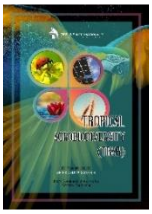
ZIBELINE INTERNATIONAL™
PUBLISHING
ISSN: 2716-7046 (Online)
CODEN: TARGCA

Tropical Agrobiodiversity (TRAB)

DOI: http://doi.org/10.26480/trab.01.2025.38.44



RESEARCH ARTICLE

PERFORMANCE EVALUATION OF CNN-BASED SOYBEAN SEED HEALTH RECOGNITION: A COMPARATIVE STUDY OF DIFFERENT DATASETS

Waqar Ahmad[a], Aftab Khaliq[b*], Ibrar Ahmad[c], Muhammad Shoaib[b], Hafiz Sultan Mahmood[b], Fiaz Ahmad[d], Bushra Siddique[c], Hajra Azeem[e], Mahmood Ali[b]

[a] Department of Bioinformatics and Biotechnology, Government College University Faisalabad
[b] Agricultural Engineering Institute, National Agricultural Research Centre (NARC), PARC, Islamabad 44000, Pakistan
[c] College of Biosystems Engineering and Food Science, Zhejiang University, Hangzhou 310058, China
[d] Department of Agricultural Engineering, Bahauddin Zakariya University, Multan 60800, Pakistan
[e] Crop Diseases Research Institute, National Agricultural Research Centre (NARC), PARC, Islamabad 44000, Pakistan
**Corresponding authors: aftabkhaliq123@gmail.com*

ARTICLE DETAILS

Article History:

Received 7 June 2025
Revised 14 July 2025
Accepted 20 July 2025
Available online 1 August 2025

ABSTRACT

Assessment of soybean seed quality is crucial when evaluating its potential for both nutritional consumption and propagation in subsequent agricultural cycles. This investigation presents an efficient, custom-designed Convolutional Neural Network (CNN) model using deep learning for assessing soybean seedling quality through non-destructive means. The proposed system is intended to function as an enhanced alternative to current human visual inspection methodologies. The performance evaluation and comparative analysis of the CNN model employed five datasets, comprising healthy soybeans and four distinct categories of diseased soybeans. The experimental results from the dataset demonstrated a classification accuracy of 98% achieved by the proposed model, as evidenced by comparisons with previous studies. The model exhibited a balanced performance between precision and recall across all categories (healthy, broken, spotted, skin-damaged, and immature), with an average F1 score of 0.95%. However, datasets D4 (Healthy and immature soybean) and D5 (healthy and mixed unhealthy soybean) exhibited a notable decline in test accuracy to 85%, accompanied by an F1-score of 82%, indicating a requirement for additional training data. Notwithstanding this variation, the model exceptional performance on most datasets suggests its potential applicability for real-time soybean quality assessment applications.

KEYWORDS

Convolutional Neural Network (CNN) model, Soybean quality, Non-destructive, deep learning

1. INTRODUCTION

The soybean (Glycine max), a species belonging to the legume family, serves as a fundamental component in the food production industry. This multifaceted crop is highly valued for its remarkable protein content, establishing an essential constituent in both human nutrition and animal feed formulations (Sable et al., 2024). Soybean seed quality is a critical determinant of both industrial demand and actual crop yield (Radchuk and Borisjuk, 2014). The evaluation of soybean quality, encompassing healthy, unhealthy, and damaged specimens, has traditionally relied on human visual inspection and mechanical sorting systems to identify and separate substandard and defective kernels (Koklu et al., 2023). The primary challenge in this procedure is the subjective nature of expert evaluation, which necessitates substantial time and effort due to lengthy, repetitive work hours. As soybean seed damage is predominantly visible on the surface, there is an increasing demand for automated systems capable of accurately and objectively identifying and categorizing soybean seed defects (Lin et al., 2019). The field of computer vision is undergoing rapid advancement, offering a versatile and transformative approach to the identification of healthy, unhealthy, and damaged soybean kernels.

Image-based processing models, developed utilizing computer vision and image processing technologies, have significantly contributed to the advancement of smart agriculture by enhancing the identification of agricultural products. These models have exerted a substantial positive influence on the field of precision farming (Lin et al., 2023). In recent years, substantial research has been conducted on non-invasive methods for evaluating agricultural products. This has been accomplished through the implementation of various technologies, including computer vision, image processing, and hyperspectral imaging. These investigations aim to elucidate the critical relationship between agriculture and Artificial Intelligence.

The objective of this research is to develop a transfer learning, custom-build model to compare five classes of soybean seed datasets obtained from an online data source. A newly developed lightweight CNN model utilizing deep learning techniques is employed for the classification of the dataset. This model is capable of automatically classifying the different aspects of seed quality, aiming to mitigate the problem of manual sorting. The segmentation of soybean seeds into four unhealthy classes broken, spotted, skin-damaged, and immature was proposed based on physical appearance, with classification performed using a CNN-based network. The dataset was divided into training, validation, and test subsets with 256*256*3 image dimensions to enhance data visualization. The availability of this lightweight model enhances the accuracy of the deep learning model for binary classification. The development and implementation of such a classification model in automated sorting mechanisms will supplant the labor-intensive and time-consuming task of assessing the quality of individual soybean grains, as well as facilitate the provision of high-quality food that safeguards public health. This approach would also contribute to the reduction of food waste while concurrently producing high-quality products.

1.1 RELATED WORKS

In agriculture, Deep Learning techniques are utilized for predicting yields, identifying plant diseases, and classifying seeds. Meanwhile, researchers

Quick Response Code	Access this article online	
	Website: www.trab.org.my	**DOI:** 10.26480/trab.01.2025.38.44

Cite the Article: Waqar Ahmad, Aftab Khaliq, Ibrar Ahmad, Muhammad Shoaib, Hafiz Sultan Mahmood, Fiaz Ahmad, Bushra Siddique, Hajra Azeem, Mahmood Ali (2025). Performance Evaluation of CNN-Based Soybean Seed Health Recognition: A Comparative Study of Different Datasets. *Tropical Agrobiodiversity, 6(1): 38-44.*

are exploring various technologies for non-invasive evaluation of food products. These include image processing methods, computer vision systems, and hyperspectral imaging techniques, which are essential for rapidly assessing product quality.

This study, successfully predicted soybean seed quality utilizing a deep learning-based sorting technique (Zhao et al., 2021). To achieve enhanced performance, a prevalent approach involves developing a robust and intricate network. However, significant constraints and computational expenses impede the transition of these methodologies to mobile devices. In a study conducted a streamlined end-to-end network was utilized for the detection of coconut diseases (Singh et al., 2021). Their custom-designed CNN model exhibited superior accuracy in comparison to conventional CNN architectures. Utilizing mask R-CNN, conducted segmentation of individual soybean seeds (Huang et al., 2022). Subsequently, they implemented ResNet for classification, achieving an accuracy rate of 96.20%. According to this study, conducted a study on lotus seed sorting using the YOLOV3 model, achieving an accuracy rate of 95% with a detection time of 45 ms (Lu et al., 2022). In a study, a modified visual geometry group network called Rice-VGG-16 was developed to classify defective rice seeds (Sun et al., 2023). This network replaced the 5th max pooling layer of the standard VGG-16 with an average pooling layer and employed Leaky-ReLU as the activation function. The enhanced Rice-VGG-16 model achieved a 99.5% accuracy rate in identifying high-quality rice. This study employed quadratic support vector to classify wheat seeds, achieving an average accuracy of 97% (Zargham et al., 2022). In contrast, utilized multiclass retinex with color restoration (MSRCR) to enhance soybean images, followed by Otsu's method for segmentation, resulting in a 98% accuracy rate (Lin et al., 2022). This study introduced an advanced technique for image acquisition, data processing, and analysis to examine soybean color and morphology, attaining a successful recognition rate of 97.36% (Baek et al., 2020). This method demonstrates significant potential for inspecting various samples that are challenging for human evaluation.

The research conducted, investigates a methodology for precise measurement of crop areas utilizing off-centered Bayesian deep learning (Wu et al., 2023). Through the incorporation of phenological features and an attention mechanism, their approach attains a high accuracy of 90.7% in the classification of maize, rice, and soybean. This methodology is both accessible and adaptable, facilitating rapid implementation across diverse regions and providing reliable agricultural yield estimates to address food security concerns. In a separate investigation, employed a modified ShuffleNet for soybean kernel identification, utilizing a combination of HIS and RGB images (Zheng et al., 2023). Their High-Resolution Feature Network (HRFN) enhances spatial fidelity in high-resolution images. Optimal recognition performance, with an ACCp of 98.36%, is achieved using ShuffleNet_COCSP and OISEW. This precise methodology demonstrates potential for assessing crop kernel quality indicators.

2. MATERIAL AND METHODS

2.1 Dataset Acquisition

This investigation utilized a dataset acquired from a public repository, comprising high-quality images of soybean seeds categorized into five classes: healthy, spotted, immature, broken, and surface-impaired. Figure 1.1 presented images of various soybean seed classes examined in this study. The comprehensive dataset encompassed 5,513 images (Lin et al., 2023). These images were classified in accordance with the Standard of Soybean Classification (GB1352-2009). To investigate various soybean seed health states, five distinct datasets (D1-D5) were established and figure 1.2 represented the number images belongs to specific dataset. Table 1 defines the sequential stages of data processing, explaining the division methodologies utilized for the model training, validation, and testing phases.

- D1: Healthy and Unhealthy (Broken soybean)
- D2: Healthy and Unhealthy (Spotted soybean)
- D3: Healthy and Unhealthy (Skin-damaged soybean)
- D4: Healthy and Unhealthy (Immature soybean)
- D5: Healthy and Unhealthy (Unhealthy mixed)

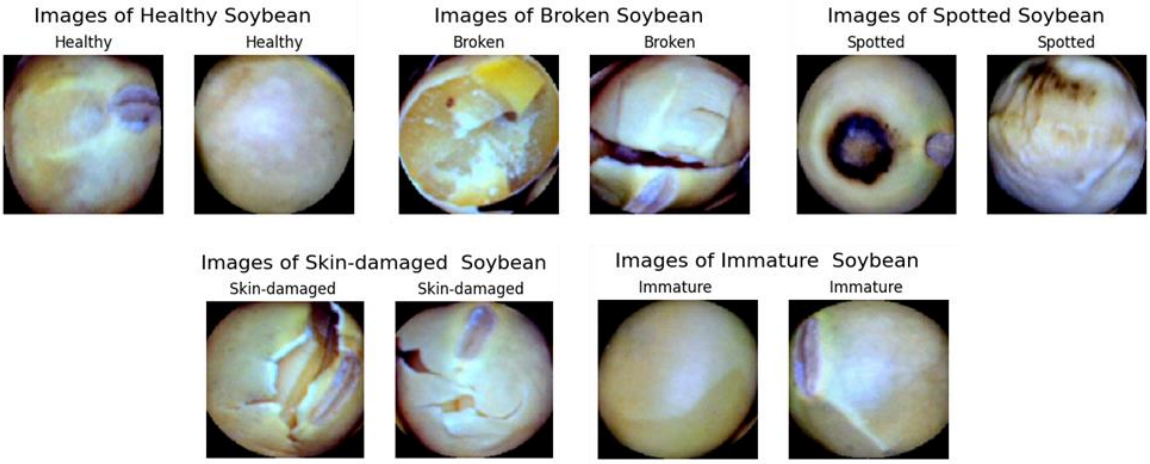


Figure 1: Randomly Selected samples of soybean seed to classify the various categories

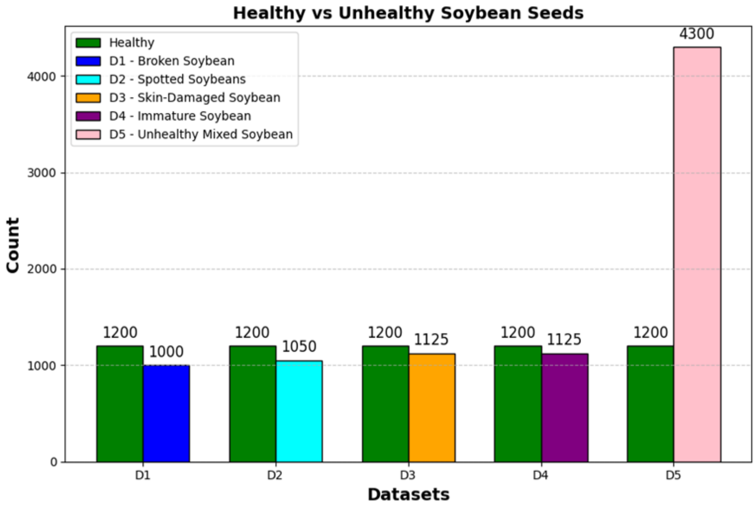


Figure 2: Images belong to different datasets

Cite the Article: Waqar Ahmad, Aftab Khaliq, Ibrar Ahmad, Muhammad Shoaib, Hafiz Sultan Mahmood, Fiaz Ahmad, Bushra Siddique, Hajra Azeem, Mahmood Ali (2025). Performance Evaluation of CNN-Based Soybean Seed Health Recognition: A Comparative Study of Different Datasets. *Tropical Agrobiodiversity, 6(1): 38-44.*

Table 1: Training, validation and test splits of datasets

Dataset	Splits	Healthy	Unhealthy
D1	Training	839	699
	Validation	240	200
	Test	121	101
D2	Training	839	734
	Validation	240	210
	Test	121	106
D3	Training	839	787
	Validation	240	225
	Test	121	113
D4	Training	839	787
	Validation	240	225
	Test	121	113
D5	Training	839	3009
	Validation	240	860
	Test	121	431

2.2 Preprocessing

Prior to model training, evaluation, and validation, the soybean dataset underwent a comprehensive preprocessing procedure. To enhance computational efficiency, all images were standardized to 256×256 pixels, irrespective of their original dimensions. To augment model generalization and mitigate overfitting, various data augmentation techniques, including random rotation and flipping, were implemented. These methods expanded the dataset by presenting diverse perspectives of soybean seeds for classification. To further address overfitting, dropout layers were incorporated during the training phase, randomly deactivating a proportion of neurons. In conjunction with these preprocessing steps, max-pooling layers were employed for spatial down sampling, optimizing the feature extraction process. This comprehensive approach incorporated transfer learning, fine-tuning, and the utilization of existing datasets to enable the model to effectively learn and identify complex features.

2.3 Model Architecture

CNN-based models are suitable for large datasets; thus, a simple architecture is proposed. However, their computational complexity limits their application on edge devices. The primary objective of this study is to perform soybean classification. The architecture of the proposed model is illustrated in the figure 3. As depicted, the architecture is primarily divided into two main stages: (a) a feature extraction network, and (b) classification.

In the initial stage of the architecture, five layers of each convolutional, batch normalization, and max pooling layers, with one global average pooling layer are utilized. Each convolutional layer employs a 3 x 3 filter size, corresponding to the feature extraction network. For feature extraction, the convolutional layers are designed such that the depth increases with each subsequent layer. The first convolutional layer extracts macro details, while the final convolutional layer extracts micro features. This design approach not only reduces the number of parameters but also transfers unnecessary features to subsequent downstream layers while maintaining optimal performance. The first two convolutional layers have depths of 16 and 32 respectively, whereas the subsequent two have depths of 64 and 128, with the final layer having a depth of 256. A rectified linear unit is employed as the activation function for each layer. To maintain the output size of the convolutional layer, the stride value is set to 1, and padding is configured to match the input size. Batch normalization is implemented to stabilize the network, which can accelerate the training process and allow for the use of higher learning rates. As illustrated in the figure, a batch normalization layer is applied after each convolutional layer. Following the first five convolutional layer groups, down-sampling is performed through max pooling and average pooling in the network. Max pooling utilizes a stride value of 2 and a filter size of 2 x 2 in the network.

In the second stage of the architecture, a flatten operation is applied subsequent to the convolution and batch normalization layers to reshape the output of the first stage into a single-dimensional feature vector for input into a fully connected layer. The flatten operation is executed by the flatten layer from the TensorFlow and Keras libraries. This layer transforms a three-dimensional matrix (6 × 6 × 256) inherited from the preceding max pooling layer into a one-dimensional vector (1 × 9216). To mitigate overfitting, a fully connected layer comprising 16 hidden units are employed, utilizing ReLU as the activation function. Subsequently, a dropout layer is implemented to randomly omit hidden units. The dropout layer serves primarily to prevent overfitting and functions as a regularization method, approximating the concurrent training of multiple nodes in the fully connected layer. During the training process, certain neurons in the fully connected layer are randomly ignored or dropped. For each epoch, the layer update during training is executed from a different perspective. Consequently, the dropout layer is advantageous in disrupting scenarios where network tiers co-adapt to rectify errors committed by prior layers, thereby enhancing model robustness. Given that the hidden fully connected layer consists of 16 units, while the output fully connected layer comprises five units, the dropout layer is positioned between these two layers. An excessively high dropout rate would be too severe for a network with a low number of units, while an extremely low dropout rate would not provide the full benefits of the dropout layer in mitigating overfitting. Therefore, in the proposed model, the dropout rate is set to 0.2, indicating that 20% of the inputs will be randomly omitted in each epoch. Following the dropout layer, a final fully connected layer with five units (classes) is implemented, employing the Softmax function for classification.

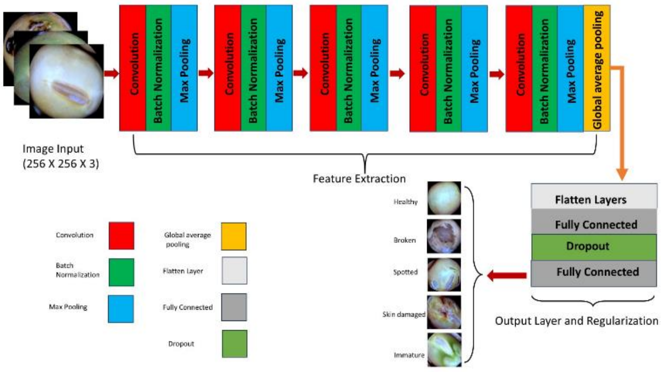


Figure 3: Architecture of the proposed CNN based soybean classification model

3. EXPERIMENT RESULTS

3.1 Model analysis considering datasets

Figure 4 showcases a comprehensive assessment of our proposed model performance through training and validation curves, displaying accuracy and loss across five datasets (D1, D2, D3, D4, and D5) over 100 epochs. These graph curves offer valuable insights into our model learning dynamics throughout the training process. The analysis reveals that training accuracy begins at 0.78 and swiftly increases in the initial epochs for all datasets, reaching stability near 1.0 after roughly 10 epochs. Following the 10th epoch, the training accuracy fluctuates between 0.95 and 1.00 until the conclusion of the training process for all datasets, indicating the model effective assimilation of the training data. The top-right plot illustrates the validation accuracy, which exhibits a comparable upward trend, rapidly improving in the early epochs and stabilizing between 0.90 and 1.10 for most datasets. Nevertheless, significant fluctuations are observed, particularly in D4 and D5, suggesting potential variations in the model applicability to these datasets.

The analysis of training and validation loss patterns revealed an inverse relationship with accuracy. The graph in the lower left quadrant illustrates the training loss trends, which exhibit a rapid decrease during the initial epochs and subsequently stabilize near zero by the conclusion of the tuning process. This observation suggests the model efficacy in error reduction throughout the training phase. Conversely, the graph in the lower right quadrant displays the validation loss, which demonstrates a sharp decline in the early epochs but subsequently exhibits fluctuations, particularly for D4 and D5. These fluctuations may indicate either overfitting or inconsistent model performance on this specific dataset.

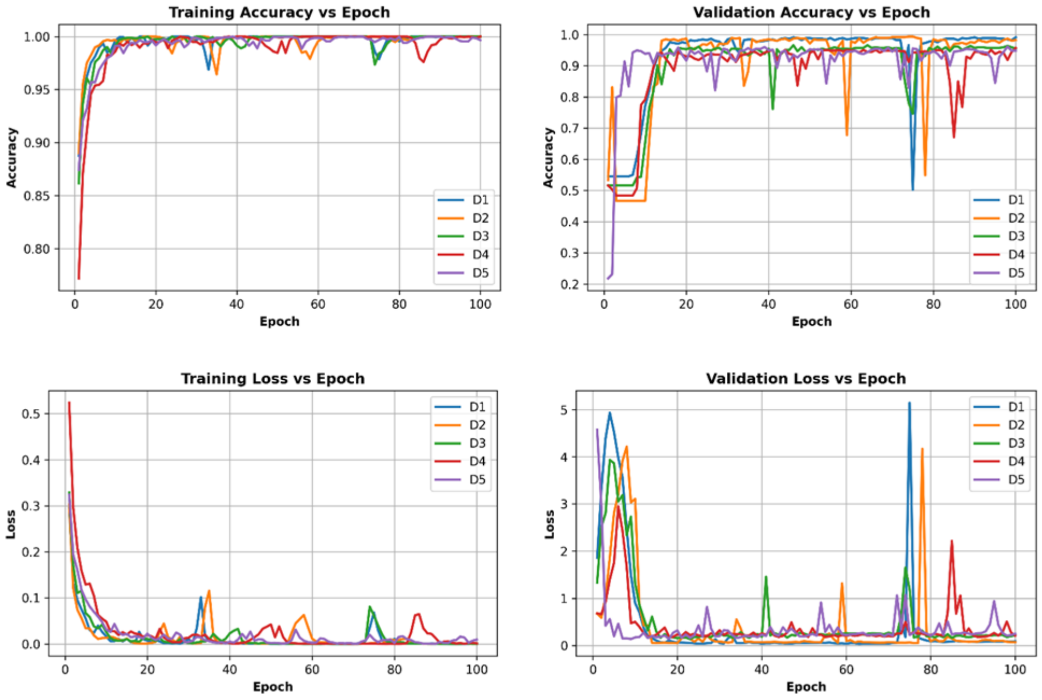


Figure 4: Graphical representation of epoch-wise accuracy and loss curves of proposed model

3.2 Evaluation metrics of different datasets

The proposed model effectiveness across various datasets is assessed utilizing a confusion matrix, which comprises four key components: True Positive (TP), False Positive (FP), True Negative (TN), and False Negative (FN). This matrix is commonly employed to calculate performance metrics such as accuracy, precision, recall, and F1-score for each dataset. The model overall performance is reflected in the dataset's accuracy. Precision indicates the model ability to correctly identify true positive samples among all positive predictions, while recall measures the model accuracy in identifying positive samples. The F1-score represents the harmonic mean of precision and recall. Additionally, the average testing time for image prediction is considered as a performance indicator.

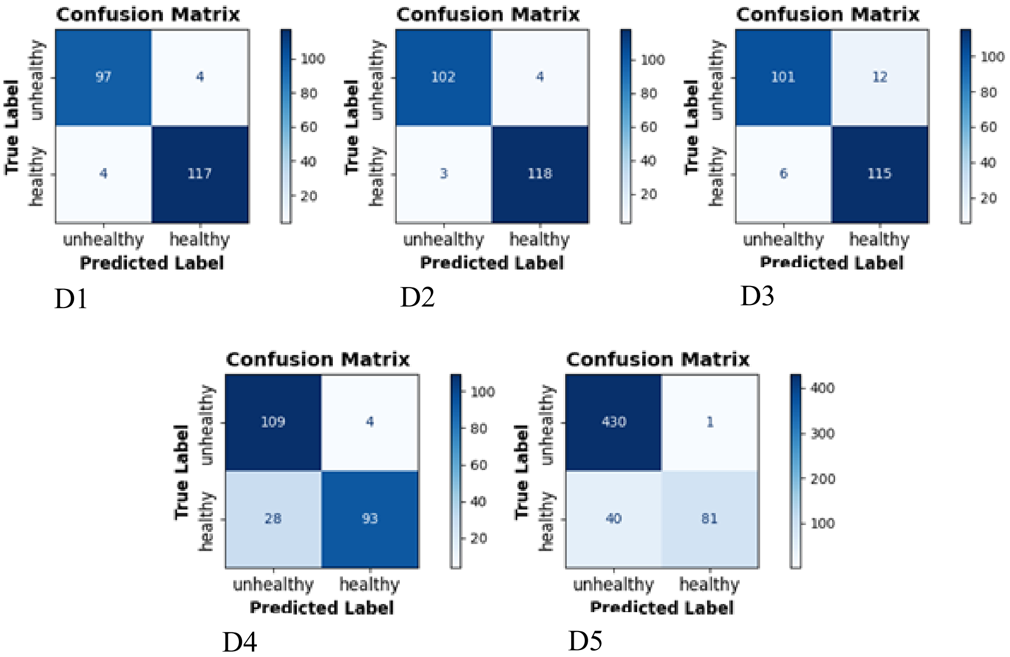
Table 2 presents the evaluation metrics for the studied datasets. The results of demonstrate that macro-precision ranges from 91% to 98%, macro-recall from 66% to 98%, macro-F1 score from 82% to 98%, and test accuracy from 85% to 98%, all derived from the confusion matrix. During the testing phase, dataset D2 achieved the highest performance matrix values of 98% with the lowest average testing time of 1418.35, while D1 scored 96% with a test accuracy of 98%. Conversely, D5 exhibited the lowest test accuracy of 85%, with an F1-score of 82%, recall of 66%, and the highest testing time among the datasets evaluated during the testing phase.

Tables 2: Evaluation indexes for the different dataset

Dataset	Macro-Precision%	Macro-Recall%	Macro-F1%	Test Accuracy%	Time (s)
D_1	96	96	96	98	1466.10
D_2	98	98	98	98	1418.35
D_3	92	92	92	92	1517.00
D_4	88	87	86	86	1512.18
D_5	91	66	82	85	3600.03

The confusion matrix illustrated in Figure 5 represents the outcomes of our proposed model testing phase. The results indicate that while the majority of the dataset was correctly classified, a small number of instances were misidentified. For dataset D1, the model demonstrated exceptional performance with a 98% accuracy rate, misclassifying only 4 out of 121 healthy seeds as false negatives and 4 out of 101 unhealthy (broken) seeds as false positives. Similarly, dataset D2 exhibited high efficiency with a 98% accuracy, precision, and recall rate, misclassifying just 3 healthy seeds as spotted (false negatives) and 4 spotted seeds as healthy (false positives). In dataset D3, the model performed well but with a slightly lower accuracy of 92% and increased testing time, misclassifying 6 healthy seeds as unhealthy (skin damaged) and 12 unhealthy seeds as healthy. Dataset D4 revealed a decline in model performance with an accuracy of 86%, showing 28 false negatives (healthy seeds misclassified as immature) and 4 false positives (immature seeds misclassified as healthy). The model performance was least effective for dataset D5, with 81 false negatives in the healthy class misclassified as unhealthy mixed. Despite the custom-built modified model high precision and accuracy in the first three datasets, the observed misclassifications highlight areas for potential improvement and optimization. These insights from the confusion matrix provide valuable information for enhancing the model performance and accuracy in soybean seed classification.

Cite the Article: Waqar Ahmad, Aftab Khaliq, Ibrar Ahmad, Muhammad Shoaib, Hafiz Sultan Mahmood, Fiaz Ahmad, Bushra Siddique, Hajra Azeem, Mahmood Ali (2025). Performance Evaluation of CNN-Based Soybean Seed Health Recognition: A Comparative Study of Different Datasets. *Tropical Agrobiodiversity, 6(1): 38-44.*



D1 D2 D3

D4 D5

Figure 5: Confusion matrices for different datasets

3.3 Class wise evaluation of datasets

Table 3 presents the precision, recall, and F1-score metrics, along with the overall accuracy of 98%, provide a comprehensive evaluation of our model performance in soybean seed classification across datasets. Datasets D1 and D2 show consistently high precision and recall of 97% for both healthy and unhealthy classes, resulting in robust F1-scores of 97%. For D3, the model maintains strong performance by achieves a precision of 91% and a recall of 95% of healthy class, resulting in an F1-score of 93%. Similarly, the unhealthy class displays a precision of 94% and a recall of 89%, leading to an F1-score of 92%, reflecting its ability to handle moderate variability. In D4, a difference arises, the healthy class achieves high precision of 96% but lower recall of 77%, while the unhealthy class favors recall of 96% over precision of 80%, yielding F1-scores of 85% and 87%, respectively. Dataset D5 presents the most challenging scenario for the model. The healthy class shows a marked decline in recall, achieving only 33%, despite a precision of 91%. This results in a low F1-score of 49%, suggesting difficulty in correctly classified the healthy instances. On the other hand, the unhealthy class attains a recall of 100%, indicating that the model captures all unhealthy instances, although with a precision of 84%. The resulting F1-score of 82% highlights a trade-off where the model prioritizes recall at the disbursement of precision for the unhealthy class.

D4 reveals a performance discrepancy, with the healthy class achieving high precision 96% but lower recall 77%, while the unhealthy class favors recall 96% over precision 80%. This results in F1-scores of 85% and 87% for healthy and unhealthy classes, respectively. D5 proves to be the most challenging dataset for the model. The healthy class experiences a significant decrease in recall 33% despite maintaining high precision 91%, resulting in a low F1-score of 49%. This suggests difficulties in correctly identifying healthy instances. Conversely, the unhealthy class in D5 achieves perfect recall 100% but lower precision 84%, yielding an F1-score of 82%. This indicates a trade-off where the model prioritizes capturing all unhealthy instances at the expense of precision.

Table 3: Class wise evaluation of studied datasets

Dataset	Class	Precision %	Recall %	F1-Score %
D1	Healthy	97	97	97
	Unhealthy	97	98	97
D2	Healthy	97	97	97
	Unhealthy	97	96	97
D3	Healthy	91	95	93
	Unhealthy	94	89	92
D4	Healthy	96	77	85
	Unhealthy	80	96	87
D5	Healthy	91	33	49
	Unhealthy	84	100	82

The accuracy distribution across all classes during testing is depicted in Figure 6. These measurements underscore the necessity of evaluating and fine-tuning models based on specific datasets. The consistently high precision, recall, and F1-scores in D1, D2, and D3 corroborate the model efficacy with these datasets. However, the variations observed in D4 and D5 indicate areas that necessitate improvement. Notably, the significant decrease in recall for the healthy class in D5 suggests a requirement for additional training data. These findings demonstrate the model robustness in handling straightforward datasets while revealing its limitations in more complex ones. This underscores the importance of dataset-specific adjustments and enhanced feature engineering to augment overall performance.

Cite the Article: Waqar Ahmad, Aftab Khaliq, Ibrar Ahmad, Muhammad Shoaib, Hafiz Sultan Mahmood, Fiaz Ahmad, Bushra Siddique, Hajra Azeem, Mahmood Ali (2025). Performance Evaluation of CNN-Based Soybean Seed Health Recognition: A Comparative Study of Different Datasets. *Tropical Agrobiodiversity, 6(1): 38-44.*

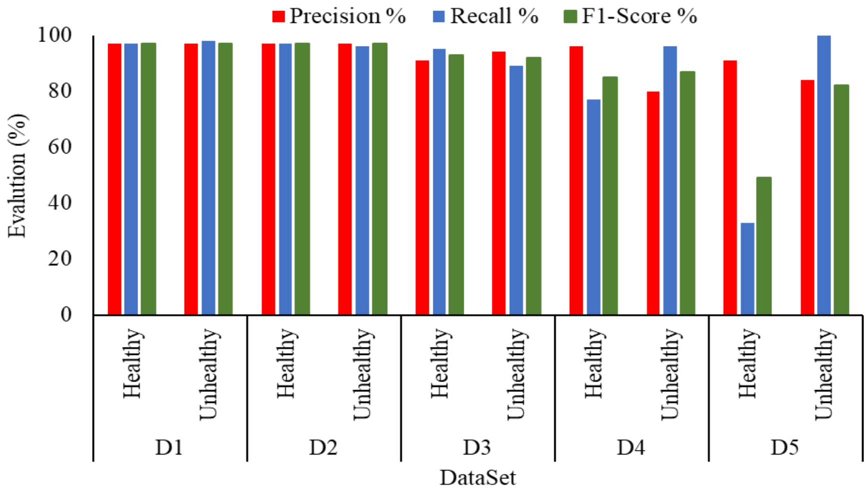


Figure 6: Class-wise accuracy of different datasets

3.4 Convolution layer visualization

Figure 7 illustrates the key output feature maps, providing insight into the learning process of our proposed model. The illustration demonstrates how the enhanced custom-built model captures feature information such as shapes, colors, textures, and edges at the initial levels. As the model progresses through deeper layers, the specific visual details become less distinct, while the information transitions to a more abstract representation.

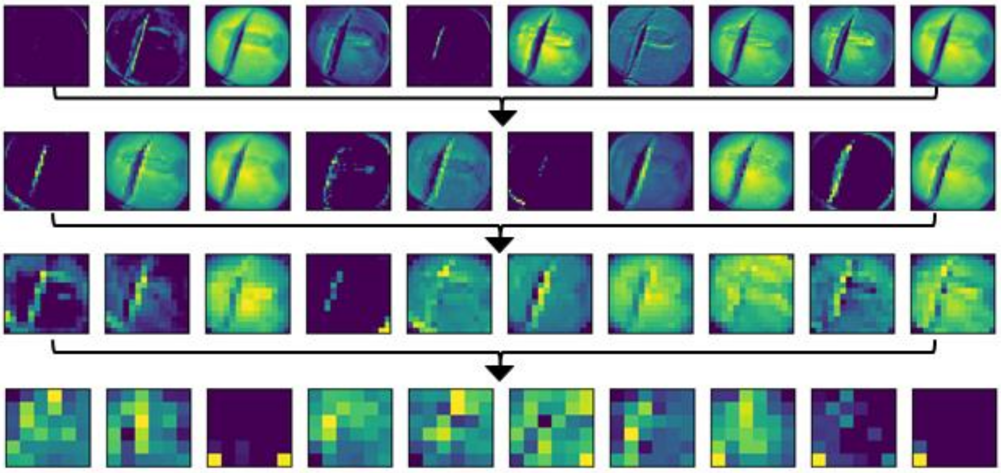


Figure 7: visualization results of output feature maps of CNN layers.

3.5 Prediction Accuracy of dataset

The CNN model proposed for comparing various soybean seed datasets demonstrated 98% accuracy with outstanding precision, recall, and F1-scores during data training and validation, exhibiting minimal variability. Figure 8 illustrates the model efficiency in distinguishing between healthy and unhealthy classes within the datasets. The model accurately identified healthy, broken, spotted, skin-damaged, and immature soybean seeds, thereby enhancing the industrial demand for high-quality products through its predictive capabilities.

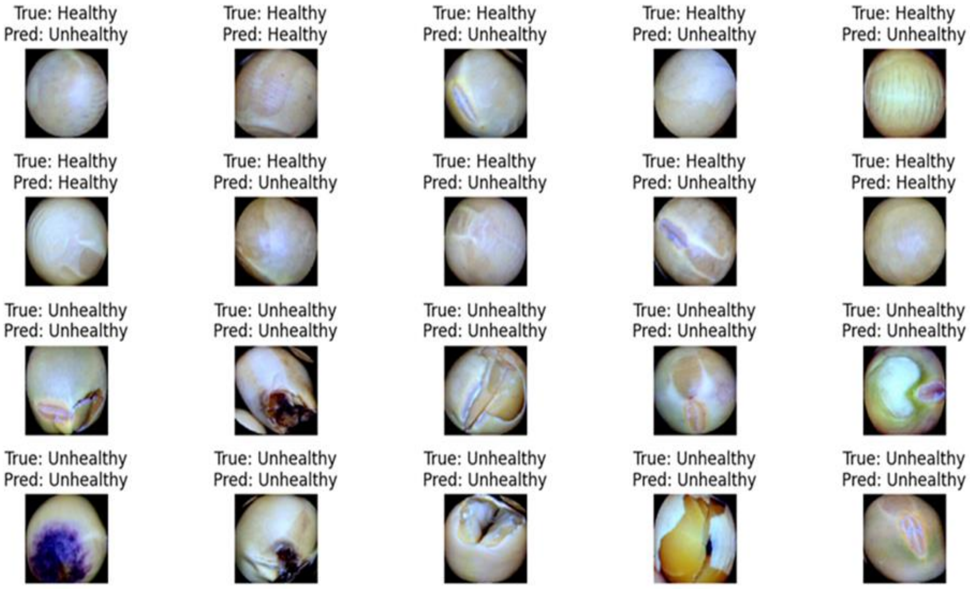


Figure 8: The results of trained model performance

Cite the Article: Waqar Ahmad, Aftab Khaliq, Ibrar Ahmad, Muhammad Shoaib, Hafiz Sultan Mahmood, Fiaz Ahmad, Bushra Siddique, Hajra Azeem, Mahmood Ali (2025). Performance Evaluation of CNN-Based Soybean Seed Health Recognition: A Comparative Study of Different Datasets. *Tropical Agrobiodiversity, 6(1): 38-44.*

4. CONCLUSIONS

This study compares five distinct soybean seed datasets to determine which data class yields the most efficient model performance. A custom-built model was adapted and trained utilizing Convolutional Neural Network (CNN) techniques for dataset classification. The proposed model demonstrated low data training and average testing times while delivering effective results with an accuracy of 98% across all datasets. Experimental findings revealed that the first three datasets achieved recall rates between 92% and 98%, contributing to a balanced F1-score. Furthermore, class-wise evaluation indicated that D1, D2, and D3 exhibited precision, recall, and F1-scores ranging from 90% to 97% for both healthy and unhealthy (broken, spotted, and skin-damaged) classes. However, D4 and D5 showed a significant decrease in recall, indicating a need for additional training data. Additionally, our comparative analysis with previous studies demonstrated that our model outperformed others in terms of accuracy, highlighting its efficacy in soybean seed classification tasks. Optimization of the model could enhance classification accuracy for D4 (healthy and immature) and D5 (healthy and mixed unhealthy) soybean seeds. The refined model possesses the potential for integration into an automated soybean seed sorting robotic system operating in real-time. Subsequent research should entail a comparative analysis of this model with novel classification techniques to maintain its position at the forefront of seed classification technology. These advancements may play a pivotal role in advancing seed classification technology and its implementation in agricultural and food production sectors.

REFERENCES

Baek, J., Lee, E., Kim, N., Kim, S. L., Choi, I., Ji, H., Chung, Y. S., Choi, M.-S., Moon, J.-K., and Kim, K.-H., 2020. High throughput phenotyping for various traits on soybean seeds using image analysis. Sensors, 20, 248. https://doi.org/10.3390/s20010248

Fazel-Niari, Z., Afkari-Sayyah, A. H., Abbaspour-Gilandeh, Y., Herrera-Miranda, I., Hernández-Hernández, J. L., and Hernández-Hernández, M., 2022. Quality assessment of components of wheat seed using different classification models. Applied Sciences, 12(8), 4133. https://doi.org/10.3390/app12084133

Huang, Z., Wang, R., Cao, Y., Zheng, S., Teng, Y., Wang, F., Wang, L., and Du, J., 2022. Deep learning based soybean seed classification. Computers and Electronics in Agriculture, 202, 107393. https://doi.org/10.1016/j.compag.2022.107393

Koklu, M., Kursun, R., Yasin, E. T., and Taspinar, Y. S., 2023. Detection of defects in soybean seeds by extracting deep features with SqueezeNet. In Proceedings of the IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS) (Vol. 1, pp. Pp. 713–717). https://doi.org/10.1109/IDAACS58523.2023.10348939

Lin, P., Li, X., Li, D., Jiang, S., Zou, Z., Lu, Q., and Chen, Y., 2019. Rapidly and exactly determining postharvest dry soybean seed quality based on machine vision technology. Scientific Reports, 9, 17143. https://doi.org/10.1038/s41598-019-53728-1

Lin, W., and Lin, Y., 2022. Soybean image segmentation based on multi-scale Retinex with color restoration. Journal of Physics: Conference Series, 2284, 012010. https://doi.org/10.1088/1742-6596/2284/1/012010

Lin, W., Fang, Y., Peng, X., Shu, L., Deng, M., Zhang, J., Zang, S., Huang, Y., Qiu, S., 2023. Soybean seeds [Data set]. Mendeley Data, V6. https://doi.org/10.17632/xyz123.6

Lin, W., Shu, L., Zhong, W., Lu, W., Ma, D., and Meng, Y., 2023. Online classification of soybean seeds based on deep learning. Engineering Applications of Artificial Intelligence, 123, 106434. https://doi.org/10.1016/j.engappai.2023.106434

Lu, A., Guo, R., Ma, Q., Ma, L., Cao, Y., ad Liu, J., 2022. Online sorting of drilled lotus seeds using deep learning. Biosystems Engineering, 221, Pp. 118–137. https://doi.org/10.1016/j.biosystemseng.2022.06.009

Radchuk, V., and Borisjuk, L., 2014. Physical, metabolic and developmental functions of the seed coat. Frontiers in Plant Science, 5, 510. https://doi.org/10.3389/fpls.2014.00510

Sable, A., Singh, P., Kaur, A., Driss, M., and Boulila, W., 2024. Quantifying soybean defects: A computational approach to seed classification using deep learning techniques. Agronomy, 14(6). https://doi.org/10.3390/agronomy14061098

Singh, P., Verma, A., and Alex, J. S. R., 2021. Disease and pest infection detection in coconut tree through deep learning techniques. Computers and Electronics in Agriculture, 182, 105986. https://doi.org/10.1016/j.compag.2021.105986

Sun, J., Zhang, Y., Zhu, X., and Zhang, Y.-D., 2023. Enhanced individual characteristics normalized lightweight Rice-VGG16 method for rice seed defect recognition. Multimedia Tools and Applications, 82, Pp. 3953–3972. https://doi.org/10.1007/s11042-022-12957-5

Wu, Y., Wu, P., Wu, Y., Yang, H., and Wang, B., 2023. Remote sensing crop recognition by coupling phenological features and off-center Bayesian deep learning. Remote Sensing, 15, 674. https://doi.org/10.3390/rs15030674

Zhao, G., Quan, L., Li, H., Feng, H., Li, S., Zhang, S., and Liu, R., 2021. Real-time recognition system of soybean seed full-surface defects based on deep learning. Computers and Electronics in Agriculture, 187, 106230. https://doi.org/10.1016/j.compag.2021.106230

Zheng, L., Zhao, M., Zhu, J., Huang, L., Zhao, J., Liang, D., and Zhang, D., 2023. Fusion of hyperspectral imaging (HSI) and RGB for identification of soybean kernel damages using ShuffleNet with convolutional optimization and cross stage partial architecture. Frontiers in Plant Science, 13, 1098864. https://doi.org/10.3389/fpls.2023.1098864

Cite the Article: Waqar Ahmad, Aftab Khaliq, Ibrar Ahmad, Muhammad Shoaib, Hafiz Sultan Mahmood, Fiaz Ahmad, Bushra Siddique, Hajra Azeem, Mahmood Ali (2025). Performance Evaluation of CNN-Based Soybean Seed Health Recognition: A Comparative Study of Different Datasets. *Tropical Agrobiodiversity, 6(1): 38-44.*