



UNLOCKING SEED POTENTIAL USING CNN

¹Dr S.Maruthu Perumal, ²Kunta Naveen, ³Kunuturi Sasi Kumar, ⁴Kutla Venkatesh Yadav, ⁵Maddukuri Poojitha

¹Assistant professor, School of Computing, Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Chennai, India- 600073.

^{2, 3, 4, 5} Student, School of Computing, Department of Computer Science and Engineering, Bharath Institute of Higher Education and Research, Chennai, India- 600073.

Abstract—Achieving a superior crop yield is crucial in the agricultural industry, with the quality of seeds being a key determinant. However, seeds are often sown without prior quality checks, leading to diminished yield and quality. This paper presents a novel method for automatically sorting seeds as either good or bad based on their visual characteristics, using a Convolutional Neural Network (CNN). The dataset used for training the model consisted of images of the top and bottom profiles of the seeds. The CNN achieved a classification accuracy of 96.875%. Additionally, a hardware solution was developed to classify seeds using the CNN model, which outperformed traditional methods by scanning both profiles of a seed. The hardware setup achieved a classification accuracy of 93.00%.

Index Terms: Seed Quality, Image Processing, Convolutional Neural Network

I. INTRODUCTION

Agriculture's foundation rests on seeds, making agriculture crucial for sustaining large populations. However, environmental factors like global warming and pollution have led to decreased crop yields, increasing the prevalence of low-quality seeds. Planting only high-quality seeds is essential for producing superior yields, which, in turn, yield high-quality seeds. Therefore, analyzing seed quality is crucial for ensuring healthy and productive crops.

Seed quality testing is vital for assessing a seed's potential output. Various methods, such as physical tests, evaluate morphological features like length, width, roundness, and diameter to assess seed quality. Vigor and germination tests determine seed viability by observing growth rates and deformities in germinating seeds.

This paper presents a machine vision-based method for analyzing and segregating soybean seeds based on morphological features. Typically, seed segregation is based on a single profile of the seed. However, this paper proposes a solution that segregates seeds based on images of both the top

bottom profiles, achieving increased efficiency and accuracy. A Convolutional Neural Network (CNN) is employed to classify seeds into two classes: good and bad. Additionally, a hardware setup, based on the CNN model, is used to automatically segregate soybean seeds based on their quality. Magnified images of the top and bottom profiles are captured using two microscopes, offering a more detailed view compared to non-magnified single profile images captured using a flatbed scanner.

II. LITERATURE SURVEY

Numerous studies have focused on evaluating seed quality and classifying seeds based on their grade and quality using various techniques. In [1], Image Processing techniques and Artificial Neural Networks were employed to classify rice grains into three grades (grade 1, grade 2, and grade 3) with a classification accuracy of approximately 98%.

Research has also explored the application of linear and non-linear machine learning algorithms on datasets containing seven morphological features (Area, Perimeter, Compactness, Length, Width, Asymmetry, and groove) [2]. Linear models such as Linear Regression and Linear Discriminant Analysis achieved accuracy scores of 92.86% and 95.8% respectively. Non-linear algorithms like K Neighbors Classifier achieved 90.4% accuracy, Decision Tree Classifier achieved 88%, Gaussian Naive Bayes' achieved 88.05%, and Support Vector Machine achieved 92.83% accuracy.

In [3], the focus was on various visual features of rice seeds. Different combinations of these features were used with machine learning techniques such as Support Vector Machine, Multi-Layer Perceptron, and Bayesian Networks to identify rice grains. The Multi-Layer Perceptron Model showed higher accuracy but similar precision compared to the Support Vector Machine Model. Each author will receive a complimentary copy of the journal showcasing their work, featuring the classification accuracy ranging from 90.27% to 90.54%. Online datasets have been taken and the techniques for the feature

extraction has been done in [4]. Three feature extractions GLCM, the Red Green Blue colour space technique and thresholding have been utilized and given to a bag-of-words model, which has been used to represent text data and this is given to the three classifiers- ANN, CNN and KNN to compare the results of using different types of neural networks. This study proves that these techniques are suitable for providing analysis on classification of seeds. Usage of neural network improves the accuracy and precision and CNN was found to be the best neural network to classify seeds

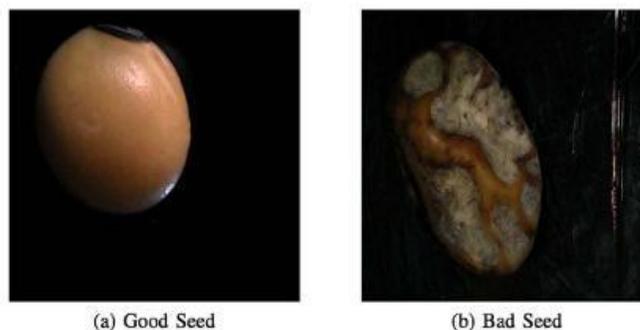


Fig. 2. Soya Seed images

II. PROPOSED METHODOLOGY

Generalised Block Diagram The general block diagram of this study is as shown in "Fig. 1". The dataset has been prepared, algorithm written and the CNN code optimized to classify the soya seeds into good or bad based on physical characteristics and the features extracted. The optimised Convolutional Neural Network model is then applied to a hardware prototype for the real-time classification of the seeds based on images of their upper and lower profiles.

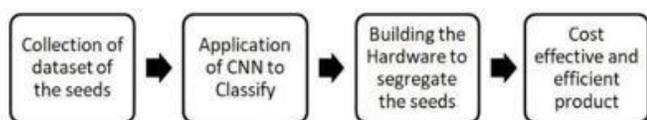


Fig. 1. The Block Diagram

B. Dataset

The soya seed image dataset comprised the following:

TABLE I
DATASET DESCRIPTION

Quality of seed	Training Images	Testing Images	Total
Good Seeds	160	50	210
Bad Seeds	160	50	210

Table I describes the number of training and testing images of the good and bad soya seeds that were used as the dataset in the project. The images in the dataset consisted of both top and bottom profiles of the seed placed in different positions on the hardware structure against a black background. "Fig. 2a" shows the image of a good soya seed. "Fig. 2b" shows the image of a bad soya seed. A total of 210 good seeds and 210 bad seeds were used to create the data set. As shown in Table I, 160 seeds of each type were used for training the model and 50 seeds each were used for testing. The images were clicked using a USB microscope with 200X zooming feature with image size 480 x 640 x 3. clicked against a black background. This way, clear, magnified images of the seeds were obtained.

C. The Convolutional Neural Network model

The Convolutional neural network shown in "Fig. 3" was used to train the model. A convolutional neural network was

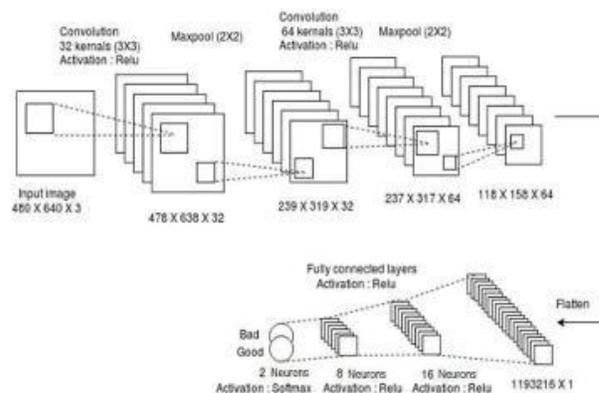


Fig. 3. The CNN model

utilized to prepare our demonstrate which was at that point utilized to classify the soya seeds into two categories- great and awful. The measurements for the input picture is 480 x 640 x 3. The input picture is passed through a convolution layer with ReLU Actuation comprising 32 channels of measure 3 x 3 each with a walk 1. This layer extricates the nearby highlights. The coming about measurement is 478 x 638 x 32. The yield of the convolutional layer was passed to a maxpool layer with pool measure 2 x 2 and a walk of 2 to decrease the number of highlights, subsequently holding as it were the imperative highlights. The coming about measurement is 239 x 319 x 32. The coming about yield of the maxpool layer was passed through a moment convolution layer with Relu Actuation comprising 64 channels of measure 3 x 3 each and walk 1 taken after by another maxpool layer with pool measure 2 x 2 and walk 2. This vector is at that point and through 3 completely associated layers of measure 16, 8 and 2 neurons individually. The completely associated layers have neurons where each neuron of a layer is associated to each neuron within the following layer, consequently the title. Consider the layer with 8 neurons taken after by the final layer with 2 neurons. The number of parameters interfacing these 2 layers are, (8 x 2 weights for each neuron + 2 inclinations for the final layer).

The final layer with 2 neurons, is the yield layer with softmax actuation. The weights and inclinations for these is appeared within the table "Table. III" This shows the extension's utilization was based on the Adam optimizer with a learning rate set at 0.00001 for 8 epochs.

Dataset Arrangement.

To produce our dataset, we utilized the maize assortment, Super 900 M - F1, which is to a great extent accessible at the Eastern Terai Locale of Nepal. We included fine seeds, harmed seeds and remote components within the dataset and categorized the seed part in rate premise (Table 1) as Great, Great, Normal, Terrible and Most exceedingly bad as in figure 1. To begin with, we clicked the pictures of the seed part with a ordinary camera with not so great determination. But the specifying that the framework require to recognize between the fine seed and harmed seeds, were not great sufficient which seem make our framework defenseless to wrong expectations. So we once more clicked 3000 pictures with high-quality camera and produced high-quality dataset with sharp specifying. We attempted to oblige all the conceivable orientations of the seeds within the seed part so that the show learns all the diverse positions of the seeds within the cluster. At last, we isolated the dataset as 2500 preparing information and 500 testing information having a place to five categories as

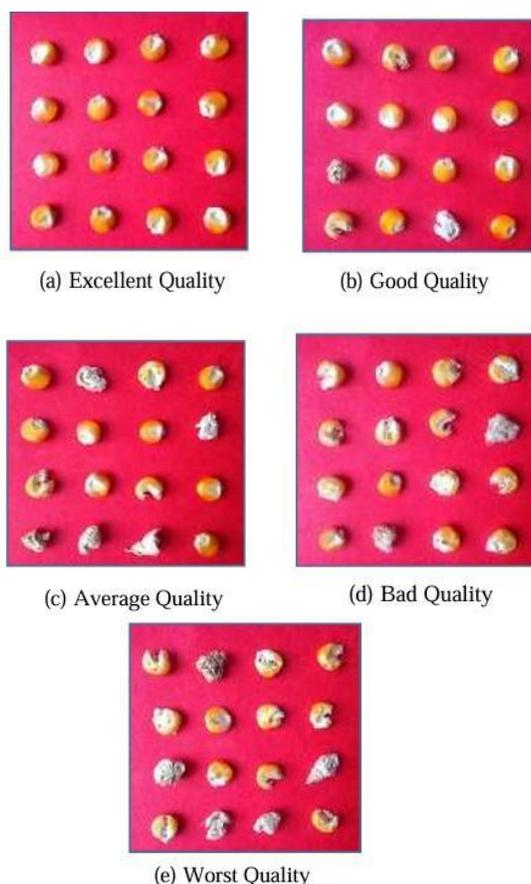


Fig 1: Different Qualities of the Seed Lots based on the percentage of Fine Quality Seeds

Categories (Seed Lot)	Fine Seeds (%)	Damaged Seeds or Foreign Elements (%)
Excellent Quality	100	0
Good Quality	61 – 99	1 – 39
Average Quality	40 – 60	40 – 60
Bad Quality	1 – 39	61 – 99
Worst Quality	0	100

Table 1: Percentage of fine seeds, and damaged seeds or foreign elements in different seed lots

Image Pre-processing

"To prepare the dataset for the neural network, we utilized the data augmentation functions provided by the Keras library. This involved processing all images to a uniform resolution of 256x256 pixels. While increasing the input shape beyond this resolution could potentially improve accuracy, it would also slow down processing speed. The chosen resolution ensures that distinguishing features of damaged seeds are preserved, aiding the machine learning process. The parameters adjusted during augmentation included rescaling, shearing, zooming, width and height shifting, rotation, and horizontal and vertical flipping. Following dataset preparation, all images have uniform attributes and are ready for input into the Convolutional Neural Network."

Building Convolutional Neural Network (CNN)

Building a Convolutional Neural Network (CNN) requires careful consideration to preserve details, especially in layers handling fine features. Our approach prioritized maintaining high resolution throughout the network architecture to accurately capture and process damaged and fine seeds in input images.

Convolution and Pooling Layers:

Our CNN architecture comprises three convolution layers, each followed by a max-pooling layer. In the convolution layers, we employed feature detectors to scan images for specific features efficiently. The pooling operation was utilized to reduce feature map resolution and achieve spatial invariance. We opted for the max-pooling function to extract dominant features from the feature map. Specifically, the convolution layers were configured with 64, 128, and 256 feature detectors, sequentially, designed to process 256x256 resolution images with a depth of 3 RGB color channels. A 3x3 feature detector matrix extensively scanned the image for relevant features. In the max-pooling layers, a 2x2 pooling matrix with a stride of 2 was used to extract maximum values, representing dominant features in the image.

ReLU Activation:

The Rectified Linear Unit (ReLU) activation function played a crucial role in our CNN architecture. ReLU effectively removes linearity from non-linear images, thereby enhancing the training process and enabling accurate predictions. Figure 2 illustrates the graphical representation of ReLU, showcasing how it prunes negative values, retaining non-linear details crucial for image analysis. By applying ReLU activation, our CNN efficiently captured and processed intricate image features, leading to improved model performance and convergence towards accurate predictions.

By integrating these components seamlessly into our CNN

architecture, we achieved a robust framework capable of preserving fine details while effectively extracting and leveraging dominant features essential for accurate image analysis and prediction tasks.

loss is consistently lower than the training loss for all 8 epochs, with the validation loss gradually decreasing compared to a steep decrease in the training loss."

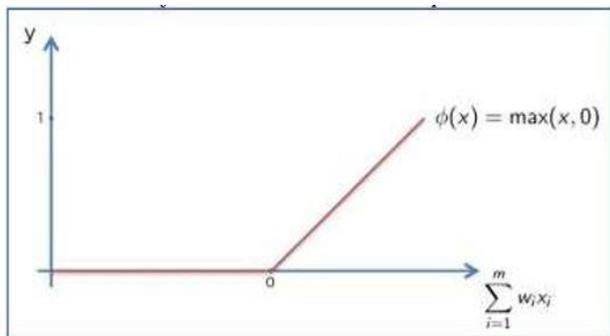


Fig : Rectified Linear Unit (ReLU) Activation Function

$$\mathcal{L} = -(y \ln(p) + (1 - y) \ln(1 - p))$$

TABLE II
CONFUSION MATRIX FOR SOYA BEAN TESTING WITH HARDWARE

	BAD ACTUAL	GOOD ACTUAL
BAD PREDICTED	TN : 43	FN : 0
GOOD PREDICTED	FP : 7	TP : 50
%Accuracy : 93.00%		

RESULTS AND DISCUSSIONS

"The CNN model underwent training with the Adam optimizer, using a learning rate of 0.00001 across 8 epochs and a batch size of 8. The dataset used for the model comprised 320 seeds, evenly distributed with 160 seeds of each type. The dataset was divided into training and validation sets, with 80% (256 seeds) used for training and 20% (64 seeds) for validation after each epoch. The model achieved an accuracy of 96.875%."

$$\%Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

The model was able to distinguish between the dominating and certain low-level features and classify the soya bean seeds.

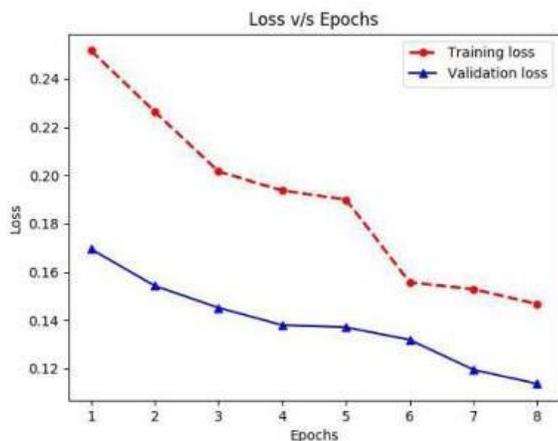


Fig. 5. Loss v/s epoch

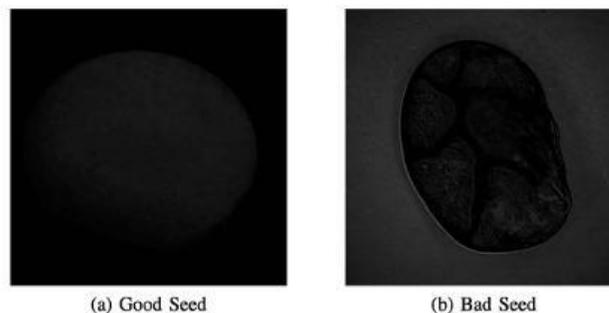


Fig. 6. Images from intermediate layers of the CNN

"Figures 6a and 6b depict images from intermediate layers of the neural network captured during 8 epochs... The weights of the last two connected layers were extracted. The penultimate layer contains 8 neurons, while the last layer contains 2 neurons, as the classification task involves categorizing soya seeds into two groups: good and bad. Table III presents the weights of each neuron in the penultimate layer to the last layer."

The model, including its structure, was tested against a set of 100 seeds, consisting of 50 seeds of each type, and achieved a correct segregation rate of 93.00%. The confusion matrix for the soya bean seeds tested using the hardware setup is presented in Table II. Out of the 50 bad seeds, 43 were correctly classified as bad and 7 as good. Among the 50 good seeds, all were correctly segregated. The classification accuracy was calculated using the formula described in Equation 2. The categorical cross-entropy loss, as shown in Equation 1, was computed using 'y' as the binary indicator of correct classification and 'p' as the predicted probability of a class. Figure 5 displays the Loss vs Epoch graph for both training and validation. It can be observed from the plot that the validation

TABLE III
WEIGHT VALUES

Neuron number	Weight 1	Weight 2
1	-0.5357173	0.5577181
2	-0.33176324	-0.34641203
3	0.6978684	0.6903747
4	0.06273707	-0.5907885
5	-0.34443784	-0.17652501
6	0.34389704	0.3173514
7	0.5729826	0.28666618
8	0.09759291	0.51804036

CONCLUSION

"The study showcases an innovative method for automated seed sorting, employing a Convolutional Neural Network (CNN) model on a hardware setup and achieving a classification accuracy of 96.875%. The experiment focused on sorting soya bean seeds based on images captured from the hardware setup's rotating disc. By leveraging a pre-trained CNN model, seeds were classified as either good or bad and segregated using servo motor mounted guides, resulting in a 93.00% classification accuracy. This approach can be extended to sort other types of grains, pulses, and oil seeds.

To enhance the system, improvements in both hardware and software can be made. Incorporating advanced microscopes with a wider field of view could increase the sorting rate by capturing images of multiple seeds simultaneously. Additionally, using faster processors could expedite the sorting process. With the right dataset, this approach could be adapted to work with any seed type."

REFERENCES

[1] Miss Shivpriya Desai, Dr. A. P. Rao, Seed Quality Analysis using Image Processing and ANN, International Journal of Trend in Scientific Research and Development, Volume 1(4), ISSN: 2456-6470.

[2] Divya Agrawal, Prof. Priyanka Dahiya, Comparisons of Classification Algorithms on Seeds Dataset Using Machine Learning Algorithm, COMPUSOFT, An International Journal of Advanced Computer Technology, 7 (5), May-2018 (Volume-VII, Issue-V)

[3] Dr. T. Avudaiappan, S.Sangamithra, A.Silpha roselin, S.Sherin farhana, KM.Visalakshi, Analysing Rice Seed Quality Using Machine Learning Algorithms, SSRG International Journal of Computer Science and Engineering (SSRG UCSE) Special Issue ICRTCRET Mar 2019

[4] Ranjith Bose M, Ranjith K, Suraj Prakash, Subham Kumar Singh, Dr Vishwanath Y, Intelligent Approach for Classification of Grain Crop Seeds Using Machine Learning, International Research Journal of Engineering and Technology (IRJET), Volume: 05 Issue: 05, May-2018, e-ISSN: 2395-0056

[5] Tekalign Tujo G., Dileep Kumar G., Elifenes Yitagesu D., Meseret Girma B., 2017, A Predictive Model to Predict Seed Classes Using Machine Learning, INTERNATIONAL JOURNAL OF ENGINEERING RESEARCH & TECHNOLOGY (IJERT) Volume 06, Issue 08 (August 2017)

[6] Hussain, Lal & Haroon Ajaz, Raja. (2015). Seed Classification using Machine Learning Techniques. Journal of Multidisciplinary Engineering Science and Technology.

D.Sivakumar, K.SuriyaKrishna, P.Akshaya, G.V.Anuja, G.T.Devadharshini, Computerized Growth Analysis of Seeds Using Deep Learning Method, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-7, Issue-6S5, April 2019

[1] A. Chaugule, S. Mali, Seed technological development a survey, Proceedings of the International Conference on Information Technology in Signal and Image Processing, pp. 71 78, ACEEE, 2013.

[2] Qiu, Z.; Chen, J.; Zhao, Y.; Zhu, S.; He, Y.; Zhang, C. Variety Identification of Single Rice Seed Using Hyperspectral Imaging Combined with Convolutional Neural Network. Appl. Sci.2018,8, 212.

[3] Kurtulmu F, Alibas, Kavdr I. Classification of pepper seeds using neural network. Int J Agric & Biol Eng, 2016; 9(1): 5162.