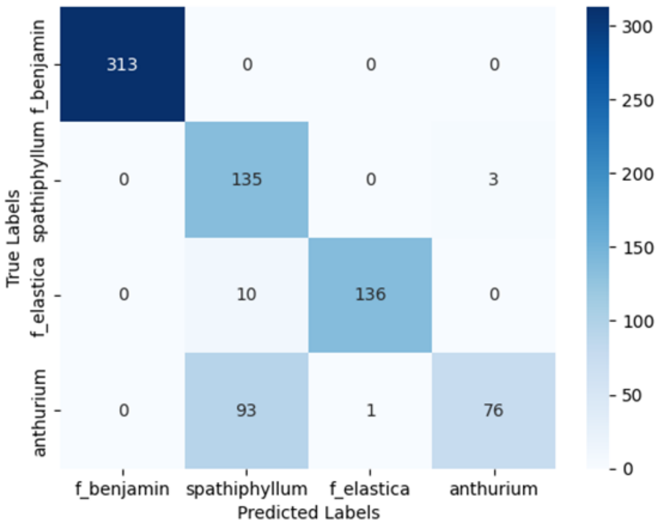

Figure 5: Confusion matrix obtained as a result of classification with the KNN method

According to the experimental results obtained, an accuracy value of 83.70% was obtained with the SVM method. Classification performance results are given in Table 3 and the confusion matrix obtained according to the prediction results is given in Figure 6.

Table 3: Performance results obtained with the SVM method

Plant Type	Precision	Recall	F1-Score	Support
ficus benjamin	1.00	1.00	1.00	313
spathiphyllum	0.54	0.88	0.67	138
ficus elastica	0.97	0.95	0.96	146
anthurium	0.81	0.41	0.54	170
accuracy			0.84	767
macro avg	0.83	0.81	0.79	767
weighted avg	0.87	0.86	0.83	767

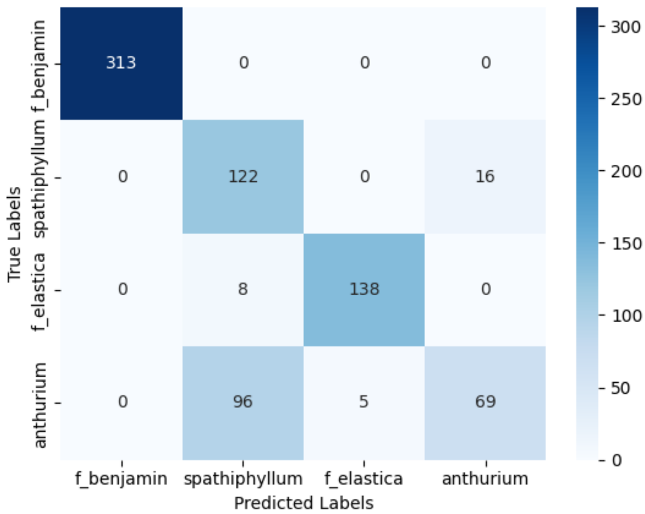

Figure 6: Confusion matrix obtained as a result of classification with the SVM method

According to the experimental results obtained, an accuracy value of 77.84% was obtained with the logistic regression method. Classification performance results are given in Table 4 and the confusion matrix obtained according to the prediction results is given in Figure 7.

Table 4: Performance results obtained by logistic regression method

Plant Type	Precision	Recall	F1-Score	Support
ficus benjamin	0.99	0.92	0.95	313
spathiphyllum	0.51	0.86	0.64	138
ficus elastica	0.81	0.87	0.84	146

anthurium	0.75	0.37	0.50	170
accuracy			0.78	767
macro avg	0.76	0.75	0.73	767
weighted avg	0.81	0.78	0.77	767

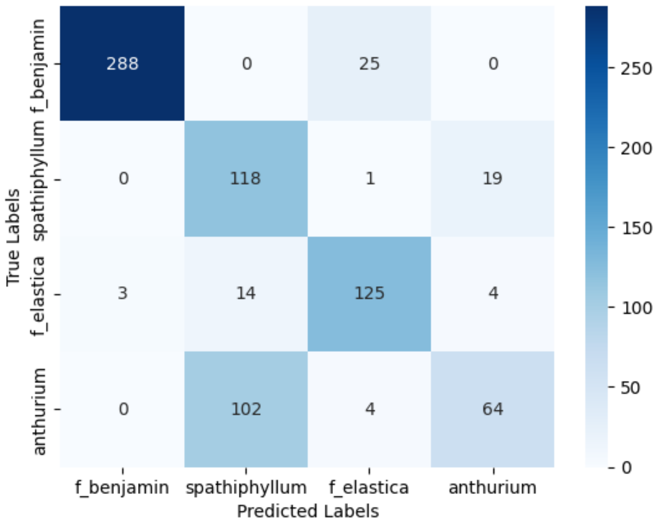

Figure 7: Confusion matrix obtained as a result of classification with the logistic regression method

According to the experimental results obtained, an accuracy value of 80.83% was obtained with the decision trees method. Classification performance results are given in Table 5 and the confusion matrix obtained according to the prediction results is given in Figure 8.

Table 5: Performance results obtained with the decision trees method

Plant Type	**Precision**	**Recall**	**F1-Score**	**Support**
ficus benjamin	0.94	0.93	0.93	313
spathiphyllum	0.64	0.75	0.69	138
ficus elastica	0.77	0.75	0.76	146
anthurium	0.77	0.68	0.72	170
accuracy			0.81	767
macro avg	0.78	0.78	0.78	767
weighted avg	0.81	0.81	0.81	767

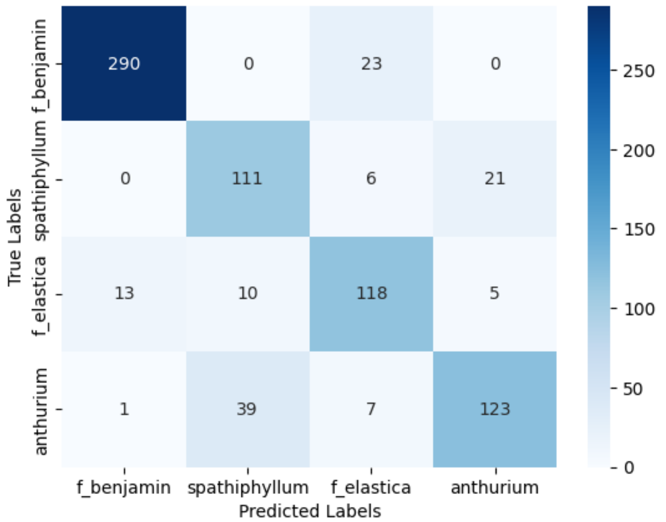

Figure 8: Confusion matrix obtained as a result of classification with the decision trees method

According to the experimental results obtained, an accuracy value of 96.74% was obtained with the random forest method. Classification performance results are given in Table 6, and the confusion matrix obtained according to the prediction results is given in Figure 9.

Table 6: Performance results obtained with the random forest method

Plant Type	**Precision**	**Recall**	**F1-Score**	**Support**
ficus benjamin	0.99	1.00	1.00	313
spathiphyllum	0.89	0.96	0.92	138
ficus elastica	1.00	0.97	0.99	146
anthurium	0.96	0.91	0.94	170
accuracy			0.97	767
macro avg	0.96	0.96	0.96	767
weighted avg	0.97	0.97	0.97	767

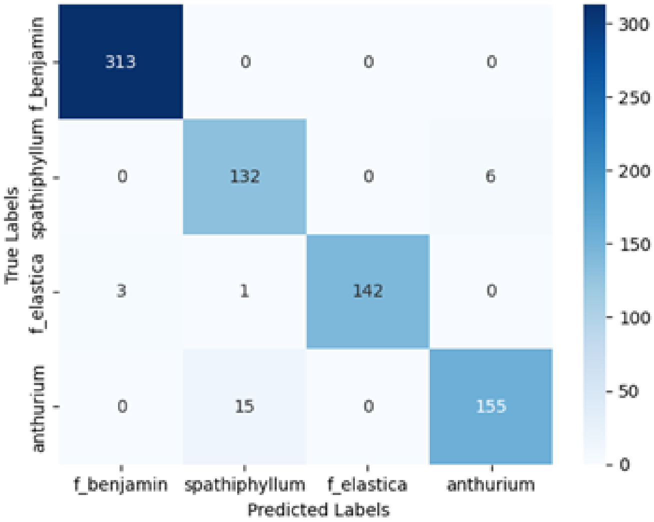

Figure 9: Confusion matrix obtained as a result of classification with the random forest method

In addition, the success accuracy rates of the machine learning methods considered in the study in classifying 4 different microscopic cell images were compared and are given in Table 7.

Table 7: Classification accuracy rates of each machine learning

Machine Learning Method	Classification Accuracy Rate
KNN	%86,05
SVM	%83,70
Logistic Regression	%77,84
Decision Trees	%80,83
Random Forest	%96,74

4. CONCLUSIONS

In this study, microscopic cell images of four different plant species (ficus benjamin, spathiphyllum, ficus elastica, and anthurium) were used to classify the species of plants using machine learning methods such as KNN, SVM, Logistic Regression, Decision Trees and Random Forest. When the experimental results obtained are examined, the highest success rate was achieved with Random Forest, with a success rate of 96.74%, and it can be said that it is very successful for classification processes with artificial intelligence. The results obtained from other methods are in the range of 77.84% - 86.05% success accuracy. Although the results obtained from these methods are low compared to the Random Forest method, they can generally be interpreted as meaningful classification rates. When the results were examined, it was seen that machine learning methods could be used effectively in plant classification based on microscopic cell images of the plant species. In particular, it has been observed that the Random Forest method provides high accuracy and improves classification performance.

As a result, it emphasizes that machine learning methods can play an important role in plant biology research and the necessity of adopting a new perspective in the field of plant species classification. In addition, the success of machine learning in classifying plant cell images also contributes to the development and increased use of artificial intelligence. Researchers can achieve successful results by classifying with artificial intelligence. These studies can guide future researchers and effective classification models can be developed using larger data sets. Additionally, general comparative analyzes can be made using more artificial intelligence methods.

REFERENCES

[1] LeCun, Y., Bottou, L., Bengio, Y., & Haffner, P. (1998). Gradient-based learning applied to document recognition. Proceedings of the IEEE, 86(11), 2278-2324, doi:10.1109/5.726791

[2] Krizhevsky, A., Sutskever, I., & Hinton, G. E. (2012). Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 25. doi:10.1145/3065386

[3] Simonyan, K., & Zisserman, A. (2014). Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, doi:10.48550/arXiv.1409.1556

[4] 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). doi: 10.1109/CVPR.2016.90

[5] Nichols, J. A., Herbert Chan, H. W., & Baker, M. A. (2019). Machine learning: applications of artificial intelligence to imaging and diagnosis. Biophysical reviews, 11, 111-118, doi:10.1007/s12551-018-0449-9

[6] Pound, M. P.; Atkinson, J. A.; Townsend, A. J; Wilson, M. H.; Griffiths, M.; Jackson, A. S.; Bulat, A.; Tzimiropoulos, G.; Wells, D. M.; Murchie, E. H.; Tony, P. P.; Andrew, P.F. (2017). Deep machine learning provides state-of-the-art performance in image-based plant phenotyping. Gigascience, 6 (10), gix083. doi:10.1093/gigascience/gix083

[7] Wu, S. G., Bao, F. S., Xu, E. Y., Wang, Y. X., Chang, Y. F., & Xiang, Q. L. (2007, December). A leaf recognition algorithm for plant classification using probabilistic neural network. In 2007 IEEE international symposium

on signal processing and information technology (pp. 11-16). IEEE, doi:10.1109/ISSPIT.2007.4458016

[8] Doğan, F., & Türkoğlu, İ. (2018). Derin öğrenme algoritmalarının yaprak sınıflandırma başarımlarının karşılaştırılması. Sakarya University Journal of Computer and Information Sciences, 1(1), 10-21.

[9] Yaman, O., & Tuncer, T. (2022). Bitkilerdeki yaprak hastalığı tespiti için derin özellik çıkarma ve makine öğrenmesi yöntemi. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 34(1), 123-132, doi:10.35234/fumbd.982348

[10] 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, doi:10.1109/ICACCS.2019.8728415

[11] Sabancı, K., Ünlerşen, M. F., & Aydın, C. (2016). Bitki yapraklarının veri madenciliği algoritmaları ile sınıflandırılması. Proceedings Book, 525.

[12] Priya, C. A., Balasaravanan, T., & Thanamani, A. S. (2012, March). An efficient leaf recognition algorithm for plant classification using support vector machine. In International conference on pattern recognition, informatics and medical engineering (PRIME-2012) (pp. 428-432). IEEE, doi:10.1109/ICPRIME.2012.6208384

[13] Deng, J., Dong, W., Socher, R., Li, L. J., Li, K., & Fei-Fei, L. (2009, June). Imagenet: A large-scale hierarchical image database. In 2009 IEEE conference on computer vision and pattern recognition (pp. 248-255). Ieee, doi:10.1109/CVPR.2009.5206848

[14] Kılınç, D., Borandağ, E., Yücalar, F., Tunali, V., Şimşek, M., & Özçift, A. (2016). KNN algoritması ve r dili ile metin madenciliği kullanılarak bilimsel makale tasnifi. Marmara Fen Bilimleri Dergisi, 28(3), 89-94, doi:10.7240/mufbed.69674

[15] Cover, T., & Hart, P. (1967). Nearest neighbor pattern classification. IEEE transactions on information theory, 13(1), 21-27, doi:10.1109/TIT.1967.1053964

[16] Atallah, D. M., Badawy, M., El-Sayed, A., & Ghoneim, M. A. (2019). Predicting kidney transplantation outcome based on hybrid feature selection and KNN classifier. Multimedia Tools and Applications, 78, 20383-20407, doi:/10.1007/s11042-019-7370-5

[17] Vapnik, V. (2013). The nature of statistical learning theory. Springer science & business media. doi:10.1007/978-1-4757-2440-0

[18] Cortes, C., & Vapnik, V. (1995). Support-vector networks. Machine learning, 20, 273-297. doi: 10.1007/BF00994018

[19] Ayhan, S. ve Erdoğmuş, Ş. (2014). Destek vektör makineleriyle sınıflandırma problemlerinin çözümü için çekirdek fonksiyonu seçimi. İibf Dergisi, 9(1), 175-198.

[20] Sengur, A. (2019). Ulticlass least-squares support vector machines for analog modulation classification, expert systems with applications, 36(3), 6681-6685, doi:10.1016/j.eswa.2008.08.066

[21] García-Gonzalo, E., Fernández-Muñiz, Z., Garcia Nieto, P. J., Bernardo Sánchez, A., & Menéndez Fernández, M. (2016). Hard-rock stability analysis for span design in entry-type excavations with learning classifiers. Materials, 9(7), 531, doi:10.3390/ma9070531

[22] Sezer, O. G., Erçil, A., & Keskinoz, M. (2005, May). Independent component based 3D object recognition using support vector machines. In Proceedings of the IEEE 13th Signal Processing and Communications Applications Conference, 2005. (pp. 99-102). IEEE. doi: 10.1109/SIU.2005.1567629

[23] Kecman, V. (2001). Learning and soft computing: support vector machines, neural networks, and fuzzy logic models. MIT press.

[24] Cox, D. R. (1958). The regression analysis of binary sequences. Journal of the Royal Statistical Society Series B: Statistical Methodology, 20(2), 215-232, doi:10.1111/j.2517-6161.1958.tb00292.x

[25] Gujarati, D. N., & Porter, D. C. (2009). Basic econometrics. McGraw-hill.

[26] Atasoy, D. (2001). Lojistik regresyon analizinin incelenmesi ve bir uygulaması. Unpublished master's thesis, Sivas Cumhuriyet University, Institute Social Science, Sivas.

[27] Quinlan, J. R. (1986). Induction of decision trees. Machine learning, 1, 81-106. doi:10.1007/BF0011625

[28] Kavzoğlu, T., & Çölkesen, İ. (2010). Karar ağaçları ile uydu görüntülerinin sınıflandırılması. Harita Teknolojileri Elektronik Dergisi, 2(1), 36-45.

[29] Akar, Ö., & Güngör, O. (2012). Classification of multispectral images using random forest algorithm. Journal of Geodesy and Geoinformation, (106), 139-146, doi:10.9733/jgg.241212.1t

[30] Erdem, F., Derinpınar, M. A., Nasirzadehdizaji, R., Selen, O., Şeker, D. Z., & Bayram, B. (2018). Rastgele orman yöntemi kullanılarak kıyı çizgisi çıkarımı İstanbul

örneği. Geomatik, 3(2), 100-107, doi:10.29128/geomatik.362179

[31] Breiman, L. (2001). Random forests. Machine learning, 45(1),5-32 doi: 10.1023/A:1010933404324

[32] Ekelik, H., & Altaş, D. (2019). Dijital reklam verilerinden yararlanarak potansiyel konut alıcılarının rastgele orman yöntemiyle sınıflandırılması. Journal Of Research İn Economics, 3(1), 28-45.