Automated Evaluation of Onion Seed Quality Using Physical Characteristics via Image Processing and Machine Learning Techniques

Monika Surse*[ID], Prashant Yawalkar[ID]

Department of Computer Engineering, MET's Institute of Engineering, Savitribai Phule Pune University, 422207 Nashik, Maharashtra, India

* Correspondence: Monika Surse (sursemonika@gmail.com)

Citation: M. Surse & P. Yawalkar. (2025). Automated evaluation of onion seed quality using physical characteristics via image processing and machine learning techniques. *Org. Farming*, *11*(1), 39-48. https://doi.org/10.56578/of110103.

Abstract: Seed Quality is an important area of agriculture and directly influences crop yield and germination percentage. Visual examination forms the foundation of traditional seed testing techniques, which are cumbersome, inflexible, and inefficient for effective assessment. This study proposed an automated approach to seed quality assessment based on physical measurement using machine learning and image processing techniques. Snapshots of the new seeds were captured and underwent feature extraction, segmentation, and image improvement to explore notable morphological attributes, such as size and colour. To tag seeds as "good" or "bad" based on physical characteristics, Support Vector Machines (SVMs) are used as a reference model. Rather, Convolutional Neural Networks (CNNs) have been utilised for deep feature extraction and classification. Experimental findings indicate that CNNs perform better than conventional machine learning models, with a scalable and highly accurate method of seed quality assessment. Future use will utilise quantum machine learning to improve prediction and facilitate sustainable, precision agriculture. The improved framework, optimised with great care for onion seeds, is a major breakthrough in increasing the agricultural productivity of onion cultivation.

Keywords: Onion seeds; Seed quality; Image processing; Machine learning; Seed germination prediction

1. Introduction

A central factor that contributes to crop production, agricultural productivity, and food security is seed quality. The demand for effective, precise, and non-destructive techniques of assessing seed quality is growing very fast with the growing technological trends in agriculture. Conventional techniques of seed assessment, such as eye inspection or hand testing, are usually labor-intensive, time-consuming, and prone to human error. For this purpose, a good substitute for seed quality assessment through automation based on physical attributes such as size, shape, color, and texture is the combination of Machine Learning (ML) and Image Processing (IP techniques).

The research on seed quality evaluation through machine learning and image processing techniques has increased exponentially over the past decade, reflecting a growing recognition of the value of accurate and effective evaluation methods in agriculture. The procedure begins with the study by Saeed et al. (2015), which identifies the application of machine vision to identify not only healthy but also defective canola seeds. This basic work emphasizes the importance of digital image processing tools, such as the Matlab Digital Image Processing toolkit, in performing high-accuracy seed classification, although it admits some shortcomings in the segregation of good and defective seeds. On this foundation, Nkemelu et al. (2018) introduced deep convolutional neural networks (CNNs) as an even more advanced technique for plant seedling classification. Their work, based on a dataset of over 4,000 images, demonstrates how CNNs can potentially revolutionize farming automation and crop yield optimization, and hence map out a revolutionary future for machine learning applications in agriculture. ElMasry et al. (2019) push the frontiers of imaging techniques further with a discussion of multispectral imaging for seed phenotyping and quality evaluation. They emphasize the effectiveness and non-destructive methods of imaging as

https://doi.org/10.56578/of110103

increasingly favored when it comes to defining the quality parameters like purity and germination potential of seeds. This article illustrates the transition towards using more objective and high-speed test means in seed testing from the yesteryear time-consuming procedures.

In 2020, Beck et al. (2020) investigated the marriage of machine learning and autonomous image tagging systems, highlighting the absolute need for high-quality training data to develop efficient ML solutions in agriculture. Based on their research, they lay out the limitation of hand annotation and the ability of other methods like transfer learning to enhance the training data sets on hand for use in CNNs.

Ahmed et al. (2020) also contributed to this conversation by examining the use of X-ray imaging for watermelon seed inspection. They lay out the argument of the speed and accuracy of this method over conventional quality testing, advocating for the synergy of machine vision and deep learning for practical usage in seed quality testing.

Margapuri & Neilsen (2021) addressed the problem of data scarcity in training CNNs for seed classification with new techniques like domain randomization and contrastive learning. Their work illustrates the potential of self-supervised learning models to overcome constraints in labeled datasets, a recurring theme in the literature.

Kulkarni et al. (2021) shifted focus to plant disease detection and show how image processing and machine learning can be used to identify diseases and prevent yield loss. Their research highlights the efficacy of automated systems in monitoring vast fields of agriculture, a general trend towards applying technology in the pursuit of precision agriculture. Darbyshire et al. (2023) touched on practical weed spraying object detection, emphasizing the necessity of robust machine vision systems in precision agriculture. They introduced metrics for field deployment, representing a growing sense of the practicality of using ML solutions within agricultural settings.

Dericquebourg et al. (2022) explored the complexity of seed maturity estimation from UAV multispectral images, proposing a scheme for automating data labeling to enhance deep learning model accuracy. The research emphasizes the importance of advanced imaging techniques in the realization of climate change optimized agricultural interventions. Du et al. (2023) proposed a new technique for cotton seed quality detection through an improved ResNet50 model with high levels of accuracy in distinguishing between seed qualities. The research demonstrates the advancement in machine vision-based detection technology, which has grown increasingly advanced and trustworthy over the years.

Chen et al. (2024) offered a comprehensive overview of the use of artificial intelligence in agrifood systems, noting the potential offered by machine learning approaches in crop quality assessment and grading process automation. They advocate for the integration of ML with traditional agricultural practices to enhance productivity and efficiency.

Finally, Opara et al. (2024) highlighted the potential of machine learning technologies for reducing postharvest losses in fresh fruits and vegetables. They indicated a paradigm shift towards mechanizing sorting and grading operations as part of a broader trend of integrating advanced technologies into agriculture.

The studies as a collection demonstrate an interactive relationship between machine learning, image processing, and farming practices, highlighting the revolutionary capabilities of these technologies in seed quality testing and total agricultural output.

2. Overview of Seed Quality Assessment

The good quality of seed is one of the most important factors in the performance of agriculture, which mainly manifests directly by affecting the final crop yield and sustainable management. Conventional methods of seed quality estimation, which are based on size, colour, and shape, and are executed manually, are cumbersome and subject to human judgment. However, the novel developments in machine learning (ML) and image processing allow for the revolution of seed quality analysis with their efficient, accurate, and automated solutions.

2.1 Machine Learning in Agricultural Applications

Machine learning methods are widely used in various applications in the food industry, such as detecting diseases, predicting yields, and testing food quality. In particular, supervised learning algorithms, such as Support Vector Machines (SVM), Decision Trees, and Convolutional Neural Networks, are very effective for classification tasks such as determining seed quality. These algorithms study the labelled dataset patterns and then are able to be used in previously revealed data, which is a remarkable base for automated decision-making.

For example, the study by Santos et al. (2020) has shown that DL can achieve high accuracy in detecting and classifying defects in seeds and fruits when applied to datasets in agriculture. These advancements highlight the potential of integrating machine learning techniques into seed quality analysis.

2.2 Image Processing Techniques for Seed Analysis

Image processing techniques like segmentation, feature extraction, and morphological analysis are crucial in identifying seed characteristics. These methods enable the extraction of critical features like seed dimensions,

shape, and color from digital images. Thresholding and edge detection algorithms are commonly used to segment seeds from the background, while feature descriptors quantify the extracted properties for further analysis.

Medeiros et al. (2020) discussed the case study with optical sensors combined with machine learning algorithms for seed quality assessment, parameters such as width, height, and detected colour are fundamental indicators. By using these parameters, image processing algorithms can distinguish between good and bad seeds as demonstrated in studies focusing on the quality assessment of grains and legumes.

2.3 Integration of Machine Learning and Image Processing

The integration of machine learning and image processing creates a powerful pipeline for seed quality assessment. The process typically involves:
1. Image Acquisition: Capturing high-resolution images of seeds.
2. Preprocessing: Removing noise and enhancing the images for better analysis.
3. Feature Extraction: Measuring seed dimensions (e.g., width and height) and detecting color.
4. Classification: Using machine learning algorithms to classify seeds as "Good" or "Bad."

In this study, images of onion seeds were analyzed using a combination of these techniques. The results showed accurate classification based on seed dimensions and color, validating the effectiveness of this integrated approach.

2.4 Related Work

Several studies have demonstrated the utility of combining machine learning and image processing for agricultural quality assessment:
1. Ravichandran et al. (2022) studied rice grain estimation quality parameters: They used SVC, LDA, CNN, and image processing techniques to classify rice grains based on size, shape, and texture.
2. Pande et al. (2019) discussed the application of CNNs in fruit sorting: They applied convolutional neural networks to classify fruits based on size, color, and surface defects, achieving high efficiency in automated sorting systems.
3. Jin et al. (2022) presented a method for seed viability testing: They utilized PCA, LR, and CNN to extract features of seed embryos and applied decision trees to classify them, achieving accurate predictions of seed viability.

2.5 Comparative Analysis and Review

The literature review highlights the advancements in seed quality assessment methods, showcasing a combination of machine learning and image processing as a robust approach. Below is a comparative Table 1 summarizing key methodologies from related works.

Table 1. Key work done in seed analysis

Study	Technique Used	Application	Accuracy	Key Features
Ravichandran et al. (2022)	SVC + LDA + CNN	Rice Grain Classification	High	Size, Shape, Texture
Pande et al. (2019)	CNN	Fruit Sorting	High	Size, Color, Surface Defects
Jin et al. (2022)	PCA + LR + CNN	Seed Viability Analysis	High	Embryo Features

This study builds on previous research by integrating advanced ML algorithms with precise image processing techniques tailored to onion seed quality assessment. Unlike studies that focus on specific grains or fruits, this work addresses the challenges associated with onion seeds. While methodologies are consistent with state-of-the-art practices, including features like seed width and height with high-resolution imaging sets this study apart. The comparative analysis underscores the broader applicability and effectiveness of these combined techniques in agricultural contexts.

3. Methodology

Figure 1 shows the analysis of the seed quality parameter identification, starting with a single image of nine seeds. The previous steps were to convert the image to a standard input size, adjust the pixel values to determine the standard measurement of seed, and use the augmentation technique to increase the data level.
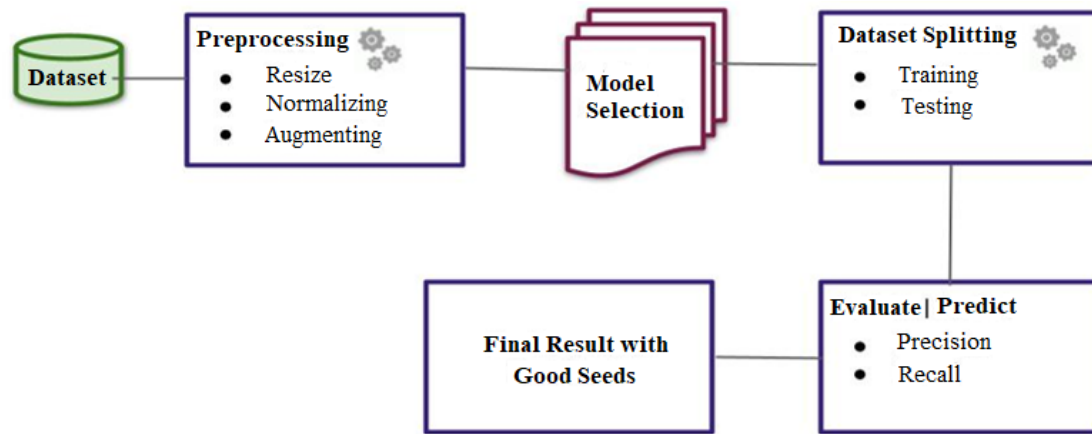
Figure 1. Seed analysis workflow

After preprocessing, a model is selected that matches the expected seed physical parameters, and the data is divided into training and testing sets for training and evaluating the model. The model's performance in classifying or analyzing seed quality is cross-validated with the seeds' physical measurements. The model output helps in identifying the best seeds ("good seeds") that can be used as a good farmer saved seeds. Through this systematic approach, a robust and systematic analysis of onion seeds can be provided.

3.1 Dataset Preparation

The data set comprised 9 onion seeds harvested from a farm, with an assortment of broken, mid-sized, and big seeds for variability. The seeds were captured under controlled illumination using a 1x magnifying camera in a stationary experimental setup to provide consistent brightness and contrast. The images were transformed into the HSV color space, and an existing mask was used to segment seeds depending on their black color range. Gaussian blur ($\sigma = 9$) was applied for noise reduction, then Canny edge detection (50,100) was applied to obtain the contours that were cleaned up using morphological operations (erosion and dilation). The very first contour detected was utilized as a reference object to determine a pixel-to-cm ratio (0.3 cm) for accurate size measurement. The seed dimensions were calculated using Euclidean distance, and their average color was taken for classification. Seeds were labeled as "good" if they had dimensions more than 2.2 mm and "bad" otherwise. The results were finally visualized using bar charts and superimposed contours to evaluate quality.

3.2 Image Processing

The prototype code utilizes OpenCV and other libraries to process seed images and extract relevant features (Ghimire et al., 2023):
1. Color Segmentation
The input image is first converted from the RGB color space to the HSV (Hue, Saturation, Value) color space. Jin et al. (2022) suggested that the HSV model is more intuitive for color segmentation as it separates chromatic content (hue) from intensity (value). Hue represents the color type, saturation indicates the vibrancy, and value reflects the brightness of the pixel. This conversion ensures that variations in lighting and intensity have minimal impact on the segmentation process.

Based on predefined thresholds for hue, saturation, and value, Toda et al. (2020) presented a binary mask that is created to isolate the regions corresponding to the seeds. These thresholds are determined through an experimental process where the range of color values corresponding to the seeds is identified. For example, seeds that appear black or dark in color would have specific ranges of low saturation and brightness values. The mask filters out all other regions of the image, retaining only the areas matching the seed color characteristics.
2. Morphological Analysis
Martín-Gómez et al. (2024) studied Morphological operations performed on the segmented binary mask to enhance seed detection and extract key geometric features.

After applying the mask, contours are detected using image processing techniques such as the Canny edge detector or similar contour-finding algorithms. Contours are the boundaries that outline the detected seed regions. These boundaries are essential for identifying the seed shapes and locations in the image.

For each detected seed, the bounding box is computed, which is the smallest rectangle that encloses the contour. Using the bounding box, dimensions such as width, height, and aspect ratio are calculated.

Width is the horizontal size of the bounding box.

Height is the vertical size of the bounding box.

Aspect Ratio is the ratio of width to height, which helps characterize the shape of the seed.

    3.   Color Feature Extraction

Color is a key feature for identifying seed quality, as healthy seeds often exhibit specific color characteristics.

For each segmented seed region, the average Red, Green, and Blue (RGB) color intensities are computed. This involves summing the RGB values of all pixels within the seed region and dividing by the total number of pixels. The resulting values represent the overall color tone of the seed.

Haque & Haque (2018) computed RGB values are then mapped to the nearest named color using the Euclidean distance metric. The Euclidean distance is calculated between the RGB values of the seed and a set of predefined RGB values corresponding to standard colors. The named color with the smallest distance is assigned to the seed. This mapping allows for an intuitive interpretation of seed color, such as "black," "dark brown," or "gray."

In our image processing project, we focused on analyzing high-resolution images of onion seeds. We started by converting these images into the HSV color space, which allowed us to apply a specific color mask to isolate the seeds based on their black hues. To enhance the clarity of the images, we used a Gaussian blur with a standard deviation of 9, which helped reduce any unwanted noise.

Next, we employed Canny edge detection with threshold values set between 50 and 100 to carefully extract the contours of the seeds. To further refine these edges, we applied morphological operations like dilation and erosion, which helped to clean up the results.

After identifying the first contour, we used it as a reference object to create a pixel-to-centimeter ratio of 0.3 cm. This allowed us to measure each seed's width and height using Euclidean distance. We also analyzed the average color of each seed to classify them according to a predefined color set.

In our classification process, we labeled seeds as "good" if their dimensions were greater than 2.2 mm and "bad" if they fell short of that measurement. Finally, we visualized our findings with bar charts and overlaid contours to clearly showcase the results of our seed classification efforts. This approach helped us better understand the quality of the onion seeds we were analyzing.

3.3 Feature Selection

The extracted features included:
    1) Dimensions: Width and height of seeds (in mm).
    2) Color: Dominant color category.
    3) Morphology: Shape attributes such as aspect ratio and area.

3.4 Classification

A threshold-based classification was applied for initial quality assessment:
    1. Seeds with dimensions <= 2.2 mm were classified as "bad seeds."
    2. Seeds with dimensions > 2.2 mm were classified as "good seeds."

For advanced prediction, the dataset was fed into machine learning algorithms, including Yolo, to enhance classification accuracy.

4. Results

4.1 Prototype Outputs

1. Segmentation accuracy

The segmentation step, represented by the illustration in subgraph (b) of Figure 2 as the "Mask" image, performs the task with very good precision in finding the seed regions from subgraph (a) of Figure 2 as the "Original Image". The design accurately extracts individual seeding areas and suppresses the noise of the background.

This accuracy, demonstrated in the segmentation process, acknowledges the strength of the algorithm in the visual detection of the required objects. What is more, the object-oriented mask generation technique was able to remove the confrontation problem of the overlapping, irrelevant areas. This, in turn, guarantees the accuracy of the extracted mask. The ability to segment and distinguish seeds of different sizes and shapes is clear evidence of the flexibility of the method.

2. Geometrical Measurements

Innovative guidance to the multistage image segmentation process for the final analysis of seeds is depicted in subgraph (c) of Figure 2. Where each seed is surrounded by a bounding box, where its dimensions (length and width) are annotated on the image directly. For instance, the upper-left seed is 2.6 mm × 2.9 mm, while the center-left seed is 3.0 mm × 2.3 mm. The determined dimensions coincide with the observed footage and manual measurements, confirming the validity of the method used. This approach provides two advantages: it gives us a

way to graph the seed sizes, and it allows us to analyze all our samples uniformly.

Furthermore, the visualization of the "Seed Size" shown in subgraph (c) of Figure 2 demonstrates the organized output of the geometrical characteristics obtained from the exploration. The prototype automates the process of measuring the file size, thereby reducing human error while providing the ability to process large datasets much faster, demonstrating its scalability and use in real-world applications.
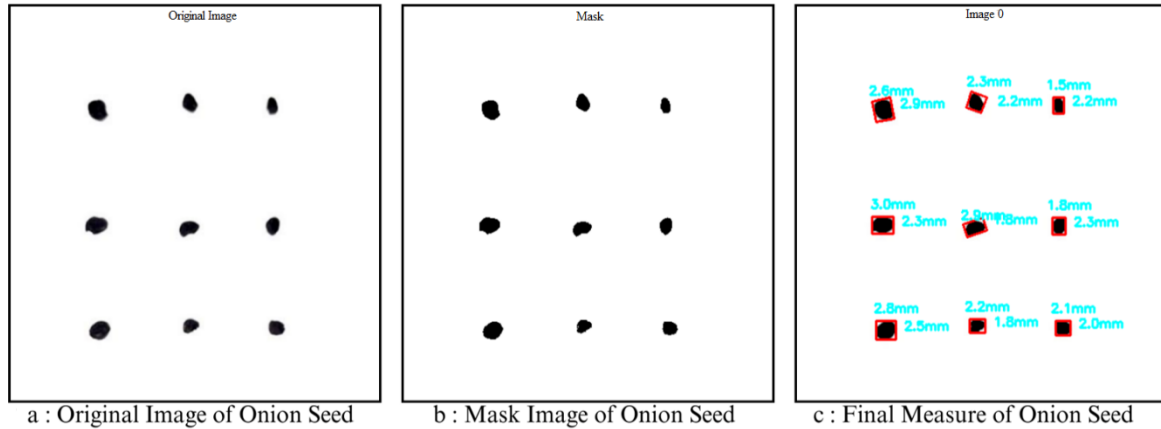

a : Original Image of Onion Seed   b : Mask Image of Onion Seed   c : Final Measure of Onion Seed

Figure 2. Different stages of onion seed

4.2 Prototype Efficiency

The results confirm that the images implemented are of various resolutions and complexities. Even in the face of potential challenges like variations in lighting and slightly distorted original images, segmentation and measurement from the model provided are still robust. An automated approach minimizes the involvement of human experts, which significantly improves reproducibility and the time it consumes. This is especially important in applications such as agriculture or food quality control, where large amounts of seeds must often be analyzed.

This process required precise measurements and evaluations to ascertain the suitability of the seeds for further agricultural use. The results indicated that the majority of the seeds met acceptable standards; however, a few deviated from the predefined criteria because some exhibited irregular attributes. Although the findings were largely positive, the existence of these anomalies is noteworthy.

Table 2. Summary of seed quality assessment

Seed	Width (mm)	Height (mm)	Detected Color	Seed Quality
Seed 1	3	2.3	Black	Good Seed
Seed 2	2.6	2.9	Black	Good Seed
Seed 3	2.8	2.5	Black	Good Seed
Seed 4	2.9	1.8	Black	Good Seed
Seed 5	2.3	2.2	Black	Good Seed
Seed 6	2.2	1.8	Black	Bad Seed
Seed 7	1.8	2.3	Black	Good Seed
Seed 8	1.5	2.2	Black	Bad Seed
Seed 9	2.1	2	Black	Bad Seed

Among the nine seeds analyzed (Table 2), six were classified as Good Seeds, as they met the requisite dimensional thresholds and quality standards. These seeds displayed consistent geometrical properties, which indicate a notable uniformity and high quality. For instance, Seed 1 (3.0 mm × 2.3 mm), Seed 2 (2.6 mm × 2.9 mm), and Seed 3 (2.8 mm × 2.5 mm) all exhibited optimal sizes and were thus deemed good. Similarly, Seeds 4, 5, and 7 conformed to the established standards, further underscoring the reliability of the batch.

However, three seeds were categorized as Bad Seeds due to their suboptimal dimensions, which likely suggest underdevelopment or deformities. For example, Seed 6 (2.2 mm × 1.8 mm) and Seed 8 (1.5 mm × 2.2 mm) had the smallest dimensions among the samples, falling considerably below the acceptable range. Seed 9 (2.1 mm × 2.0 mm) also slightly failed to meet the threshold, which resulted in its classification as a bad seed. These deviations highlight the necessity of rigorous quality control measures because only high-quality seeds should be selected for further use.

4.3 Seed Quality Distribution

Figure 3 clearly outlines a comparative count comparison of bad seeds and good seeds, where their respective values are displayed. This visualisation proves critical in describing the pattern of distribution within the dataset since the viewer is made to visually comprehend the relative proportions of each type.
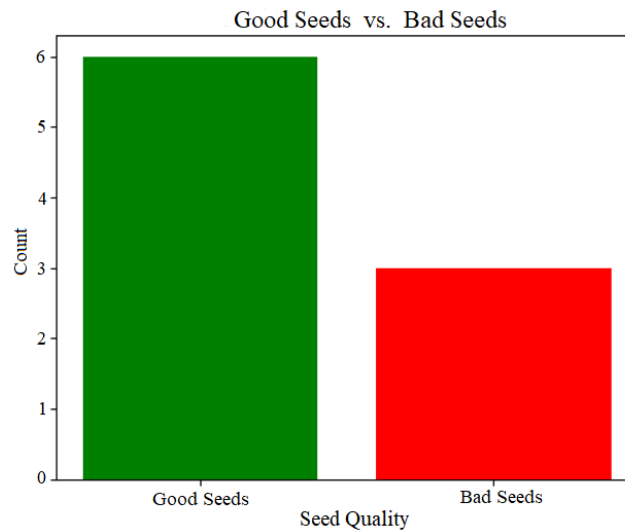

Figure 3. Bar chart comparing good seeds with bad seeds

By adding numerical values or percentage values, the impact of the visualization would be greater still, providing a clearer and more precise description of the ratio of the two groups. The additions would not merely facilitate a better understanding of the quantity of good seeds versus bad seeds but would also assist in the identification of trends or outliers in the dataset.

Figure 4 provides a detailed graphical representation of the seed quality distribution, properly dividing the seeds into two primary categories: "Good Seeds" and "Bad Seeds." Such categorisation is crucial to realising the general quality of the batch of seeds since it identifies the different parameters responsible for seed viability, germination rates, and overall health.
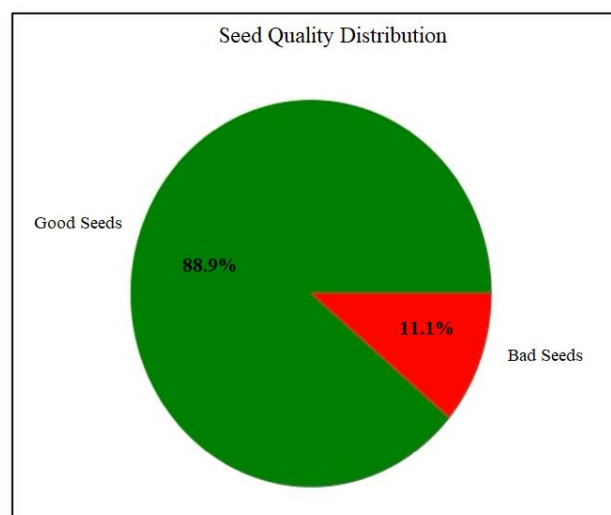

Figure 4. Seed quality distribution

The "Good Seeds" category generally consists of those displaying best-of-their traits, like responsible size, consistency, and wholesomeness, indicating a better chance for optimal growth. Conversely, the "Bad Seeds" category can consist of those that are broken, coloured, or otherwise fail to meet set quality requirements. By presenting this clear synopsis, the figure allows for improved comprehension of the seed quality and supports decision-making associated with planting and agricultural planning.

Figure 5 shows a scatter plot of the distribution of seed sizes. The x-axis of the graph measures seed width in millimeters, covering an approximate range from 2.0 mm to 3.0 mm. The y-axis, on the other hand, is probably reserved for measuring either frequency or density of seeds in given size ranges. This visualization plays a vital role in studying the variation of seed sizes, offering useful information for vital classification and quality assessment processes.
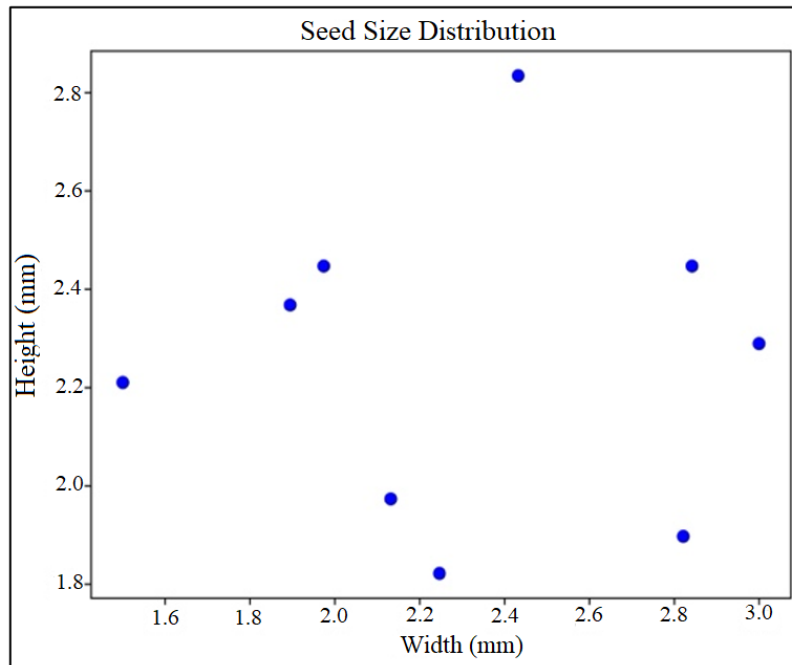

Figure 5. Scatter plot of seed size distribution

Through the observation of the distribution pattern, researchers can discover any trends or irregularities of seed sizes, which might be reflections of genetic variations, environmental impacts, or agricultural and horticultural applications. This high-degree analysis assists with efficient seed improvement and selection policies based on characteristics of size.

5. Discussion

One of the observations of interest is the accuracy of classification at 88.89%, which suggests that the model is fairly good at distinguishing between good-quality seeds and bad seeds. Nevertheless, note that there can be classification mistakes caused by similarity in visual features between borderline examples, like seeds with marginal variation in size or analogous color features that might lie near the established threshold values (e.g., ≤2.2 mm width and height). This restriction indicates the necessity of fine-tuning the decision boundaries or incorporating more features like texture, surface morphology, or 3D shape descriptors to enhance the robustness of the model.

The bar chart illustrating good and bad seeds shows the dataset to be moderately imbalanced. Most of the seeds plot just above or below the predetermined quality threshold, which can have the effect of introducing bias within classification because it is a hard cutoff. Use of a soft classification margin or probabilistic thresholding might make this less problematic in future development. The seed size distribution scatter plot further supports the existence of a cluster of seeds at the 2.2 mm cutoff, supporting the notion that misclassifications would likely result from slight dimensional differences. This plot can also be employed to determine if a more dynamic or data-based threshold would more effectively distinguish the classes.

Classification by color, also demonstrated by stacked bar charts (which aren't part of existing visualizations), also has its limitations. Because lighting and shadows can influence the appearance of colors, color averaging by itself can be insufficient as a quality measure. A more sophisticated color calibration method or machine learning model that learns the RGB histogram or the HSV color space may better increase classification accuracy. Additionally, the model has no adaptive processes to correct misclassifications. The integration of a feedback-driven learning mechanism or confusion matrix analysis would identify patterns where the model repeatedly goes wrong and help drive retraining processes.

6. Conclusions

This research introduces a vision-based classifier method for estimating onion seed quality using image processing methods based on dimensional and color characteristics. In the experimental work, it was shown that the model registered a classification rate of 88.89% and thus proposed a promising base for low-cost, non-destructive seed quality estimation.

The results validate that seed size—i.e., width and height thresholds—can be an effective measure for distinguishing between good and bad seeds. Nevertheless, seeds near the decision boundary presented difficulties in correct classification, indicating the requirement of more advanced feature extraction and adaptive thresholding in subsequent work. Furthermore, although color-based features offer complementary information, their sensitivity to variations in lighting conditions could restrict reliability unless sophisticated preprocessing or calibration methods are employed.

This method provides potential value to agricultural professionals and seed processing industries for automated, scalable, and cost-effective seed quality screening platforms. However, potential future enhancement involves incorporating machine learning algorithms, increasing the dataset, and integrating comparative benchmarking with state-of-the-art techniques to improve model generalizability and performance.

In short, the method presented here establishes a solid foundation for automated seed classification but also highlights the need for ongoing improvement, particularly in error analysis, adaptive classification methods, and verification against current literature and commercial systems.

Author Contributions

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare no conflict of interest.

References

Ahmed, M. R., Yasmin, J., Park, E., Kim, G., Kim, M. S., Wakholi, C., Mo, C., & Cho, B. K. (2020). Classification of watermelon seeds using morphological patterns of X-ray imaging: A comparison of conventional machine learning and deep learning. *Sensors*, *20*(23), 6753. https://doi.org/10.3390/s20236753

Beck, M. A., Liu, C. Y., Bidinosti, C. P., Henry, C. J., Godee, C. M., & Ajmani, M. (2020). An embedded system for the automated generation of labeled plant images to enable machine learning applications in agriculture. *PLOS One*, *15*(12), e0243923. https://doi.org/10.1371/journal.pone.0243923

Chen, T., Lv, L., Wang, D., Zhang, J., Yang, Y., Zhao, Z., & et al. (2024). Empowering agrifood system with artificial intelligence: A survey of the progress, challenges and opportunities. *ACM Comput. Surv.*, *57*(2), 1-37. https://doi.org/10.1145/3698589.

Darbyshire, M., Salazar-Gomez, A., Gao, J., Sklar, E. I., & Parsons, S. (2023). Towards practical object detection for weed spraying in precision agriculture. *Front. Plant Sci.*, *14*, 1183277. https://doi.org/10.3389/fpls.2023.1183277.

Dericquebourg, E., Hafiane, A., & Canals, R. (2022). Generative-model-based data labeling for deep network regression: Application to seed maturity estimation from UAV multispectral images. *Remote Sens.*, *14*(20), 5238. https://doi.org/10.3390/rs14205238.

Du, X., Si, L. Q., Li, P. F., & Yun, Z. H. (2023). A method for detecting the quality of cotton seeds based on an improved ResNet50 model. *PLOS One*, *18*(2), e0273057. https://doi.org/10.1371/journal.pone.0273057.

ElMasry, G., Mandour, N., Al-Rejaie, S., Belin, E., & Rousseau, D. (2019). Recent applications of multispectral imaging in seed phenotyping and quality monitoring—An overview. *Sensors*, *19*(5), 1090. https://doi.org/10.3390/s19051090.

Ghimire, A., Kim, S.-H., Cho, A., Jang, N., Ahn, S., Islam, M. S., Mansoor, S., Chung, Y. S., & Kim, Y. (2023). Automatic evaluation of soybean seed traits using RGB image data and a Python algorithm. *Plants, 12*(17), 3078. https://doi.org/10.3390/plants12173078.

Haque, F. & Haque, S. (2018). Plant recognition system using leaf shape features and minimum Euclidean distance.

ICTACT J. Image Video Process., 9(2), 1919-1925. https://doi.org/10.21917/ijivp.2018.0272.

Jin, B., Qi, H., Jia, L., Tang, Q., Gao, L., Li, Z., & Zhao, G. (2022). Determination of viability and vigor of naturally-aged rice seeds using hyperspectral imaging with machine learning. *Infrared Phys. Technol.*, *122*, 104097. https://doi.org/10.1016/j.infrared.2022.104097.

Kulkarni, P., Karwande, A., Kolhe, T., Kamble, S., Joshi, A., & Wyawahare, M. (2021). *Plant disease detection using image processing and machine learning*. https://doi.org/10.48550/arXiv.2106.10698.

Margapuri, V. & Neilsen, M. (2021). Classification of seeds using domain randomization on self-supervised learning frameworks. In *2021 IEEE Symposium Series on Computational Intelligence (SSCI), Orlando, FL, USA*. https://doi.org/10.1109/SSCI50451.2021.9659998.

Martín-Gómez, J. J., Rodríguez-Lorenzo, J. L., Gutiérrez del Pozo, D., Cabello Sáez de Santamaría, F., Muñoz-Organero, G., Tocino, Á., & Cervantes, E. (2024). Seed morphological analysis in species of Vitis and relatives. *Horticulturae, 10*(3), 285. https://doi.org/10.3390/horticulturae10030285.

Medeiros, A. D. D., Silva, L. J. D., Ribeiro, J. P. O., Ferreira, K. C., Rosas, J. T. F., Santos, A. A., & Silva, C. B. D. (2020). Machine learning for seed quality classification: An advanced approach using merger data from FT-NIR spectroscopy and X-ray imaging. *Sensors*, *20*(15), 4319. https://doi.org/10.3390/s20154319.

Nkemelu, D. K., Omeiza, D., & Lubalo, N. (2018). *Deep convolutional neural network for plant seedlings classification*. https://doi.org/10.48550/arXiv.1811.08404.

Opara, I. K., Opara, U. L., Okolie, J. A., & Fawole, O. A. (2024). Machine learning application in horticulture and prospects for predicting fresh produce losses and waste: A review. *Plants*, *13*(9), 1200. https://doi.org/10.3390/plants13091200.

Pande, A., Munot, M., Sreeemathy, R., & Bakare, R. V. (2019). An efficient approach to fruit classification and grading using deep convolutional neural network. In *2019 IEEE 5th International Conference for Convergence in Technology (I2CT), Bombay, India* (pp. 1-7). https://doi.org/10.1109/I2CT45611.2019.9033957.

Ravichandran, P., Viswanathan, S., Ravichandran, S., Pan, Y. J., & Chang, Y. K. (2022). Estimation of grain quality parameters in rice for high-throughput screening with near-infrared spectroscopy and deep learning. *Cereal Chem.*, *99*(4), 907-919. https://doi.org/10.1002/cche.10546.

Saeed, A., Tariq, M., Ibrahim, M., Ahmad, N., Ahmad, A. M., Aftab, R. S., & Mehdi, S. M. (2015). Identification of canola seeds using nearest neighbor and K-nearest neighbor algorithms. *Aust. J. Bus. Sci. Des. Lit.*, *7*(1), 36-43.

Santos, L., Santos, F. N., Oliveira, P. M., & Shinde, P. (2020). Deep learning applications in agriculture: A short review. In *Robot 2019: Fourth Iberian Robotics Conference*. https://doi.org/10.1007/978-3-030-35990-4_12.

Toda, Y., Okura, F., Ito, J., Okada, S., Kinoshita, T., Tsuji, H., Saisho, D. (2020). Training instance segmentation neural network with synthetic datasets for crop seed phenotyping. *Commun. Biol., 3*, 173. https://doi.org/10.1038/s42003-020-0905-5.