GRAIN PALETTE-A DEEP LEARNING ODYSSEY IN RICE TYPE CLASSIFICATION THROUGH TRANSFER LEARNING

ABSTRACT

Rice is a widely produced grain product globally and exists in numerous genetic varieties. These varieties possess distinguishing characteristics such as texture, shape, and color, leading to their separation from one another. The classification and evaluation of seed quality can be achieved through these distinguishing features of rice varieties. The research employed five distinct types of rice, including Arborio, Basmati, Ipsala, Jasmine, and Karacadag, which are commonly cultivated in Turkey. The dataset used in the study consisted of 75,000 grain images, with 15,000 images from each of the five rice varieties. The study employed the Convolutional Neural Network (CNN) algorithms to build models for the image dataset, and subsequently performed classification processes. Previous research has predominantly concentrated on the classification of intact or head kernels, with minimal attention paid to the inclusion of broken kernels that can affect classification accuracy. This research aims to address this gap by incorporating both broken and intact rice kernels in the dataset to replicate real-world field-testing conditions. The study calculated statistical results such as

sensitivity, specificity, prediction, F1 score, accuracy, false positive rate, and false negative rate using the confusion matrix values of the models. The results of each model were presented in tables. The findings indicate that the models used in the research for classifying rice varieties can be effectively applied in this field.

1. INTRODUCTION

Principal cereal grains
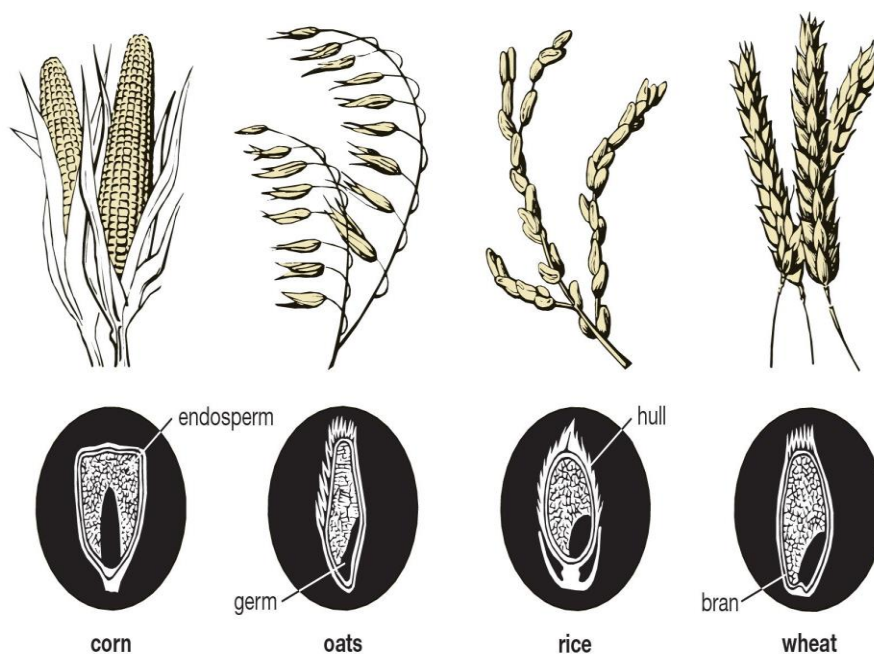
© 2010 Encyclopædia Britannica, Inc.

The use of image processing and computer vision applications in agriculture has gained interest due to their nondestructive nature and cost-effectiveness when compared to manual methods . Compared to traditional manual methods, computer vision applications based on image processing offer several advantages . Manual evaluation or classification of grains can be time-consuming and expensive, and can also be influenced by the evaluator's experience. Furthermore, manual methods may lead to inconsistent evaluation results, especially when assessments are made on a large scale, making rapid decision-making difficult. Rice is a widely produced and consumed grain product in many countries around the world, and is priced based on various parameters in the market, such as texture, shape, color, and fracture rate. To determine

these parameters and perform classification operations, digital images of rice products are acquired and various machine learning algorithms are used. These algorithms enable large amounts of data to be analyzed quickly and reliably. Using such methods in rice production is important for improving the quality of the final product and meeting food safety criteria in an automated, economical, efficient, and non-destructive manner .The evaluation and classification of rice quality have been increasingly conducted using digital image features in recent years. These features can include geometric parameters (such as length and perimeter), fracture rate, whiteness, and the detection of cracks in rice grains. Various image processing systems can be used to extract these features, which can then be classified using machine learning algorithms such as ANN, SVM, LR, DNN, and CNN. Several studies have employed these algorithms for rice quality assessment, and a summary of these studies is presented in Table 1.

The primary objective of this research is to develop a non-destructive model that can enhance the classification success rate, not just in the laboratory setting but also in the harsh field-testing environment

Table 1 Similar Studies found in the literature.

No	Crop	Data Pieces	Classifier	Class	Accuracy %
1.	Soybean	1670	BPNN	4	90.00
2.	Soybean	1200	SVM	4	99.03
3.	Soybean	4366	SVM	5	95.90
4.	Dry bean	13,611	SVM	7	93.13
5.	Rice	1700	SVM	2	98.50
6.	Rice	843	SVM	16	87.18
7.	Rice	7399	DCNN	3	95.50

8.	Rice	3810	LR	2	93.02
9.	Rice	200	CNN	3	88.07
10.	Wheat	640	ANN	2	87.50
11.	Wheat	180	SVM	2	95.00
12.	Wheat	7000	SVM	2	86.81
13.	Wheat	150	ANN	16	72.80
14.	Wheat	6400	SVM	40	88.33
15.	Wheat	200	ANN	2	99.93
16.	Wheat	3000	ANN	2	93.46

where the likelihood of broken rice kernels is high and can affect the model's accuracy. The proposed models employ CNN networks with varying numbers of layers for classification. The dataset is divided into two categories: one includes broken rice kernels, and the other comprises intact rice kernels. Additionally, 75,000 rice images from five different classes were evenly distributed to theCNN method, which is capable of classifying raw images without preprocessing. The CNN methods' classification success rates were compared.

The paper is structured as follows: Section two provides a description of the dataset, performance metrics, crossvalidation, and methods used in the study. Section three outlines the experimental results obtained in the study. The final section includes an evaluation of the experimental results and presents recommendations.

2. MATERIAL AND METHODS

The study utilized various CNN algorithms with different layers architecture to develop models that can classify rice varieties using modified datasets that include both broken and head kernels. Figure 1 illustrates the flow chart of the

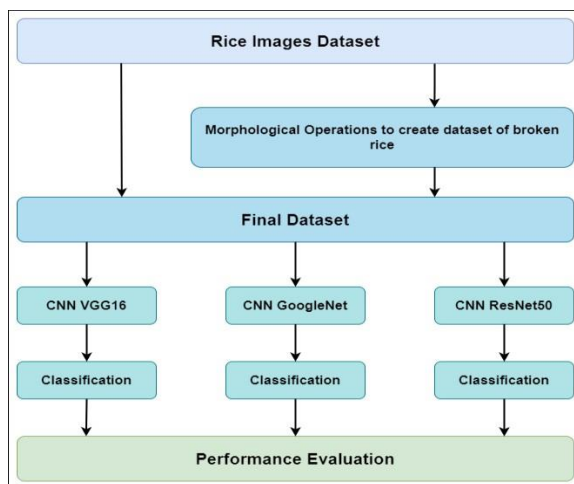proposed   models   for   rice   variety   classification.

Figure 1 : FlowChart of the Proposed Models

DATASETS : The study utilized datasets of five rice varieties commonly grown in Turkey, namely Arborio, Basmati, Ipsala, Jasmine, and Karacadag. The first dataset contained 75,000 rice grain images, with 15,000 images from each variety[. The images in this dataset were RGB and had a size of 250 x 250 pixels, with each rice grain located within the image. In addition, a filter was applied to all images using MATLAB environment to synthetically convert them into broken kernels, resulting in an additional 75,000 rice kernel images. The model used to convert intact rice kernels into broken ones was designed to produce images that closely resemble actual field scenarios. The original rice varieties used in the study, along with their corresponding broken kernels, are depicted in Fig. 2. The work carried out in this study involved pre-processing of the dataset, which included skew correction of all rice kernels and positioning them vertically. This was achieved by predicting the major axis of the kernel using its morphological features and determining the horizontal orientation. Additionally, a random ratio ranging from 10% to 40% was used to cut the same kernel and produce a broken kernel to be included in the dataset, making it more realistic.
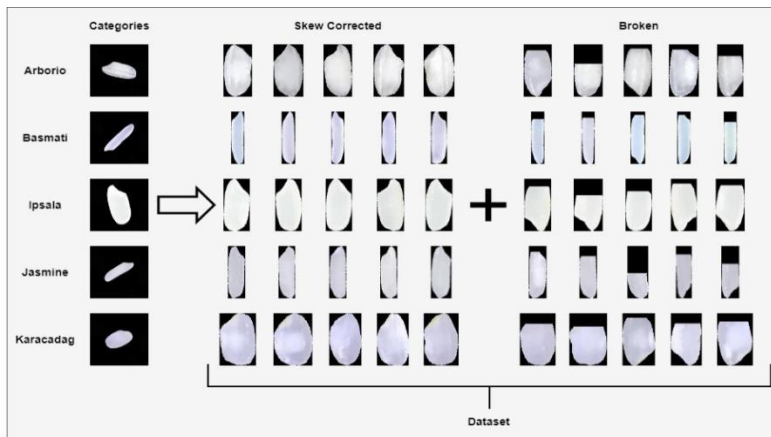
Figure 2 Original Rice Varities and corresponding broken kernels

PERFORMANCE METRICS : The confusion matrix is a tool used to evaluate the performance of machine learning methods in classification tasks. It helps to identify the connections between the classifier's performance and the results of the tests. This matrix provides information on both correct and incorrect classifications of positive and negative samples.explained the benefits of the confusion matrix. In Table 2, a two-class confusion matrix is presented.

Table 2 Confusion Matrix

		Predicted Class	
		Positive	Negative
Actual Class	Positive	TP (True Positive)	FN (False Negative)
	Negative	FP (False Positive)	TN (True Negative)

Table 3 Five Class Confusion Matrix

		Predicted Class				
		$CLASS_1$	$CLASS_2$	$CLASS_3$	$CLASS_4$	$CLASS_5$
Actual Class	$CLASS_1$	$TRUE_1$	$FALSE_{12}$	$FALSE_{13}$	$FALSE_{14}$	$FALSE_{15}$

CLASS$_2$	FALSE$_{21}$	TRUE$_2$	FALSE$_{23}$	FALSE$_{24}$	FALSE$_{25}$
CLASS$_3$	FALSE$_{31}$	FALSE$_{32}$	TRUE$_3$	FALSE$_{34}$	FALSE$_{35}$
CLASS$_4$	FALSE$_{41}$	FALSE$_{42}$	FALSE$_{43}$	TRUE$_4$	FALSE$_{45}$
CLASS$_5$	FALSE$_{51}$	FALSE$_{52}$	FALSE$_{53}$	FALSE$_{54}$	TRUE$_5$

Table 3 presents the five-class confusion matrix used in the classification processes of the rice dataset in this study. The values of true positive (TP), false positive (FP), false negative (FN), and true negative (TN) are represented in the matrix.

The statistical performance of classifiers can be analyzed in detail by performing calculations using TP, FP, FN and TN values in the confusion matrix]. Table 4 shows the metrics obtained from statistical calculations for the two-class confusion matrix, along with the formulas used to calculate these metrics and information about their purpose. On the other hand, Table 5 shows the calculation of TP, TN, FP and FN values in a five-class confusion matrix.

Other performance metrics for evaluating classification algorithms exist in the literature, in addition to those presented in Table 4. However, since the dataset used in this study has an equal number of data in each class and the classification success is already high, there is no need to use additional metrics. The range of values for the metrics used is from 0 to 1, but since the classification success of the three models used is high, rounding these values does not allow for a proper comparison. To address this, the metric values are presented as percentages.

Table 5: Statistical Metrics

Metrics	Formula	Symbol	Description
Sensitivity	tp / (tp + fn)	SNS	Provides the count of accurate positive estimates

Specificity	tn / (tn + fp)	SPC	refers to the number of negative predictions that were correctly classified
Precision	tp / (tp + fp)	PRE	Provides a value indicating a positive estimate
F1-Score	(2*tp) / (2*tp + fp +fn)	F1S	The F1 score is the harmonic mean of precision and recall values. It is used to evaluate the model's performance on datasets with uneven class distributions
Accuracy	(tp + tn) / (tp + tn + fp + fn)	ACC	This metric gives the overall accuracy of the model's classification
False Positive Rate	fp / (tn + fp)	FPR	Provides the number of positive estimates that were incorrectly classified as negative
False Negative Rate	fn / (tp + fn)	FNR	Provides the number of negative estimates that are incorrectly classified

CROSS VALIDATION : Cross-validation is a widely used method for assessing the accuracy of classification models in a rigorous and objective way. It involves dividing the dataset into k equal parts, with 1/k of the data reserved for testing and k-1/k used for training. This process is repeated k times, with each part used once as the test set. The overall classification accuracy of the model on the test set is then obtained by taking the arithmetic mean of the classification accuracies obtained from each round of testing . The study employed a value of k equal to 10 for the cross validation method.

CONVOLUTIONAL NEURAL NETWORK (CNN) : Convolutional Neural Network (CNN) is a deep learning method that is commonly used in various fields, such as image processing, natural language processing, voice recognition, and datasets that contain a large number of data. CNN can function as an end-to-end classifier, where it can extract features with its layers, and learn and classify with these features. It is composed of five main layers, namely, the convolution layer, pooling layer, activation layer, fully connected layer, and softmax layer. The convolution layer applies various filters to extract image features from each region of the image. Meanwhile, the pooling layer reduces the size of the image feature to transfer it to the next layer while also optimizing the adjustments to avoid negatively affecting the classification.

The pooling layer is responsible for reducing the dimensionality of the feature maps obtained from the convolution layer. This helps to reduce the computational complexity of the network while preserving the important features. The pooling operation can be performed using different methods such as max pooling, average pooling, and L2-norm pooling. The optimal adjustment of the pooling layer is important to prevent information loss and avoid overfitting.The activation layer introduces non-linearity to the network by applying a specific activation function to each neuron's output. The fully connected layer is responsible for processing the extracted features from previous layers and performing the final classification. In this layer, each neuron is connected to all neurons in the previous layer, and learning operations are performed to make inferences. Finally, the softmax activation function is used in the output layer to parse classes and produce the final classification results .

Transfer learning is a technique used in deep learning where a pre-trained model is used as a starting point for a new model. In this technique, the knowledge gained by training a model on a large dataset is transferred to a new model to help it learn and generalize better on a smaller dataset.

In this study, the VGG16 network structure and trained VGG16 model were used as a starting point for training the CNN network. VGG16 is a convolutional neural network architecture that was developed for image classification tasks. The pre-

trained VGG16 model was trained on a large dataset of images, and its learned features were transferred to the new model to help it classify the images in the present dataset. This approach reduces the time and resources required for training a new model from scratch and improves the performance of the model.

The modified VGG16 architecture was utilized to extract features from rice images, where the fc8 layer, the last fully connected layer of VGG16, was removed, and the fc7 layer's features were utilized. Additionally, the fully dependent fc_optimized layer was added to the network, which produced five outputs for classification.

Another CNN network used in this study is GoogLeNet, also known as Inception v1, is a deep convolutional neural network (CNN) architecture that was introduced in 2014 by a team of researchers at Google. It was designed with the goal of increasing both the accuracy and efficiency of image classification tasks, and achieved this by utilizing a unique "inception module" that allowed the network to capture features at multiple scales without dramatically increasing the number of parameters. GoogLeNet achieved state-of-the-art results on the ImageNet Large Scale Visual Recognition Challenge (ILSVRC) 2014, solidifying its place as one of the most influential CNN architectures in the field of deep learning.

Last but not least the CNN network ResNet50, is a deep residual neural network that was introduced in 2015. It is a 50-layer convolutional neural network that has significantly improved the performance of image classification tasks compared to its predecessors. The key innovation of ResNet50 is the introduction of skip connections, or shortcuts, that allow information to flow more easily through the network. These connections enable ResNet50 to train deeper networks while minimizing the vanishing gradient problem. ResNet50 has been shown to achieve state-of-the-art performance on a range of image classification benchmarks, including the ImageNet Large Scale Visual Recognition Challenge.

3. EXPERIMENTAL RESULTS

In this section, the classification results obtained by the CNN method are presented. The dataset consists of features extracted from 75,000 rice grain images, which also include 75,000 broken rice grain images, and it includes Arborio,Basmati, Ipsala, Jasmine, and Karacadag rice classes for classification. The images in the dataset were used as an input to the CNN. Table 7 presents the hardware specifications and network structures used to run these algorithms.To evaluate the performance of the classification algorithms used in this study, a confusion matrix was employed. For each algorithm, a confusion matrix was created, and values such as TP, TN, FP, and FN were used for performance evaluations. Metrics such as SNS, SPC, PRE, F1S, ACC, FPR, and FNR were used to assess performance. To objectively evaluate the success of the models, the cross-validation method was utilized, with a k value of 10 being determined. The study utilized the CNN method, where 150,000 rice images (consisting of head and broken kernels) for each rice variety were fed as input to the CNN model. To enhance the success rate of the CNN method, the weights of the VGG16 network were utilized to train the model using transfer learning. To mitigate the overfitting problem of the trained CNN model, dropout was incorporated. As a result, the classification accuracy has increased to nearly 100%, and overfitting issues have been avoided. The confusion matrix obtained from this classification process can be found in Table 8.

Table 7 Specifications of hardware used in the study and network parameters of classifiers

HARDWARE UNIT	SPECIFICATIONS
CPU Unit	AMD Ryzen 7 4800H 2.90GHz
Graphic Card	Nvidia RTX 2070

Table 8 CNN Confusion Matrix

VGG		Predicted Class				
		Arborio	Basmati	Ipsala	Jasmine	Karacadag
Actual Class	Arborio	29997	2 29995	0	2	1
	Basmati	1		0	2	1
	Ipsala	0	2	29999	1	0
	Jasmine	1	1	0	29994	0
	Karacadag	1	0	1	1	29998

GoogleNet		Predicted Class				
		Arborio	Basmati	Ipsala	Jasmine	Karacadag
Actual Class	Arborio	29998	2	0	0	1
	Basmati	1	29997	0	0	1
	Ipsala	0	0	29999	1	0
	Jasmine	0	1	0	29998	0
	Karacadag	1	0	1	1	29998

ResNet		Predicted Class				
		Arborio	Basmati	Ipsala	Jasmine	Karacadag
Actual Class	Arborio	29999	1 29998	0	0	0
	Basmati	0		0	0	0
	Ipsala	0	0	29999	1	0
	Jasmine	0	1	0	29998	0
	Karacadag	1	0	1	1	30000

Table 9 Statistical Results based on results from CNN method (%)

VGG	Arborio	Basmati	Ipsala	Jasmine	Karacadag
SNS	99.99	99.98	99.99	99.98	100
SPC	100	100	100	100	99.99
PRE	100	99.99	100	99.99	100
FIS	100	99.99	100	99.99	100
ACC	99.99	99.99	99.99	99.98	99.99
FPR	0	0	0	0	0
FNR	0.01	0.02	0	0.02	0

GoogleNet	Arborio	Basmati	Ipsala	Jasmine	Karacadag
SNS	99.99	99.99	100	99.99	99.99
SPC	100	100	99.99	100	100
PRE	100	99.99	100	99.99	100
FIS	100	99.99	100	99.99	100

ACC	99.99	99.99	99.99	99.99	99.99
FPR	0	0	0	0	0
FNR	0.01	0.01	0	0.01	0

4. CONCLUSIONS

The study evaluated the performance of CNN models using rice images and statistical measurements of confusion matrix were used as performance metrics. The study used SNS, SPC, PRE, F1S, ACC, FPR, FNR values to compare the performance of each method and class. These metrics were used to calculate the training and testing success of the algorithms. The study demonstrates that CNN networks have a high capacity to accurately predict rice kernel varieties, even in cases where they are broken by up to 40% of their original length. The average classification success of all CNN networks is nearly 100%, with ResNet50 exhibiting a higher accuracy rate compared to the other models. Although there is a slight variation in ResNet50's accuracy rate, it is significantly higher than the other models by a few decimal points. Specifically, ResNet50 achieved 100% accuracy in predicting Arborio, Ipsala, and Karacadag varieties, with only one or two samples failing for other varieties.