iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

AGRICULTURAL SEED IMAGE CLASSIFICATION USING CNN

Dr. A. Mary Sowjanya (Associate Professor)
Mrs. Varsha Kumari (Research Scholar)
Department of Computer Science and System Engineering,
Andhra University College of Engineering(A), Visakhapatnam, Andhra Pradesh.

Malireddi Sai Lahari (MCA Student)
Department of Information Technology and Computer Applications,
Andhra University College of Engineering(A), Visakhapatnam, Andhra Pradesh.

Corresponding Author: Malireddi Sai Lahari, Email: malireddilahari1201@gmail.com

ABSTRACT
The rapid growth of the agricultural supply chain has increased the need for fast and reliable identification of seed varieties, with manual visual inspection being one of the most significant bottlenecks in seed quality assessment. Different seed varieties can appear visually similar, and traditional inspection methods often struggle to provide consistent, scalable classification. This paper presents a Deep Learning-based Agricultural Seed Image Classification System that analyzes seed images and classifies them into one of nine varieties: Black Gram, Chickpea, Corn, Moong Bean, Pinto Bean, Rice, Soyabean, Wheat, and White Kidney Bean. The dataset of 900 labelled images is preprocessed through resizing, normalization, and data augmentation (rotation, shifting, shearing, zooming, and horizontal flipping), and a Convolutional Neural Network based on the AlexNet architecture is trained and evaluated. The experimental results show that the CNN model achieves a validation accuracy of 94.00%, with precision of 0.9636, recall of 0.9629, F1-Score of 0.9635, and AUC-ROC of 0.9812.

Keywords:
Seed Classification, Convolutional Neural Network, Deep Learning, AlexNet, Image Processing, Data Augmentation, Precision Agriculture.

INTRODUCTION
The increasing scale of the agricultural supply chain has led to a rapid rise in the need for automated seed-quality assessment. Among these challenges, distinguishing visually similar seed varieties is particularly difficult because manual inspection depends on subjective human judgement and does not scale well across large quantities of seeds. Traditional identification methods commonly depend on manual visual features such as color, shape, and size and therefore face difficulty when varieties share overlapping characteristics.
Deep Learning provides a data-driven approach for identifying discriminative visual patterns directly from images. In the proposed system, labelled seed images are processed and used to train a classification model that distinguishes nine different seed varieties. A Convolutional Neural Network based on the AlexNet architecture is selected because it is well suited to automatically learn hierarchical visual features such as shape, texture, and color patterns without manual feature engineering.
The literature survey in the project considers commonly used approaches for agricultural image classification, including traditional Machine Learning methods such as Support Vector Machine and K-Nearest Neighbors using hand-crafted features, as well as deep learning approaches using Convolutional Neural Networks.The project particularly focuses on a custom AlexNet-based CNN and evaluates its effectiveness for multi-class seed classification using the prepared seed image dataset.

OBJECTIVES
- To collect and organize a labelled agricultural seed image dataset covering nine seed varieties.
- To preprocess seed images through resizing, normalization, and data augmentation to improve model robustness.
- To design and train a Convolutional Neural Network based on the AlexNet architecture for seed classification.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

- To evaluate the model using Accuracy, Precision, Recall, F1-Score, AUC-ROC, classification reports, and a confusion matrix.

METHODOLOGY

Data Collection and Preprocessing

The dataset consists of 900 seed images organized into nine class folders: Black Gram, Chickpea, Corn, Moong Bean, Pinto Bean, Rice, Soyabean, Wheat, and White Kidney Bean. During preprocessing, all images are resized to 224 x 224 pixels and pixel values are rescaled to the range [0, 1]. The dataset is divided into 720 training images (80%) and 180 validation images (20%), with class distribution preserved across the split.

Fig. 1. Sample images from the nine seed classes used in the dataset.

Data Augmentation and Feature Learning

Real-time data augmentation is applied using Keras' ImageDataGenerator, including rotation (±20°), width and height shifts (0.2), shear (0.2), zoom (0.2), and horizontal flipping, to increase training diversity and reduce overfitting. Unlike traditional pipelines that rely on hand-crafted color, shape, and texture features, the CNN automatically learns discriminative feature maps directly from the training images.

CNN Model Architecture

The classification model is based on the AlexNet architecture. It begins with a 96-filter 11x11 convolution (stride 4) followed by max-pooling and batch normalization, a 256-filter 5x5 convolution block, three consecutive 3x3 convolution layers (384, 384, and 256 filters), and a final max-pooling/normalization stage. Extracted feature maps are flattened and passed through two fully connected layers of 4096 neurons each, with a dropout rate of 0.6 after every major block. A final Softmax layer with nine output neurons produces class probabilities for the predicted seed variety.

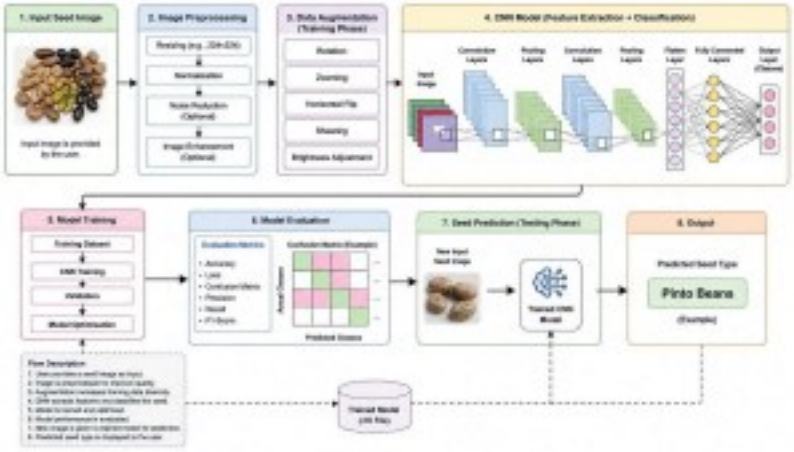

Fig. 2. System Architecture of the Proposed CNN-Based Seed Classification Pipeline.

IMPLEMENTATION

The implementation is performed in Python using TensorFlow and Keras for CNN design and training, OpenCV and NumPy for image loading and preprocessing, and Matplotlib/Seaborn for visualization. The model is trained

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

using the Adam optimizer (initial learning rate 0.0001) with categorical cross-entropy loss, a batch size of 64, and 10 training epochs.
A learning-rate scheduler reduces the learning rate by 10% after epoch 10, while ReduceLROnPlateau (factor 0.2, patience 5) and EarlyStopping (patience 29, restoring best weights) are used to stabilize training and control overfitting. Predictions and class probabilities are generated for the validation data, followed by calculation of evaluation metrics and visualization of the confusion matrix, training curves, and comparative performance charts.

RESULTS AND DISCUSSION

The proposed system was successfully implemented and evaluated using the 180-image validation split covering all nine seed classes. Table 1 summarizes the final evaluation metrics for the proposed CNN model.

Table 1. CNN Model Performance

Metric	Value
Validation Accuracy	94.00%
Validation Loss	0.1394
Precision	0.9636
Recall	0.9629
F1-Score	0.9635
AUC-ROC (OvR)	0.9812

The results indicate that the CNN model achieves consistently high performance across all nine seed classes, with precision, recall, and F1-Score values all above 0.96, confirming balanced and reliable classification rather than performance skewed toward a subset of classes.

Fig. 3. Training and Validation Accuracy Across Epochs.

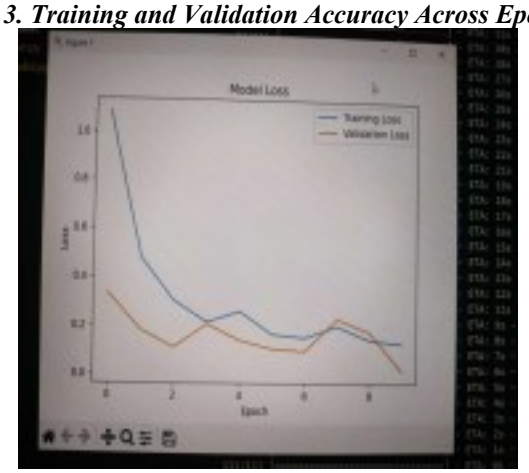

Fig. 4. Training and Validation Loss Across Epochs.

Both accuracy curves converge closely without a widening gap, and both loss curves decline steadily without divergence, indicating good generalization with minimal overfitting.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

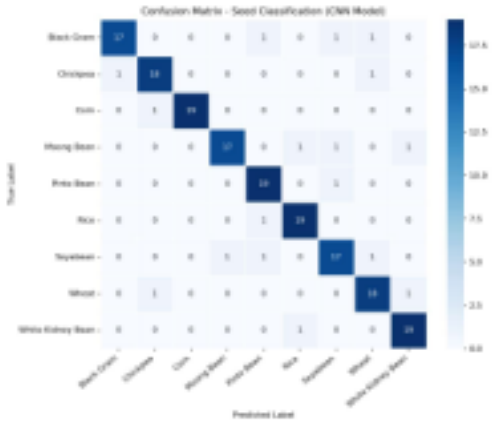

Fig. 5. Confusion Matrix on the Validation Set (9 Classes, 20 Samples Each).

The confusion matrix contains strong diagonal dominance across all nine classes, with only minor confusion between visually similar seed types such as Black Gram, Soyabean, and Wheat. This is consistent with the reported 94% validation accuracy and indicates that the CNN produces few misclassifications on the evaluated validation set.

Fig. 6. Sample Prediction Output -A Black Gram Image Classified with 99.85% Confidence.

CONCLUSION

A Deep Learning-based Agricultural Seed Image Classification System was developed to classify seed images into nine varieties: Black Gram, Chickpea, Corn, Moong Bean, Pinto Bean, Rice, Soyabean, Wheat, and White Kidney Bean. The workflow includes dataset preparation, preprocessing, data augmentation, CNN model training, testing, and performance evaluation. The proposed AlexNet-based CNN achieved strong results, with 94.00% validation accuracy, 0.9636 precision, 0.9629 recall, 0.9635 F1-Score, and 0.9812 AUC-ROC. The results demonstrate the effectiveness of the proposed CNN for automated seed classification.

FUTURE SCOPE

• Collect larger and more diverse seed image datasets to improve model generalization. • Explore advanced architectures such as ResNet, EfficientNet, and Vision Transformers for improved accuracy. • Apply transfer learning using pretrained models to improve performance with limited training data. • Integrate real-time seed classification into a mobile or web-based application for field use. • Extend the system to assess seed quality and detect seed-borne diseases in addition to variety classification. • Deploy the model in seed-certification and precision-farming environments for large-scale use.

REFERENCES

1) [1] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "ImageNet Classification with Deep Convolutional Neural Networks," Advances in Neural Information Processing Systems (NIPS), vol. 25, pp. 1097-1105, 2012.

 © Author(s) | CC BY 4.0 International License.

iJETRM
INTERNATIONAL JOURNAL OF ENGINEERING TECHNOLOGY RESEARCH & MANAGEMENT (IJETRM)
https://ijetrm.com/

2) [2] I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning, MIT Press, 2016.
3) [3] F. Chollet, Deep Learning with Python, 2nd ed., Manning Publications, 2021.
4) [4] R. C. Gonzalez and R. E. Woods, Digital Image Processing, 4th ed., Pearson Education, 2018.
5) [5] C. M. Bishop, Pattern Recognition and Machine Learning, Springer, 2006.
6) [6] TensorFlow Developers, TensorFlow Documentation, https://www.tensorflow.org
7) [7] Keras Team, Keras Documentation, https://keras.io
8) [8] OpenCV Team, OpenCV Documentation, https://opencv.org
9) [9] NumPy Documentation, Scientific Computing with Python, https://numpy.org
10) [10] Scikit-learn Developers, Scikit-learn Documentation, https://scikit-learn.org
11) [11] Pandas Documentation, Python Data Analysis Library, https://pandas.pydata.org
12) [12] Matplotlib Documentation, Visualization with Python, https://matplotlib.org
13) [13] Seaborn Documentation, Statistical Data Visualization, https://seaborn.pydata.org

 © Author(s) | CC BY 4.0 International License.