ISSN 2976-1190 (Online)
Journal of Engineering and Sciences
Vol. 1 Issue 1

Varietal Identification of Rice Seed Using Deep Convolutional Neural Network

Amrit Poudel[1], Bidur Devkota[2]

[1] *Department of Software Engineering, Gandaki College of Engineering and Science, Pokhara University, Nepal*
[2]*Department of Electronics and Computer Engineering, IOE, Pashchimanchal Campus, Nepal*

(Manuscript Received 16/08/2022; Review: 30/09/2022; Revised 03/10/2022; Accepted 05/10/2022)

Abstract

Speed and accuracy in agricultural production are essential for long-term economic growth, competitiveness, and sustainability. Traditional manual rice seed classification operations are expensive and unreliable because human decisions regarding the identification of objects and problems are inconsistent, subjective, and slow. Process automation using machine vision technology is quick, accurate, inexpensive, and non-destructive. In this study, a hybrid CNN-SVM model was employed to classify rice seeds. SVM was used for classification, and the VGG16 pretrained model was used to extract features. For the classification of similar and dissimilar rice seed species, the National Center of Protection of New Varieties and Goods of Plants (NCPNVGGP) of Vietnam's rice dataset was implemented. The VNRICE seed dataset was the other dataset that was used. However, a new dataset with three varieties (Chhomrong, Machhapuchhre-3, and Lumle-2) was generated and used for the classification of rice seeds from Nepal. A F1 score of 0.78 was obtained for the classification of rice seeds from similar species. The F1 score for the classification of dissimilar species using rice seeds was 0.93. The F1 score for rice seeds in the VNRICE dataset was 0.92. However, CNN-SVM performed classification with an F1 score of 0.91 for the dataset of Nepali rice seeds. These findings show that the CNN-SVM classifier is a useful tool for the classification of various rice seed varietals.

Keywords: CNN-SVM; Hybrid Model; Rice Dataset; Rice Seed Classification

---

1. Introduction

Classification of seed varieties is one of the foremost importance for seed producers and farmers to ensure the purity of a variety and crop output [1]. Rice is the monocot plant Oryza Sativa's (Asian rice) seed [2][3]. In addition to being incredibly nutrient-dense and delectable, rice is the main food consumed by 3 billion people worldwide [4]. Rice seeds of different varieties are very similar in terms of morphology and color[5].

Human judgments in determining the variety of seed are contradictory, arbitrary, and slow [6]. Additionally, traditional manual seed classification operations are inefficient and costly [7]. However, machine vision technology provides a nondestructive, cost-effective, quick, and reliable solution [8].

Corresponding author. Amrit Poudel, Bidur Devkota

E-mail address: pdlamrit7@gmail.com , im.bidur@gmail.com

Image preprocessing, segmentation, extraction, and classification are the four blocks that make up computer vision-based classification [9], with feature extraction playing a crucial role in classification accuracy [10]. Therefore, several feature extraction and machine machine learning techniques have been proposed for seed variety classification. The features such as shape [11], texture [12], wavelet [13] and color [14] are used for seed variety classification. Several classifiers based on k-means clustering [15], backpropagation neural network (BPNN) [13] and Support Vector Machine (SVM) [12] have been developed. Previous research has shown that the computer vision technique works well for the classification of seeds with different features and classifiers.

1.1 Problem Statement

It's still a huge challenge to classify rice seed variants with similar visual appearances. Hybridization has lately resulted in the production of a number of rice seed varieties. It created an issue with their classification, which has grown more difficult owing to their strong resemblance. To solve this problem, previous research focused on extracting the handcrafted features of rice seeds [11][12][14][16]. Unfortunately, this classification technique necessitates a specific seed orientation in a sorting system, which makes industrial scaling difficult.

1.2 Motivation and Contributions

The hypothesis that higher-level features with more discriminative information could dramatically improve rice seed classification motivated our research. In recent years, traditional methods for extracting low-level features from images have been replaced by CNNs, which extract robust high-level features automatically and hierarchically [17]. In comparison to traditional feature-engineered methods, numerous studies have shown that using CNNs as a generic extractor can significantly improve the accuracy of computer vision tasks [18]. For example, in the research [19], a CNN architecture VGG-16 was used to extract features that were then fed into an SVM classifier to identify 40 different wheat grain varieties. On the proposed model, this study achieved 100% accuracy. In another research [20], the feasibility of haploid corn seed classification using hyperspectral images was studied by extracting features using a CNN architecture VGG-19 and a classification accuracy of 96.32% was achieved. The research [21] used a deep convolutional neural network (CNN) as a generic feature extractor for corn seeds. Artificial neural networks (ANN), cubic support vector machines (SVM), quadratic SVM, weighted k-nearest-neighbor (KNN), boosted tree, bagged tree, and linear discriminant analysis were used to classify the extracted features (LDA). Corn seed varieties were classified more accurately by models trained with CNN-extracted features than by models trained with only simple features. The CNN-ANN classifier performed the best, correctly classifying 2250 test cases with accuracy, precision, recall and F1 score of 98 percent.

In research [22][23], for image classification tasks, a hybrid CNN-SVM model is able to improve the performance over an original CNN, in both binary classification and multi-class classification problems.

However, for rice seed variety classification, in the recent research, the hybrid CNN SVM architecture was not used where CNN was used for feature extraction and SVM was used for classification using those CNN extracted features [24][25].

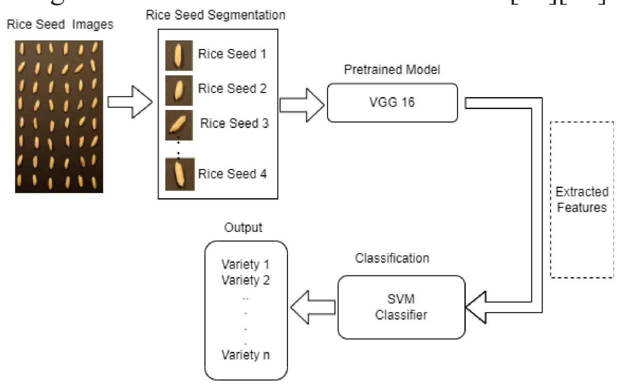

Figure 1: System Model

Hence in this research, a hybrid CNN-SVM architecture is used for rice seed variety classification where CNN is used for feature extraction and SVM is used for classification.

The major contributions of this research are:

- Use of CNN-extracted features for rice seed classification
- Creation of new dataset for rice seed classification

The following is how this section is organized: The research methodology is discussed in Section 2, which includes a detailed description of feature extraction techniques and classifiers. The results of rice seed varieties are presented in section 3. Section 4 concludes with a summary of current research and recommendations for future research.

2. Materials and Method

The system model that was used for the classification of rice seeds in this research is shown in Figure 1. In this research, the transfer learning method was employed to extract features using the VGG-16 model pretrained on the ImageNet dataset. The pre-trained weights were used as initial weights in deep learning architecture to extract features from the input image. Using the VGG-16 pretrained model, deep extracted features are extracted and these features are fed into the SVM classifier.

2.1 Sample Preparation

The RGB images of known rice seed varieties provided by the National Center of Protection of New Varieties and Goods of Plants (NCPNVGGP) in Vietnam were used exclusively in the research. Based on the research[6][3], in this research, the rice seed varieties were divided into similar and dissimilar species to test how the model performs on both similar and dissimilar rice species.NDC1, NepCoTien, NepDacSanLienHoa, NV1, and NepThomBacHai, NepThomHungYen were similar rice varieties. Dissimilar rice varieties included BC15, BQ10, NH92, NT16, PC10, and VietHuong8. There were 96 RGB images in each rice seed variety. VNRICE was the other dataset used in this study. Six paddy varieties were included in the VNRICE[7][8][26][27]. BC15, Huongthom-1, Nep-87, Q-5, Thien-uu-8, and X–23 were the rice varieties included in this dataset. In total, the VNRICE dataset contained 10508 RGB images of rice seeds.

Three varieties of rice seeds were used to classify Nepali rice seeds: Chhomrong, Machhapuchhre3, and Lumle-2. These rice seeds were obtained from Regional Agricultural Research Station (RARS), Lumle, Nepal. There were 3400 RGB images in each rice seed variety

2.2 The convolutional neural network feature extraction

The structure of CNN model consists of three main neural layers: convolution, fully connected, and pooling, each of them has different tasks in the network architecture. The kernel of the CNN structure is the convolutional layer that performs the heaviest computational operations. The convolutional layers closer to the input extract basic features such as edges of different orientations, while deeper convolution layers extract complex and abstract features such as specific subsampled object regions [18]. Convolution layers are frequently followed by activation layers in order to capture more complex features of the input image and increase the nonlinearity of the deep learning structure. A pooling layer is placed between successive convolutional layers to reduce the number of parameters and computational complexity of the model [28]. As a result, this layer aids generalization by preventing overfitting [29]. VGG-16 is made up of 13 convolutional layers, 5 pooling layers, and 3 fully connected layers [30].

2.3 Training and testing procedure

To develop classification models, the datasets were randomly categorized into different training and testing subsets in the ratio of 5:5, 6:4, 7:3, 8:2 and 9:1.

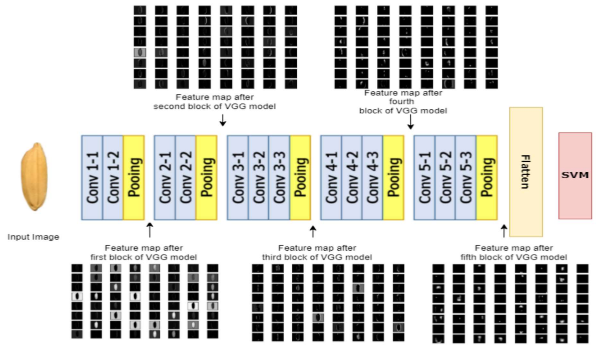

Figure 2 The overall architecture of the feature extraction

2.4 Support Vector Machine

The SVM is a maximum margin classifier. SVM attempts to separate training vectors by directly searching appropriate boundaries between classes, rather than modeling their probability distribution. A decision line with the maximum distance from the nearest training points of each class achieves a good separation or lower SVM generalization error [19]. The hybrid CNN-SVM architecture established for the research is shown in Figure 1.

3. Results and Discussion

After extracting features from the images of the samples, the extracted features were fed into the SVM classifier and the performance of the classifier was evaluated in terms of precision, recall, F1-score and accuracy.

Table 1, Table 2, Table 3 and Table 4 showed the classification result of rice seeds on similar species, dissimilar species, VNRICE dataset and on Nepali rice dataset respectively on different train to test data ratio. From Table 1 and 2, it was found that the hybrid CNN-SVM model performs better classification on dissimilar species rather than similar species. However, for the VNRICE and Nepali rice seed dataset the CNN-SVM model performed best when the train to test data split ratio was 9:1.

Table 1. Performance of classification of the test data on dissimilar species

Split	Precision	Recall	F1-score	Accuracy
5:5	0.89	0.88	0.88	0.88
6:4	0.90	0.89	0.89	0.89
7:3	0.89	0.89	0.89	0.89
8:2	0.93	0.93	0.93	0.93
9:1	0.89	0.89	0.88	0.88

Table 2. Performance of classification of the test data on similar species

Split	Precision	Recall	F1-score	Accuracy
5:5	0.78	0.75	0.75	0.78
6:4	0.81	0.77	0.78	0.78
7:3	0.70	0.76	0.76	0.78
8:2	0.79	0.78	0.78	0.76
9:1	0.77	0.75	0.75	0.78

Table 3. Performance of classification of the test data on VNRICE dataset

Split	Precision	Recall	F1-score	Accuracy
5:5	0.90	0.90	0.90	0.90
6:4	0.89	0.89	0.89	0.89
7:3	0.89	0.88	0.88	0.88
8:2	0.90	0.89	0.90	0.89
9:1	0.92	0.92	0.92	0.91

Table 4. Performance of classification of the test data on Nepali dataset

Split	Precision	Recall	F1-score	Accuracy
5:5	0.90	0.90	0.90	0.90
6:4	0.89	0.89	0.89	0.89
7:3	0.90	0.90	0.90	0.90
8:2	0.90	0.90	0.90	0.90
9:1	0.91	0.91	0.91	0.91

4. Conclusions

In this study, we created a hybrid CNN-SVM classification model to categorize rice seeds on a dataset of four different types of rice seeds (similar, dissimilar, VNRICE and Nepali).

For Nepali rice seed classification, three varieties of rice seeds were used: Chhomrong, Machhapuchhre3, and Lumle-2. The performance of CNN-SVM model for the classification of rice seeds was evaluated in terms of precision, recall, F1-score and accuracy. It is found that CNN-SVM classifies rice seeds of dissimilar species with higher accuracy than that of similar rice seed species. For the VNRICE and Nepali dataset, the CNN-SVM performed classification with the accuracy of 0.91. The result showed that CNN-SVM classifier is an efficient tool for the classification of different rice seed varieties.

As the next step, following future works can be carried out:

- explore the effect of adding more varieties of rice seeds on the classification performance (precision, accuracy, recall and F1 Score).
- explore the effect of dimensionality reduction on classification time and accuracy.

References

[1] Shouche S. P., Rastogi R., Bhagwat S. G., & Sainis J. K. Shape analysis of grains of Indian wheat varieties, *Computers and Electronics in Agriculture*, 33(1) (2001) 55–76.

[2] Maheshwari C. QUALITY ASSESSMENT OF ORYZA SATIVA SSP INDICA (RICE) USING COMPUTER VISION, *International Journal of Innovative Research in Computer and Communication Engineering,* (2013).

[3] Vu H., Duong V. N., & Nguyen T. T. Inspecting rice seed species purity on a large dataset using geometrical and morphological features, *Proceedings of the Ninth International Symposium on Information and Communication Technology - SoICT* , (2018).

[4] Cui W., Liu G., Bezerra M., Lagos D. A., Li Y., & Cai Y. (2017), Occurrence of Methylmercury in Rice-Based Infant Cereals and Estimation of Daily Dietary Intake of Methylmercury for Infants, *Journal of Agricultural and Food Chemistry,* 65(44) (2018) 9569–9578).

[5] Weng S., Tang P., Yuan H., Guo B., Yu S., Huang L., & Xu C. Hyperspectral imaging for accurate determination of rice variety using a deep learning network with multi-feature fusion, *Spectrochimica Acta Part A: Molecular and Biomolecular Spectroscopy*, (2020) 234.

[6] Pourreza A., Pourreza H., Abbaspour-Fard M.-H., & Sadrnia H. Identification of nine Iranian wheat seed varieties by textural analysis with image processing, *Computers and Electronics in Agriculture*, (2012) 102–108.

[7] Bakhshipour A., Sanaeifar A., Payman S. H., & de la Guardia M. Evaluation of Data Mining Strategies for Classification of Black Tea Based on Image-Based Features, *Food Analytical Methods*, (2017)1041–1050.

[8] Ropelewska E. The use of seed texture features for discriminating different cultivars of stored apples, *Journal of Stored Products Research*, (2020).

[9] Sharif M., Khan M. A., Iqbal Z., Azam M. F., Lali M. I. U., and Javed M. Y. Detection and classification of citrus diseases in agriculture based on optimized weighted segmentation and feature selection, *Computers and Electronics in Agriculture*, 150, (2018) 220–234.

[10] Iqbal Z., Khan M. A., Sharif M., Shah J. H., Rehman M. H., & Javed K. An automated detection and classification of citrus plant diseases using image

processing techniques: A review, *Computers and Electronics in Agriculture*, 153, (2018) 12–32.

[11] Huang K.-Y., & Chien M.-C. A Novel Method of Identifying Paddy Seed Varieties. *Sensors*, 17(4), (2017) 809.

[12] Sun, J., Lu, X., Mao, H., Jin, X., & Wu, X. A Method for Rapid Identification of Rice Origin by Hyperspectral Imaging Technology. *Journal of Food Process Engineering*, *40*(1), (2015) e12297.

[13] Singh Ksh. R., & Chaudhury, S. Efficient technique for rice grain classification using back-propagation neural network and wavelet decomposition, *IET Computer Vision*, 10(8), (2016) 780–787.

[14] Liu Z., Cheng F., Ying Y., & Rao X. Identification of rice seed varieties using neural network, *Journal of Zhejiang University SCIENCE*, 6B(11), (2005), 1095–1100.

[15] Ma D., Cheng H., & Zhang W. Maize Embryo Image Acquisition and Variety Identification Based on OTSU and K-Means Clustering Algorithm, *2013 International Conference on Information Science and Cloud Computing Companion*, (2013)

[16] Thu Hong P. T., Thanh Hai T. T., Lan L. T., Hoang V. T., Hai V., & Nguyen T. T. Comparative Study on Vision Based Rice Seed Varieties Identification, *2015 Seventh International Conference on Knowledge and Systems Engineering (KSE)*, (2015)

[17] Zhang S., Zhang S., Zhang C., Wang X., & Shi Y. Cucumber leaf disease identification with global pooling dilated convolutional neural network, *Computers and Electronics in Agriculture*, 162 (2019) 422–430.

[18] Kozłowski M., Górecki P., & Szczypiński P. M. Varietal classification of barley by convolutional neural networks, *Biosystems Engineering*, 184 (2019) 155–165.

[19] Özkan K., Işık Ş., & Yavuz B. T. Identification of wheat kernels by fusion of RGB, SWIR, and VNIR samples. *Journal of the Science of Food and Agriculture*, 99(11), (2019), 4977–4984.

[20] Liao W., Wang X., An D., & Wei Y. Hyperspectral Imaging Technology and Transfer Learning Utilized in Haploid Maize Seeds Identification, *2019 International Conference on High Performance Big Data and Intelligent Systems (HPBD&IS)*, (2019)

[21] Javanmardi S., Miraei Ashtiani S.-H., Verbeek F. J., & Martynenko A. *Computer-vision classification of corn seed varieties using deep convolutional neural network. In Journal of Stored Products Researc*h, 92 (2021) 101800, Elsevier BV

[22] Jiang S., Hartley R., & Fernando B. Kernel Support Vector Machines and Convolutional Neural Networks, *Digital Image Computing: Techniques and Applications (DICTA)*, *IEEE*, (2018)

[23] Leng, Jiabing and Li, Tao and Bai, Gang and Dong, Qiankun and Dong, Han, Cube-CNN-SVM: A Novel Hyperspectral Image Classification Method,*2016 IEEE 28th International Conference on Tools with Artificial Intelligence (ICTAI)*,(2016)

[24] Kantip Kiratiratanapruk and Pitchayagan Temniranrat and Wasin Sinthupinyo and Panintorn Prempree and Kosom Chaitavon and Supanit Porntheeraphat and Anchalee Prasertsak, Development of Paddy Rice Seed Classification Process using Machine Learning Techniques for Automatic Grading Machine, *Journal of Sensors,* (2020)

[25] Samson Damilola Fabiyi and Hai Vu and Christos Tachtatzis and Paul Murray and David Harle and Trung Kien Dao and Ivan Andonovic and Jinchang Ren and Stephen Marshall, Varietal Classification of Rice Seeds Using {RGB} and Hyperspectral Images, *Institute of Electrical and Electronics Engineers ({IEEE})*,(2020)

[26] Duong H.-T., & Hoang V. T. Dimensionality Reduction Based on Feature Selection for Rice Varieties Recognition, *2019 4th International Conference on Information Technology (InCIT)*, (2019)

[27] TRUONG HOANG, V., Van Hoai, D. P., Surinwarangkoon, T., Duong, H.-T., & Meethongjan, K , A comparative study of rice variety classification based on deep learning and hand-crafted features, *ECTI Transactions on Computer and Information Technology (ECTI-CIT)*, *14*(1), (2020),1–10.

[28] Zhang K., Xu Z., Dong S., Cen C., & Wu Q. Identification of peach leaf disease infected by Xanthomonas campestris with deep learning, *Engineering in Agriculture, Environment and Food*, 12(4), (2019), 388–396.

[29] Zhu Z., Gao B., and He F. Near-Infrared Hyperspectral Imaging Combined with Deep Learning to Identify Cotton Seed Varieties, *Molecules*, 24(18), (2019) 3268.

[30] Simonyan K., & Zisserman A. Very Deep Convolutional Networks for Large-Scale Image Recognition, *ArXiv*, (2014)