AN EFFICIENT QUALITY ANALYSIS OF RICE GRAINS USING SUPPORT VECTOR MACHINE OVER CONVOLUTIONAL NEURAL NETWORK WITH IMPROVED ACCURACY

Shaik Imam Rahaman Basha[1], N. Bharatha Devi[2*]

Article History: Received: 12.12.2022 **Revised:** 29.01.2023 **Accepted:** 15.03.2023

Abstract

Aim: The aim of the study is to use an efficient quality analysis of rice grains using a Novel Support Vector Machine over a Convolutional Neural Network with improved accuracy.
Materials and Methods: Novel Support Vector Machine algorithm (N=10) and artificial neural network (N=10) was iterated 20 times and analyzed rice grains. The sample size was calculated using G-power of 80% for two groups.
Results: Novel Support Vector Machine has significantly better accuracy of 67.41% compared to the Convolutional Neural Network of 90.38%. The statistical significance of the analysis of rice grains difference is $p=0.001$ ($p<0.05$) and Independent sample T-test value states that the results in the study are significant. **Conclusion:** The accuracy performance parameter of the Convolutional Neural Network appears to be better than the Novel Support Vector Machine algorithm .

Keywords: Convolutional Neural Network, Land, Machine Learning, Population, Production, Rice grains, Novel Support Vector Machine.

[1]Research Scholar, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105.
[2*]Project Guide, Department of Computer Science and Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Saveetha University, Chennai, Tamilnadu, India, Pincode: 602105.

1. Introduction

India is the country's second rice producer, with 44 million hectares under cultivation. 65 percent of the Indian population eats rice, and the country produces 40% of the world's rice. Polished white rice (ponni) and Kavuni (Black) rice are the most common rice varieties grown in south India (Kamath 1956). Moisture content, land, grain purity, cracks, grain dimensions, whiteness, milling degree, and chalkiness are all factors that affect rice quality. These are the primary characteristics that affect rice grade and price. As a result, determining the quality of rice is critical . Manually analyzing grain kind, grading, and evaluating quality characteristics.The manual inspection process is more difficult. It is dependent on working circumstances on land, human variables, cleaning pace, and salvage recovery. Other factors that affect rice quality include the presence of impurities such as stones, damaged seeds, and fractured granules (Zhou et al. 2022). Because manual examination by human inspectors is less accurate, there is a greater chance of contaminants being mixed together. Rice quality suffers as a result of this (de Oliveira, Pegoraro, and Viana 2020). To solve these challenges, automated classifier systems based on image processing are really being developed to test the quality of rice grains. It can be classified into three categories: nutrition, hygiene, and production amount. The level of recovery of grain products in order of value is controlled by milling quality. The ability to attract buyers is determined by appearance quality; sensory and nutritional quality are established by edible properties; and health quality is determined by immunity from internal and external contamination. Despite their ubiquitous use in a variety of applications, rice grains are still exported to other nations such as South America, Australia, and the United States, and this may be advantageous to the Indian economy (D'Adamo et al. 2022).

In this research area, several research papers are available in IEEE and Science direct. In Science direct, a total of 124 papers have been published and in IEEE, 42 papers had been published then in google scholar, a total of 426 papers were published. Rice is a source of food for more than half of the global population. The most cited articles as follows (de Oliveira, Pegoraro, and Viana 2020) have citations 8 times. but its sustainable production is increasingly being challenged by a growing population worldwide, fluctuating climate conditions, and decreases in the arable land area and the availability of water for irrigation (Y. H. Li and Tuong 2001). Rice grain yield is determined by the number of panicles per plant, number of grains per panicle, and grain weight. As the major determinant of grain weight, grain size is one of the most important yield-related traits in cereal plants, influencing both yield potential and grain quality, with implications for the commercial value of the grain.Our team has extensive knowledge and research experience that has translated into high quality publications(K. Mohan et al. 2022; Vivek et al. 2022; Sathish et al. 2022; Kotteeswaran et al. 2022; Yaashikaa, Keerthana Devi, and Senthil Kumar 2022; Yaashikaa, Senthil Kumar, and Karishma 2022; Saravanan et al. 2022; Jayabal et al. 2022; Krishnan et al. 2022; Jayakodi et al. 2022; H. Mohan et al. 2022)

In the existing system, there is a research gap. Because evaluating the quality of rice grains can be complicated. Machine learning techniques for rice quality have been the subject of some research. It is, therefore, critical to examine and compare the various classifications algorithms that provide greater accuracy. As a result, the purpose of this research is to compare the accuracy of Novel Support Vector Machines and Convolutional Neural Network methods for rice grain quality analysis.

2. Materials and Methods

The research study was done in an Internet Programming Lab, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences. The number of groups identified for the study are two. Group-1 is a Novel Support Vector Machine algorithm and Group-2 is the k-nearest neighbor algorithm. Sample size for each group was calculated by using previous study results in clinical.com by keeping g power as 85%, threshold 0.05 and confidence interval as 95% (Edmonston 2011). According to that, the sample size of the Novel Support Vector Machine (N=10) and k-nearest neighbor algorithm (N=10) were calculated.

Support Vector Machine

Support Vector Machine is a supervised Machine Learning model that is able to solve the regression and classification problems. It is, however, mostly used to solve classification problems. Every data point placed on N-dimensional space represents space. Every point that is pointed in the dimensional space represents the feature that is the value of a specific coordinate in the Support Vector Machine. The objective of the Support Vector Machine algorithm is to find a hyperplane in an N-dimensional space that distinctly classifies the data points. The dimension of the hyperplane depends upon the number of features. If the number of input features is three, then the hyperplane becomes a 2-D plane. It becomes difficult to imagine when the number of features exceeds three.

Pseudo Code

Input - rice grain assignment dataset

Output - Accuracy

1. Import the packages and import the dataset into programs using anaconda.
2. Load the dataset bold text and read the dataset with 205 rows and 26 columns.
3. Removing the irrelevant features and checking for null values.
4. Converting categorical values to numeric values.
5. Split the data into training and testing sets and model building and training.
6. Check the Accuracy.

Convolutional Neural Network

Convolutional Neural Network is a class of deep neural networks which is most commonly applied to analyzing visual imagery. Their other applications include video understanding, speech recognition and understanding natural language processing. The method is trained using training data, such as a series of photos in the input and their labels at the output. It entails feeding images of the training set, x, and their associated labels (targets), y, to a convolutional neural network in order to learn the network's function, y=f (x). The sequence of words we use Recurrent Neural Networks more precisely an LSTM, similarly for image classification we use Convolutional Neural Network. In this blog, we are going to build a basic building block for a Convolutional Neural Network.

Pseudo Code

Input - rice grain assignment dataset

Output - Accuracy

1. Import the packages and import the dataset into programs using anaconda.
2. Load the dataset bold text and read the dataset with 205 rows and 26 columns.
3. Split the data into sets (x_test,y_test) and (y_train,x_train).
4. Classify the fitting into two sets: x,y train and x,y test.
5. Predict and Check the Accuracy.

The research work was an experiment in google colab, the Hardware and Software requirements for experimenting the work includes i5 processor, 500GB SSD, 16GB RAM, Windows OS, python: colab/jupiter. The dataset was divided into two parts: Training and testing sets. Then the algorithm is experimented on and the training sets are varied 10 times based on the set size. The dataset contains the information about the rice grain and full data about the quality of rice grains. Based on the different rice image data, it is said that the quality of rice grain is used to predict the rice grain by applying the algorithms. The Dataset was collected from kaggle (https://www.kaggle.com/ntnu-testimon/paysim1).

Statistical Analysis

In this research study the quality analysis of rice grains was done using Anaconda Navigator – Spyder and IBM SPSS. It is a statistical software tool (Shukla, Strooper, and Carrington 2005) used for data analysis. For both proposed and existing algorithms 10 iterations were done with a maximum of 80-90 samples and for each iteration, Dependent variables are batch size. The predicted accuracy was noted for analyzing accuracy and for every change in the input. Independent T-Test is used to compare Novel Support Vector Machine algorithms and Convolutional Neural Network algorithms to analyze the quality of rice grains.

3. Results

Table 1 shows the quality analysis of rice grains with respect to its Mean and Standard Deviation of Novel Support Vector Machine algorithm and Convolutional Neural Network algorithm, which describes that Novel Support Vector Machine has an accuracy mean of 67.41%, Standard Deviation of 20 for the sample size of N=10, where the Convolutional Neural Network algorithm has an accuracy Mean of 90.38%, Standard Deviation of 20 for the sample size of N=10. Based on the above results the statistical significance of the Convolutional Neural Network algorithm is high. Table 2 shows the quality analysis of rice grains with respect to its Mean Difference, Standard Error Difference and Significant Difference of Novel Support Vector Machine algorithm and Convolutional Neural Network algorithm, which describes that there is a significant difference between the two groups with $p=0.01$ ($p<0.05$). Table 3 shows the accuracy values of SVM algorithm and CNN algorithm. The mean accuracy of SVM and CNN are 67.41 and 90.38 respectively.

Figure 1 shows the bar graph comparison of the mean accuracy of Convolutional neural network algorithm quality analysis of rice grains. The mean and standard deviation of CNN is better than SVM and this shows that the performance of CNN performs well in the analysis of rice grains effectively.

4. Discussion

The accuracy of Convolutional Neural Network algorithm quality analysis of rice grains appears to be better than the Novel Support Vector Machine algorithm (Parhami 2006; Y. Li et al. 2018) quality analysis of rice grains accuracy. The statistical significant difference between the two groups obtained is p=0.001 ($p<0.05$). Consumers of today are very conscious about the quality of food grains (Poutanen et al. 2021). In order to ensure the quality of rice grains, an automated rice grain quality assessment system based on CNN and SVM classifiers has been addressed in this research work. Two types of rice grains viz. Ponni and Matta are taken up for study. The proposed system classified the rice grains based on their morphological (Szymanek 2012) and geometrical features (Polster 2012). The experimental analysis showed that the proposed CNN classifier (Sheela et al. 2016) has an overall accuracy of 90.38%, whereas SVM resulted in 67.41%. On the other hand, when the other parameters are considered, such as computational time and parallel processing (Parhami 2006) (classifying more than one sample at a time). The limitation performance of SVM is analyzed in rice quality from the dataset obtained from kaggle. The work has independent values which are present in the data set which makes the process slow and the performance of SVM is far lesser than its counterpart. The future study should analyze the quality of rice and the accuracy of the rice with more accuracy rate and with less computational time when compared to the present study and these are the reasons that may be due to an imbalanced training ratio. It is concluded that the proposed system can effectively classify the type of rice grains.

5. Conclusion

Today's consumers are very concerned about the quality of their food grains. An automated rice grain quality evaluation system based on Machine Learning classifiers has been addressed in this research effort in order to ensure the quality of rice grains. Based on Novel Support Vector Machine and Convolutional Neural Network techniques, the suggested system recognised and categorized the quality analysis of rice grains. Finally, the experimental results demonstrated that the Novel Support Vector Machine has a 67.41 percent accuracy and the Convolutional Neural Network has a 90.38 percent accuracy.

Declarations

Conflicts of interest

No conflicts of interest in this manuscript.

Authors Contribution

Author SIRB was involved in data collection, data analysis, and manuscript writing. Author NBD was involved in conceptualisation, data validation and critical review of manuscripts.

Acknowledgements

The authors would like to express their gratitude towards Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (Formerly known as Saveetha University) for providing the necessary infrastructure to carry out this work successfully.

Funding

We thank the following organizations for providing financial support that enabled us to complete the study.

1. BrainoVision Solutions Pvt. Ltd.
2. Saveetha University.
3. Saveetha Institute of Medical and Technical Sciences.
4. Saveetha School of Engineering.

6. References

D'Adamo, Idiano, Gianluca Lupi, Piergiuseppe Morone, and Davide Settembre-Blundo. 2022. "Towards the Circular Economy in the Fashion Industry: The Second-Hand Market as a Best Practice of Sustainable Responsibility for Businesses and Consumers." Environmental Science and Pollution Research International, February. https://doi.org/10.1007/s11356-022-19255-2.

Edmonston, Phil. 2011. Lemon-Aid Used Cars and Trucks 2011-2012. Dundurn.

Jayabal, Ravikumar, Sekar Subramani, Damodharan Dillikannan, Yuvarajan Devarajan, Lakshmanan Thangavelu, Mukilarasan Nedunchezhiyan, Gopal Kaliyaperumal, and Melvin Victor De Poures. 2022. "Multi-Objective Optimization of Performance and Emission Characteristics of a CRDI Diesel Engine Fueled with Sapota Methyl Ester/diesel Blends." Energy. https://doi.org/10.1016/j.energy.2022.123709.

Jayakodi, Santhoshkumar, Rajeshkumar Shanmugam, Bader O. Almutairi, Mikhlid H. Almutairi, Shahid Mahboob, M. R. Kavipriya, Ramesh Gandusekar, Marcello Nicoletti, and Marimuthu Govindarajan. 2022. "Azadirachta Indica-Wrapped Copper Oxide Nanoparticles as a Novel Functional Material in Cardiomyocyte Cells: An Ecotoxicity Assessment on the Embryonic Development of Danio Rerio." Environmental Research 212 (Pt A): 113153.

Kamath, M. G. 1956. Rice in India.

Kotteeswaran, C., Indrajit Patra, Regonda Nagaraju, D. Sungeetha, Bapayya Naidu Kommula, Yousef Methkal Abd Algani, S. Murugavalli, and B. Kiran Bala. 2022. "Autonomous Detection of Malevolent Nodes Using Secure Heterogeneous Cluster Protocol." Computers and Electrical

Engineering. https://doi.org/10.1016/j.compeleceng.2022.107902.
Krishnan, Anbarasu, Duraisami Dhamodharan, Thanigaivel Sundaram, Vickram Sundaram, and Hun-Soo Byun. 2022. “Computational Discovery of Novel Human LMTK3 Inhibitors by High Throughput Virtual Screening Using NCI Database.” Korean Journal of Chemical Engineering. https://doi.org/10.1007/s11814-022-1120-5.
Li, Y. H., and T. P. Tuong. 2001. Water-Saving Irrigation for Rice: Proceedings of an International Workshop Held in Wuhan, China, 23-25 March 2001. IWMI.
Li, Yunfeng, Peijun Du, Zhaohui Xue, Le Gan, Xin Wang, and Hao Liang. 2018. “A Method of Rice Information Extraction Based on Particle Swarm Optimization SVM Algorithm.” Proceedings of the 2018 10th International Conference on Machine Learning and Computing. https://doi.org/10.1145/3195106.3195155.
Mohan, Harshavardhan, Sethumathavan Vadivel, Se-Won Lee, Jeong-Muk Lim, Nanh Lovanh, Yool-Jin Park, Taeho Shin, Kamala-Kannan Seralathan, and Byung-Taek Oh. 2022. “Improved Visible-Light-Driven Photocatalytic Removal of Bisphenol A Using V2O5/WO3 Decorated over Zeolite: Degradation Mechanism and Toxicity.” Environmental Research. https://doi.org/10.1016/j.envres.2022.113136.
Mohan, Kannan, Abirami Ramu Ganesan, P. N. Ezhilarasi, Kiran Kumar Kondamareddy, Durairaj Karthick Rajan, Palanivel Sathishkumar, Jayakumar Rajarajeswaran, and Lorenza Conterno. 2022. “Green and Eco-Friendly Approaches for the Extraction of Chitin and Chitosan: A Review.” Carbohydrate Polymers 287 (July): 119349.
Oliveira, Antonio Costa de, Camila Pegoraro, and Vívian Ebeling Viana. 2020. The Future of Rice Demand: Quality Beyond Productivity. Springer Nature.
Parhami, Behrooz. 2006. Introduction to Parallel Processing: Algorithms and Architectures. Springer Science & Business Media.
Polster, Burkard. 2012. A Geometrical Picture Book. Springer Science & Business Media.
Poutanen, Kaisa S., Anna O. Kårlund, Carlos Gómez-Gallego, Daniel P. Johansson, Nathalie M. Scheers, Ingela M. Marklinder, Anne K. Eriksen, et al. 2021. “Grains - a Major Source of Sustainable Protein for Health.” Nutrition Reviews, November. https://doi.org/10.1093/nutrit/nuab084.
Saravanan, A., P. Senthil Kumar, B. Ramesh, and S. Srinivasan. 2022. “Removal of Toxic Heavy Metals Using Genetically Engineered Microbes: Molecular Tools, Risk Assessment and Management Strategies.” Chemosphere 298 (July): 134341.
Sathish, T., R. Saravanan, V. Vijayan, and S. Dinesh Kumar. 2022. “Investigations on Influences of MWCNT Composite Membranes in Oil Refineries Waste Water Treatment with Taguchi Route.” Chemosphere 298 (July): 134265.
Sheela, J., Department of Computer Application, National Institute of Technology, Tiruchirappalli, Tamil Nadu, India-, and A. Vadivel. 2016. “Criminal Event Detection and Classification in Web Documents Using ANN Classifier.” International Journal of Signal Processing Systems. https://doi.org/10.18178/ijsps.4.5.382-388.
Shukla, R., P. Strooper, and D. Carrington. 2005. “Tool Support for Statistical Testing of Software Components.” 12th Asia-Pacific Software Engineering Conference (APSEC’05). https://doi.org/10.1109/apsec.2005.105.
Szymanek, Bogdan. 2012. Introduction to Morphological Analysis.
Vivek, J., T. Maridurai, K. Anton Savio Lewise, R. Pandiyarajan, and K. Chandrasekaran. 2022. “Recast Layer Thickness and Residual Stress Analysis for EDD AA8011/h-BN/B4C Composites Using Cryogenically Treated SiC and CFRP Powder-Added Kerosene.” Arabian Journal for Science and Engineering. https://doi.org/10.1007/s13369-022-06636-5.
Yaashikaa, P. R., M. Keerthana Devi, and P. Senthil Kumar. 2022. “Algal Biofuels: Technological Perspective on Cultivation, Fuel Extraction and Engineering Genetic Pathway for Enhancing Productivity.” Fuel. https://doi.org/10.1016/j.fuel.2022.123814.
Yaashikaa, P. R., P. Senthil Kumar, and S. Karishma. 2022. “Review on Biopolymers and Composites – Evolving Material as Adsorbents in Removal of Environmental Pollutants.” Environmental Research. https://doi.org/10.1016/j.envres.2022.113114.
Zhou, Tongtong, Liang Zhang, Ruixuan Zhao, Qiannan Liu, Wei Liu, and Honghai Hu. 2022. “Effects of Particle Size Distribution of Potato Starch Granules on Rheological Properties of Model Dough Underwent Multiple Freezing-Thawing Cycles.” Food Research International 156 (June): 111112.

Tables and Figures

Table 1. The above table represents or shows the group statistical which contains Mean, Std.Deviation and Std.Error Mean of SVM and CNN algorithm using deep learning

	Groups	**N**	**Mean**	**Std.Deviation**	**Std.Error Mean**
Accuracy	**SVM**	10	67.41	1.73309	0.54805
	CNN	10	90.38	3.31705	1.04894

Table 2. The below table shows the independent sample test which consists of Levene's test for equality of variables and t-test equality of means

Accuracy	**Levene's Test for Equality of Variances**		**T-test for Equality of Means**					
	F	**Sig.**	**t**	**df**	**Mean Difference**	**Std. Error Difference**	**95% Conf. Interval Lower**	**95% Conf. Interval Upper**
Equal Variances assumed	5.75	0.001	-19.4	18	-22.970	1.18349	-25.4564	-20.48359
Equal Variances not assumed	5.75	0.001	-19.4	13.5	-22.970	1.18349	-25.5158	-20.42416

Table 3. Accuracy values of SVM algorithm and CNN algorithm (Mean accuracy of SVM and CNN = 67.41&90.3) respectively.

group_id	**SVM**	**CNN**
1	61.00	81.01
1	62.05	81.18
1	63.10	81.22
1	64.19	81.25
1	65.14	82.32
1	66.71	83.55

1	67.51	84.45
1	67.61	84.51
1	67.89	84.71
1	68.13	85.67
2	69.31	85.89
2	69.44	88.99
2	69.51	89.88
2	69.69	90.13
2	69.73	90.24
2	70.01	90.31
2	70.11	90.43
2	70.19	90.91
2	70.24	92.13
2	70.26	94.16

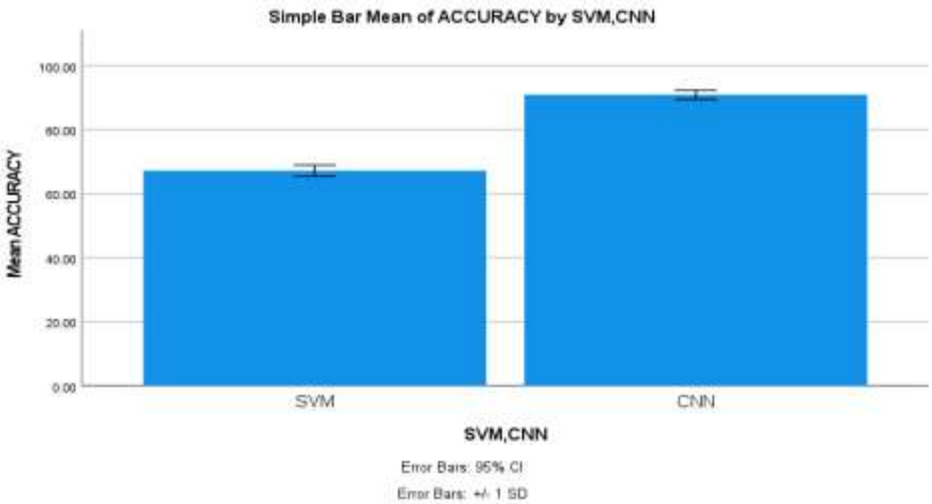

Fig. 1. Simple Bar Graph for Comparison of Accuracy. Mean accuracy of CNN algorithm is better than SVM algorithm and standard deviation of CNN algorithm is better than SVM algorithm. X axis: CNN algorithm vs SVM algorithm, Y axis: Mean accuracy of detection. Error Bars: 95% CI and Error Bars: +/-1 SD.