Journal of Physical Chemistry and Functional Materials Volume 7, Issue 2 (2024) 22-26

Journal of Physical Chemistry and Functional Materials

Home Page of Journal: https://dergipark.org.tr/jphcfum

A New Hybrid Method for Classification of Rice Leaf Diseases: SVM+NCA+Resnet50

Harun Bingol[a]* and Serpil Aslan[a]

[a]*Malatya Turgut Ozal University, Department of Software Engineering, Malatya, Türkiye*

* *Corresponding author: E-mail: harun.bingol@ozal.edu.tr*

ABSTRACT

Rice is extremely important for individuals and countries, both in terms of nutritional value and financial value. It is necessary to protect such an important plant from diseases and increase the yield. However, early detection of diseases on plant leaves can prevent the spread of this disease and is also very important in terms of treating the plant. Artificial intelligence has become very popular in recent years thanks to its success in terms of disease classification. Convolutional Neural Network (CNN) architectures used in image classification perform very successful work. Within the scope of this study, it is recommended that the diseases on rice leaves be classified using artificial intelligence techniques, without mixing them with each other, with very high accuracy values, and without any problems caused by humans. With this proposed model, a support vector machine-based model is proposed that classifies five (5) of the most common rice diseases with a very high accuracy of %98.

ARTICLE INFO

Keywords:
Artificial Intelligence,
Classification,
Resnet50,
NCA,
SVM

Received: 2024-06-11
Accepted: 2024-07-20
ISSN: 2651-3080
DOI: 10.54565/jphcfum.1499620

1. Introduction

Rice is a very important plant for the nutrition of individuals, especially in Asia and Africa [1]. It is the agricultural product with the third highest production worldwide, after sugar cane and corn. It is undesirable for a plant with such a serious production level to be affected by diseases and lose productivity. There are some well-known diseases of the rice plant. Many of these diseases can be identified from the leaves of the plant. Classifying these diseases using manual methods can lead to serious time loss and erroneous results [2]. Classifying these diseases also requires serious expert knowledge and experience. It is a known fact that it is difficult to access experts to diagnose this disease in rice in rural areas. It is important not to confuse these diseases observed in rice leaves and to diagnose them correctly, both in terms of choosing the drug to be applied for the treatment of the disease and in terms of the course and spread of the disease. Misdiagnosing the disease may lead to incorrect drug treatment and will not cure the disease. In such a case, there will be loss of time, money, and labor. These are undesirable situations. In order to prevent these situations, artificial intelligence models that can classify diseases very precisely in recent years have been used in pharmaceutical, chemical, health, aviation, defense industries, etc. started to be used.

The performance of CNN, a sub-branch of artificial intelligence, in image classification is known [3]. CNN architectures do not work with vectors due to their nature. Instead, they work with raw images, that is, matrices. For this reason, the features of the images to be classified are automatically extracted without the need for expert knowledge [4]. Some studies on classification of rice leaf with CNN architectures are as follows;

Yucel and Yildirim used a 6-class data set in their study. Using Efficientb0 [5], Shufflenet [6], and Resnet101 [7] deep architectures, they extracted the features of the images in the data set and obtained a feature map by combining these features. They also stated that they used the support vector machine as a classifier in their applications. They reported that they reached the highest accuracy rate of 98% [2].

Ghosal and his colleagues used VGG 16, one of the most well-known CNN architectures, to classify rice leaf diseases [8]. As a result of the experiments, they stated that they achieved the highest classification accuracy of 92.46% [9].

Ahad et al. used 6 different CNN architectures to classify 9 different rice leaf diseases and stated that they achieved the highest accuracy value of 98%. It was stated that the original dataset used in this study contained 900 images and this was increased to 42876 with data augmentation techniques [10].

Bhattacharya et al. stated that in their study, they managed to distinguish diseased and healthy leaves with a 94% accuracy rate by using a data set containing a total of 1500 rice leaf images, including three rice leaf diseases. They also stated that they categorized the diseased leaves with an accuracy rate of 78.44% [11].

In this study, a hybrid model was developed to classify images of rice leaf disease using deep learning architectures. In this model we developed, the feature maps of the Resnet50 deep model were extracted, then the most important features were selected with the Neighborhood Component Analysis (NCA) method and the model was optimized. The optimized feature map was classified using the Support Vector Machine (SVM) classifier.

This article is organized as follows. In Section 2, deep learning architectures and NCA dimensionality reduction methods used in classifying flower images are explained. Additionally, the hybrid model we proposed is given in detail in this section. In Section 3, results and discussions performed to classify flower images are presented. This section comments are made on the performance of the algorithms, and finally, in Section 4, the results are interpreted and information about future applications is given.

2. Material and Method

In this section, the dataset used in the study, CNN architectures, NCA dimensionality reduction method, and the hybrid model we proposed are explained.

Dataset

Rice leaf disease dataset consists of a total of 6 classes, 1 of which is healthy. There are 2627 images in this dataset. The class names in the data set are Bacterial Leaf Blight, Brown Spot, Healthy, Leaf Blast, Leaf Scald, Narrow Brown Spot [12].

CNN Architectures and NCA

During the experiments, state-of-the-art architectures Alexnet, Googlenet, Mobilenetv2, and Resnet50 architectures were used. Alexnet, one of these architectures, won the ImageNet competition in 2012. In addition, it accepts the images in the data set as 227x227 and consists of 25 layers [13]. Googlenet architecture is designed keeping in mind low computational cost [14]. Mobilenetv2 architecture is an architecture that produces very efficient results on devices that are quite limited in terms of hardware features [15]. Resnet50 architecture is the winning architecture of the ImageNet competition held in 2015. In this competition, Resnet50 architecture came first with an error rate of 3.6% [7].

When images in the dataset are used by deep architectures, 1000 features of each image are used. However, using all of these features can often cause extra computational costs and the best features must be selected. In this application, the neighborhood component analysis (NCA) method was used as the feature selection method [3, 16].

Proposed Model

In this study, Resnet50 architecture, which has been very popular in recent years, was used as the base. Thanks to this architecture, different features of the images in the rice leaf disease dataset were obtained. The size of the feature map obtained with this architecture is 2627x1000. The NCA dimensionality reduction method was used to reduce the size of the resulting feature map. After applying the NCA dimension reduction method, the size of the feature map obtained was 2627x156. Finally, the optimized feature map was classified in classical machine learning classifiers. The proposed model is shown in Figure 1.

3. Results and Discussions

In our study, first of all, the data set was divided into 70% training and 30% test data to compare the performance of the model we proposed. Results were obtained on 4 different pre-trained CNN architectures. The accuracy values obtained in these architectures are given in Table 1.

Table 1 Accuracy values (%) from CNN architectures

Alexnet	**Googlenet**	**Mobilenetv2**	**Resnet50**

94.53	92.62	94.40	96.31

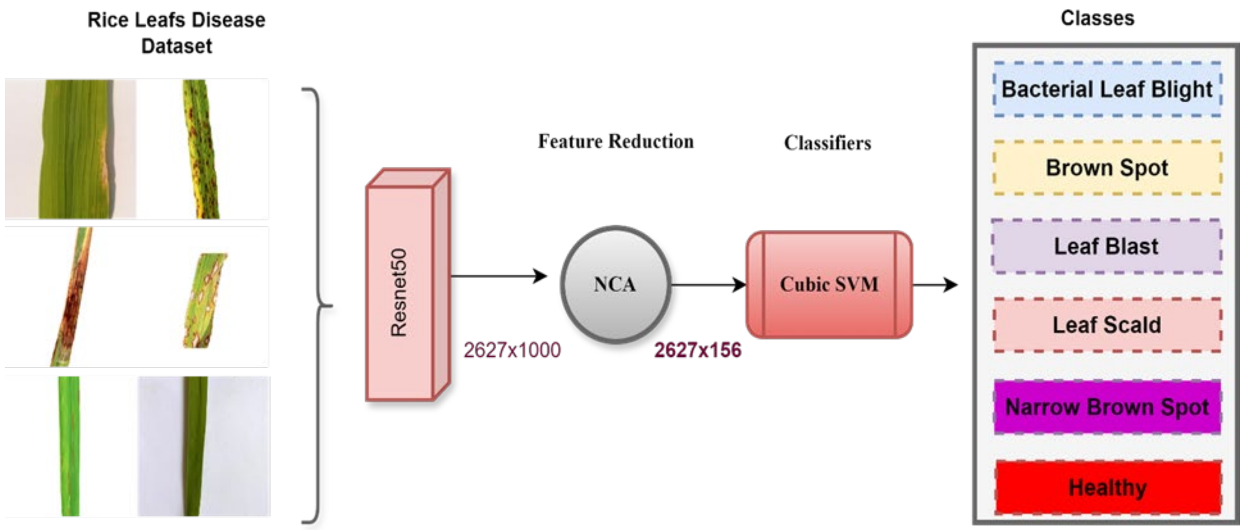

Figure 1 Block diagram of the proposed model

When Table 1 is examined, among 4 different CNN architectures, Resnet50 gave the highest accuracy value of 96.31%, and Googlenet architecture gave the lowest accuracy value of 92.62%. Table 2 shows the complexity matrices showing the accuracy values of 4 different CNN architectures. In Table 2, Bacterial Leaf Blight, Brown Spot, Healthy, Leaf Blast, Leaf Scald, and Narrow Brown Spot classes are values 1, 2, 3, 4, 5, and 6, respectively.

Of the 4 different CNN architectures used in the study, the highest accuracy value was obtained in the Resnet50 architecture. In this architecture, 757 of the 786 rise leaf images were predicted correctly, while 29 were predicted incorrectly by the model.

It is known that machine learning classifiers have been used especially with deep learning methods in recent years and have achieved good results. In the second stage of this study, 1000 features obtained with the Resnet50 architecture were reduced to 156 with the NCA feature selection method. The resulting feature map was classified in the Cubic SVM classifier. The confusion matrix for this process is shown in Table 3.

Table 3 Confusion matrix of the proposed model

Resnet50+NCA+CubicSVM

True Class \ Predicted Class	1	2	3	4	5	6
1	437				1	
2	1	418	3	15		1
3	1	3	432	2		
4		16	4	416		2
5					435	3
6					1	437

Table 2 Confusion matrices of CNN architectures

Alexnet

True Class	1	2	3	4	5	6
1	130				1	
2		123	1	7		
3		4	124	3		
4		13	7	110		1
5	1				130	
6		4		1		126

Predicted Class

Googlenet

True Class	1	2	3	4	5	6
1	127				2	2
2	1	114	6	10		
3		3	127	1		
4	1	17	6	105	2	
5		1			130	
6		5			1	125

Predicted Class

MobilenetV2

True Class	1	2	3	4	5	6
1	129					2
2		121	3	7		
3		1	129	1		
4	1	20	3	104		3
5	1				128	
6						131

Predicted Class

Resnet50

True Class	1	2	3	4	5	6
1	131					
2		122	4	4		1
3			131			
4		14	3	114		
5					131	
6				1	2	128

Predicted Class

The model we proposed correctly classified 2575 of 2627 rice leaf images and misclassified 52 of them. The average accuracy value obtained in the model we proposed was 98%. This average accuracy value obtained by the proposed model is higher than the accuracy values produced by the standard CNN architectures used during the experiments.

4. Conclusion

Few nutrients that are as widely used in human nutrition as rice are found in nature. It is a known fact that there are huge problems caused by diseases of such widely consumed products. Therefore, detection of the disease in rice leaves becomes extremely important.

Deep learning methods are used extensively in image processing and classification applications today. The need to classify Rice Leaves Disease images using deep learning methods has increased in recent years [17]. In this study, a classification application was carried out using a public dataset containing 6 rice leaf classes consisting of 2627 Rise Leaves Disease images, using deep learning methods and feature selection methods together. The accuracy of the proposed model is 98%.

The CNN architectures, NCA dimension reduction method, and Cubic SVM classifier used in this study showed very effective results on rice leaf disease. This study has shown that instead of classifying by using all the features of an image, a better-performing classification can be achieved by selecting a much smaller number of qualified features. As a result, within the scope of this study, the newest and most effective models in the field of scientific artificial intelligence were used to solve this classification problem and effective results were obtained.

Competing interests

The authors declare that they have no competing interests.

References

[1] Koklu, M., I. Cinar, and Y.S. Taspinar, Classification of rice varieties with deep learning methods. Computers and electronics in agriculture, 2021. 187: p. 106285.

[2] Yucel, N. and M. Yildirim, Automated deep feature fusion based approach for the classification of multiclass rice diseases. Iran Journal of Computer Science, 2024. 7(1): p. 131-138.

[3] Bingol, H., Derin Öğrenme Modellerinde Komşuluk Bileşen Analizi Yöntemi Kullanarak Çiçek Görüntülerinin Sınıflandırılması. Fırat Üniversitesi Mühendislik Bilimleri Dergisi, 2022. 34(1): p. 439-447.

[4] Sonka, M., V. Hlavac, and R. Boyle, Image processing, analysis and machine vision. 2013: Springer.

[5] Tan, M. and Q. Le. Efficientnet: Rethinking model scaling for convolutional neural networks. in International conference on machine learning. 2019. PMLR.

[6] Zhang, X., et al. Shufflenet: An extremely efficient convolutional neural network for mobile devices. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[7] He, K., et al. Deep residual learning for image recognition. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2016.

[8] Simonyan, K. and A. Zisserman, Very deep convolutional networks for large-scale image recognition. arXiv preprint arXiv:1409.1556, 2014.

[9] Ghosal, S. and K. Sarkar. Rice leaf diseases classification using CNN with transfer learning. in 2020 IEEE Calcutta Conference (Calcon). 2020. IEEE.

[10] Ahad, M.T., et al., Comparison of CNN-based deep learning architectures for rice diseases classification. Artificial Intelligence in Agriculture, 2023. 9: p. 22-35.

[11] Bhattacharya, S., A. Mukherjee, and S. Phadikar, A deep learning approach for the classification of rice leaf diseases. Intelligence Enabled Research: DoSIER 2019, 2020: p. 61-69.

[12]14\05\2024 Available from: https://www.kaggle.com/datasets/dedeikhsandwisaputra/rice-leafs-disease-dataset?resource=download.

[13] Krizhevsky, A., I. Sutskever, and G.E. Hinton, Imagenet classification with deep convolutional neural networks. Advances in neural information processing systems, 2012. 25.

[14] Szegedy, C., et al. Going deeper with convolutions. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2015.

[15] Sandler, M., et al. Mobilenetv2: Inverted residuals and linear bottlenecks. in Proceedings of the IEEE conference on computer vision and pattern recognition. 2018.

[16] Goldberger, J., et al., Neighbourhood components analysis. Advances in neural information processing systems, 2004. 17.

[17] Yildirim, M., et al., Automatic classification of particles in the urine sediment test with the developed artificial intelligence-based hybrid model. Diagnostics, 2023. 13(7): p. 1299.