Survey Paper on Assessing Rice Seed Quality Through Convolutional Neural Networks

Mahesh J. Kanase[1], Ravikant A. Hatgine[2], Bharti Vivek Bandgar[3]
[1,2] *Department of Computer Science & Engineering, DKTE's Yashwantrao Chavan Polytechnic, Ichalkaranji-416115, Maharashtra, India*
[3]*KJ College of Engineering and Management Research, Pune- 411148, Maharashtra, India*

***Abstract*—Seed quality inspection is a foundational requirement for plant nurseries and agricultural operations, as it directly affects seedling growth and overall crop performance. Conventional inspection approaches rely on trained personnel to manually examine seed samples, a process that is costly, slow, and prone to human error — especially when dealing with large volumes.**

This paper presents a machine learning-based framework designed to automate the classification of rice seed quality with higher speed and precision. The proposed system follows a standard machine learning pipeline comprising data collection, model training, validation, and testing. A dataset of 2,000 images was assembled, equally split between high-quality (superior) and low-quality (non-superior) rice seeds, with physical properties such as color and shape forming the basis for classification.

The model was developed and evaluated using Convolutional Neural Network (CNN) techniques on Google Colaboratory, applying cross-validation with an 80:20 training-to-validation split. The resulting Deep CNN model classifies individual rice seed images fed into the system. Evaluation on 30 test samples yielded a precision of 93% and a recall of 95%, indicating strong capability in distinguishing between superior and non-superior seeds.

***Index Terms*—Artificial Intelligence, CNN, Rice Seed Classification, Deep CNN, Supervised Learning**

I. INTRODUCTION

Rice stands as one of the world's most essential agricultural commodities, serving as the primary dietary staple for a large portion of humanity — particularly across Asian countries like Indonesia. Because of its central role in nutrition, ensuring the quality of rice (Oryza sativa) seed varieties is a critical step in sustainable agricultural production.

One of the recurring challenges in rice cultivation is maintaining the varietal integrity of seed batches. When seeds from different varieties are inadvertently mixed, it introduces what is known as varietal contamination — a problem that can give rise to weeds, off-types, unwanted crossbreeding, and genetic mutations, ultimately degrading crop yields. Traditionally, trained specialists have been responsible for detecting these impurities, using basic tools such as forceps, magnifying glasses, and flashlights to examine seed morphology, color, and texture.

However, the differences between many paddy seed varieties are subtle, making accurate visual identification both challenging and time-consuming. The limitations of manual inspection become more pronounced when inspectors must process thousands of seed samples simultaneously, where even minor perceptual errors can lead to significant quality control failures.

Advancements in technology, particularly those driven by the fourth industrial revolution, are increasingly being adopted in the agricultural sector to address these challenges. Technologies designed for automated quality evaluation can substantially improve the speed and consistency of seed quality testing while also enabling the development of export-grade processing standards.

Computer vision has seen widespread adoption across many domains over the past decade, and deep learning has emerged as a powerful tool within this field — outperforming traditional statistical approaches in tasks like image recognition and object detection. Unlike conventional machine learning models that require manual feature engineering, deep learning frameworks can automatically extract complex patterns and features from raw image data, making them well-suited for visual classification tasks.

This study employs supervised learning through a Convolutional Neural Network architecture to develop a system capable of identifying the quality of individual rice grains from digital images. The goal is to enable a reliable, scalable, and automated alternative to manual inspection, with the potential to support large-scale agricultural quality assurance operations.

II. LITERATURE REVIEW

A range of research efforts have explored machine learning and image analysis techniques for evaluating rice seed quality. The following section reviews key contributions in this area.

[1] Chatnuntawech et al. (2018) proposed a novel rice classification approach using a Spatio-Spectral Deep Convolutional Neural Network (SS-DCNN). By combining both spectral and spatial characteristics of high-dimensional image data, the model demonstrated superior accuracy compared to conventional classification methods. The study also highlighted the model's scalability, suggesting its applicability in food quality assessment and agricultural automation.

[2] Kiratiratanapruk et al. (2020) developed a paddy rice seed classification pipeline specifically designed for integration with automatic grading machines. The study explored several machine learning algorithms targeting key seed attributes such as size, shape, and color. By combining advanced feature extraction with optimized classification models, the system achieved strong accuracy and was shown to be effective for large-scale agricultural quality control.

[3] Fabiyi et al. (2020) investigated the use of both RGB and hyperspectral imaging for varietal classification of rice seeds. The study demonstrated that combining these two imaging modalities captured complementary features — including texture, color, and spectral reflectance — and achieved better classification accuracy than either technique used in isolation. The authors highlighted its suitability for automated seed quality assessment.

[4] Gupta and Neelam (2015) proposed a computer vision-based approach for identifying rice varieties using the Mahalanobis Distance metric. After acquiring and preprocessing rice grain images, morphological and texture features were extracted and classified using this statistical measure. The results confirmed that computer vision combined with statistical methods can effectively distinguish rice varieties.

[5] Joshi et al. (2021) introduced a non-destructive, label-free approach using Optical Coherence Tomography (OCT) alongside a deep neural network for rice seed classification. The system leveraged OCT's ability to capture high-resolution internal microstructures of seeds, which were then analyzed by the deep learning model. Experimental results showed reliable classification without requiring any external labeling or physical damage to the samples.

[6] Ahmed et al. (2021) presented a dual-phase CNN system for identifying diseases in rice grains, specifically designed to perform well on small datasets. The first phase focused on feature extraction while the second refined the classification step. Data augmentation was used to compensate for limited training data, and the system demonstrated competitive accuracy and robustness even under constrained data conditions.

[7] Mookdarsanit and Mookdarsanit (2021) introduced PhosopNet, an improved grain localization and classification system that leverages advanced image augmentation strategies. The approach addressed common challenges in grain detection by boosting model generalization through extensive augmentation. Evaluation results showed significant improvements over prior methods in both detection accuracy and classification performance.

[8] Lurstwut and Pornpanomchai (2017) proposed an image analysis technique for evaluating rice seed germination based on three visual features: color, shape, and texture. Using digital image processing, the system extracted these features and assessed germination viability with high accuracy, offering a practical, non-invasive alternative to conventional germination evaluation methods.

[9] Aznan et al. (2017) developed a rice seed variety identification system that combined color feature extraction from images with an Artificial Neural Network (ANN) for classification. The results showed effective identification of different seed

varieties, illustrating how the pairing of image processing and neural networks can be applied to automated agricultural tasks.

[10] Aukkapinyo et al. (2020) proposed a region proposal-based CNN method for simultaneously localizing and classifying rice grains in images. By incorporating region proposals to identify areas of interest before applying CNN-based classification, the system achieved high efficiency and accuracy across various grain types and imaging conditions.

[11] Kido et al. (2018) explored the use of CNN and Regions with CNN Features (R-CNN) for detecting and classifying lung abnormalities in medical images. The system applied CNN for feature extraction and used R-CNN to focus analysis on regions likely to contain pathology, improving diagnostic accuracy. While primarily a medical imaging study, its methodology provides relevant insights for region-based classification tasks in other domains.

[12] Liu et al. (2021) applied a LASSO (Least Absolute Shrinkage and Selection Operator) logistic regression model combined with hyperspectral imaging to evaluate the purity of rice seeds. Spectral features extracted from hyperspectral images were analyzed to distinguish pure seeds from contaminated ones. The method proved effective for automated seed purity assessment and quality certification in agriculture.

III. PROPOSED WORK

The framework adopted in this study for rice seed quality assessment is built on supervised learning, using a Convolutional Neural Network (CNN) as the core classification architecture. CNN-based techniques are incorporated into the supervised learning pipeline to enable the system to classify rice quality at the level of individual grains. Rice seed images serve as the input, and the system produces a corresponding quality category as its output.

1. Classification System

The quality detection system operates through two distinct phases. The first is the Model Training Phase, in which a dataset of labeled rice seed images is used to build a classification model based on the visual characteristics of the seeds. The second is the Testing Phase, where images with unknown quality classifications are passed through the trained model to generate predictions. This two-stage structure ensures that the model is well-calibrated before it is deployed for real-world quality evaluation.

2. Image Acquisition

A dedicated image acquisition setup was used to capture photographs of rice seeds for both training and evaluation purposes. Images were taken using a camera mounted on a three-foot support brace to maintain a stable and consistent shooting angle. Each seed was placed against a white background, and a fixed camera-to-subject distance was maintained throughout the process to ensure uniform image quality and reduce variability in the captured data.

3. Dataset

The dataset used in this study consists of labeled rice seed images from the Situbagendit local variety. It is divided into two categories — superior seeds and non-superior seeds — each represented across the dataset columns. Prior to training, a data cleaning step was performed to remove irrelevant or erroneous entries, ensuring that only high-quality data was used for model development. The integrity of the dataset is particularly important, as noisy or inaccurate data can adversely affect both model accuracy and generalization. A clean dataset supports more reliable predictions and stronger overall system performance.

4. Supervised Learning

Supervised learning is a category of machine learning in which a model learns from labeled training data to make predictions on new, unseen inputs. This study applies supervised learning to classify rice seeds as either superior or non-superior based on their visual properties. The machine learning algorithm processes the labeled dataset and extracts patterns associated with seed size, shape, and appearance. These patterns form the foundation of the predictive model, which can then classify previously unseen seed images into the appropriate quality category.

IV. CONCLUSION

This study demonstrates that automated rice seed quality evaluation is achievable through digital image processing combined with a Convolutional Neural Network approach. By analyzing the physical characteristics of individual rice grains — including color and shape — the proposed system can reliably classify seeds as superior or non-superior without manual intervention.

The model was trained and validated on the Google Colaboratory platform and achieved an overall accuracy of 88%, with a precision of 93% and a recall of 95%. These metrics confirm the system's effectiveness for practical quality assessment tasks.

While the results are encouraging, there remains room for improvement. Future work should focus on expanding the dataset, exploring additional CNN architectures, and refining the model to handle greater diversity in seed types and imaging conditions. With further development, this technology has strong potential to support scalable, accurate, and cost-effective rice seed quality management in agricultural settings.

REFERENCES

[1] I. Chatnuntawech, K. Tantisantisom, P. Khanchaitit, T. Boonkoom, B. Bilgic, and E. Chuangsuwanich, "Rice Classification Using Spatio-Spectral Deep Convolutional Neural Network," Comput. Vis. Pattern Recognit. Cornell Univ., vol. 3, no. 1, pp. 1–22, 2018. [Online]. Available: http://arxiv.org/abs/1805.11491

[2] K. Kiratiratanapruk et al., "Development of Paddy Rice Seed Classification Process using Machine Learning Techniques for Automatic Grading Machine," J. Sensors, vol. 2020, 2020, doi: 10.1155/2020/7041310.

[3] S. D. Fabiyi et al., "Varietal Classification of Rice Seeds Using RGB and Hyperspectral Images," IEEE Access, vol. 8, no. 1, pp. 22493–22505, 2020, doi: 10.1109/ACCESS.2020.2969847.

[4] J. Gupta and Neelam, "Identification and Classification of Rice Varieties using Mahalanobis Distance by Computer Vision," Int. J. Sci. Res. Publ., vol. 5, no. 5, pp. 1–6, 2015.

[5] D. Joshi et al., "Label-free non-invasive classification of rice seeds using optical coherence tomography assisted with deep neural network," Opt. Laser Technol., vol. 137, p. 106861, 2021, doi: 10.1016/j.optlastec.2020.106861.

[6] T. Ahmed, C. R. Rahman, and M. F. M. Abid, "Rice Grain Disease Identification Using Dual Phase Convolutional Neural Network Based System Aimed at Small Dataset," agriRxiv, vol. 2021, no. 1, pp. 1–14, 2021, doi: 10.31220/agrirxiv.2021.00062.

[7] P. Mookdarsanit and L. Mookdarsanit, "PhosopNet: An improved grain localization and classification by image augmentation," Telkomnika, vol. 19, no. 2, pp. 479–490, 2021, doi: 10.12928/TELKOMNIKA.v19i2.18321.

[8] B. Lurstwut and C. Pornpanomchai, "Image analysis based on color, shape and texture for rice seed (Oryza sativa L.) germination evaluation," Agric. Nat. Resour., vol. 51, no. 5, pp. 383–389, 2017, doi: 10.1016/j.anres.2017.12.002.

[9] A. A. Aznan, R. Ruslan, I. H. Rukunudin, F. A. Azizan, and A. Y. Hashim, "Rice seed varieties identification based on extracted colour features using image processing and ANN," Int. J. Adv. Sci. Eng. Inf. Technol., vol. 7, no. 6, pp. 2220–2225, 2017, doi: 10.18517/ijaseit.7.6.2990.

[10] K. Aukkapinyo, S. Sawangwong, P. Pooyoi, and W. Kusakunniran, "Localization and Classification of Rice-grain Images Using Region Proposals-based Convolutional Neural Network," Int. J. Autom. Comput., vol. 17, no. 2, pp. 233–246, 2020, doi: 10.1007/s11633-019-1207-6.

[11] S. Kido, Y. Hirano, and N. Hashimoto, "Detection and classification of lung abnormalities by use of CNN and R-CNN," in Proc. 2018 Int. Workshop Adv. Image Technol. (IWAIT), 2018, pp. 1–4, doi: 10.1109/IWAIT.2018.8369798.

[12] W. Liu, S. Zeng, G. Wu, H. Li, and F. Chen, "Rice seed purity identification technology using hyperspectral image with lasso logistic regression model," Sensors, vol. 21, no. 13, p. 4384, 2021, doi: 10.3390/s21134384.