JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

Seed Quality Assessment and Leaf disease Prediction Through Non-Destructive Techniques

Dr.B.V.D.S.Sekhar
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Gokarakonda Mani JayaSurya (Lead)
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Bhupathiraju Sujith Varma
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Chinnam Ajay Ratnam
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Achanta Pavan Kalyan
Department of Information Technology
S.R.K.R. Engineering College(A)
SRKR Marg, Bhimavaram.

Abstract— **Seeds, as living products, require proper cultivation, harvesting, and processing to optimize their viability and crop productivity. Chemical and physical techniques have proven to be effective and reliable, with high accuracy rates. However, these techniques are often accompanied by certain drawbacks such as high costs, extended processing times, and a requirement for skilled operators. To overcome these limitations, alternative approaches are being explored, including the use of machine learning algorithms and image analysis techniques. These innovative methods have the potential to significantly improve the efficiency, speed, and accuracy of seed quality assessment while reducing the need for manual intervention.. Non-destructive technologies are rapid, accurate, reliable, and simple methods for assessing the quality of the seeds. Some research based on image processing and analysis has been explored in terms of the assessing external and internal quality of a variety of seeds. Another important aspect of agricultural productivity is detection and elimination of diseases during growth period. Generally, leaves are the most effected part during a disease since they are food factories to plants/crops. This can cause a significant decline in production if the disease is not identified and eliminated. Research supports that detection of these diseases can be done easily by employing machine vision where there is less scope for error. Hence, we aim to develop a model using the above-mentioned techniques which can be used to assess the quality of seeds, and a model for prediction of leaf disease.**

I. INTRODUCTION

Seeds are the primary component of plant and crop development, and because of their enormous biological and economic significance, farmers, producers, and seed testing facilities frequently take considerable care to ensure good quality. Much study has been concentrated on seed quality evaluation due to increased concerns about food supply and improving the overall quality of agricultural products. All parties participating in agricultural production chains, including farmers, traders, distributors, and a number of other parties, view high seed quality as an essential element of successful crop production. By the way, while the definition of "seed quality" varies depending on who you ask in the agricultural production systems, the term itself typically refers to purity, homogeneity, and germinability,vigor, viability, seed health, and freedom from physiological disorder, insect infestation or any other deterioration symptoms. Even when using a healthy and viable seed for agriculture, the plants still need to be carefully watched over. Agricultural diseases pose a serious threat to the security of our food supply. Pests and illnesses cause crops to be destroyed or plant parts to wither, which lowers food output and increases food insecurity. Also, little is known about diseases and pest management or control in many less developed nations. One of the main causes of decreased food production is toxic pathogens, poor disease control, and dramatic climate changes. A farmer's transition from one disease control strategy to another is quite challenging. The classic method used in practise for the discovery and identification of plant disease is expert observation with the unaided eye. But this procedure is more expensive in large plantations and sometimes this may be less accurate. In some countries like India, farmers may have to show the specimen to experts, this makes time consuming and more expensive Due to lack of infrastructure rapid identification remains a problem in most parts of the world.

Hence, the development of non-contact and faster methods with minimum human interposition is critically required to pledge seeds of the highest quality for production and trading purposes for disease detection, the image processing methods are suitable and efficient with the help of plant leaf images. Though continuously monitoring of health and disease detection of plants increases the quality and quantity of the yield, it is costly. Machine learning algorithms are experimented due to their better accuracy. However, selection of classification algorithms appears to be a difficult task as the accuracy varies for different input data.

The objectives are to detect leaf disease portion from the image, extract features of an exposed part of the leaf, and recognize diseased leaf through developed model. Therefore, it is essential to have a working machinery to detect the seed quality through non-destructive methods. Seed Germ is a cost-effective phenotyping platform for automated seed imaging and machine learning based phenotypic analysis of crop seed germination. It uses ML-based analytical software for measuring both germination and other establishment related traits during the quality determination. Various methods have been proposed to detect leaf diseases. A simple economic system is the end goal of all the research. By using machine learning algorithms, a novel approach is proposed to assess seed and leaves by means of machine vision. Seed quality assessment module requires additional hardware since the study of external features limits the attributes needed. In contrast most leaf diseases can be identified through external features.

Though we have some existing systems to assess the quality of seeds we propose to improve the reliability of such models by means of adding additional features and integrate these with leaf disease detection module, to improve assistance to the agricultural stake holders.

II. Literature Survey (Related Work)

A. *Salome Hema Chitra,S.Suguna and S.Naganandini Sujatha, "A Survey on Image Analysis Techniques in Agricultural Product", Indian Journal of Science and Technology.*

The authors of the study suggested a thorough examination of image analysis methods for assessing agricultural products. They specifically designed a five-step seed identification processing mechanism. Seed sample photos were initially gathered as the framework for additional processing. Noise removal and picture enhancement techniques were used during the pre-processing of the seed photographs. After segmentation, edge detection was used to improve the image. To distinguish between normal and defective seeds, the segmented image was examined to extract attributes like colour, shape, and texture. Then, to enable precise seed identification, these traits were applied to image analysis tools

B. *D. Wang ,F. E. Dowell,M. S. Ram &W. T. Schapaugh . Classification of Fungal-Damaged Soybean Seeds Using Near-Infrared Spectroscopy*

Showed Using PLS models with NIR, healthy and fungal-damaged soybean seeds may be distinguished with ease. When a two-class model was applied to the wavelength range of 490-1690 nm, classification accuracy was greater than 89%. To distinguish between distinct kinds of fungal damage, five-class neural network models were created.

C. S.Verma, A. Chug, and A. P. Singh, "Application of convolutional neural networks for evaluation of disease severity in tomato plant.

A collection of several forms of tomato leaf imagery from the PlantVillage dataset was used to highlight the usage of three different CNN models, AlexNet, SqueezeNet, and Inception V3, to assess the level of late blight pathogenicity in tomato plants in the early, middle, and end-stages. Models were implemented utilising feature extraction and transfer learning, where the multiclass support vector machine was trained using the obtained features, for better performance. The accuracy of the classifier, mean F1-score, and recall were the criteria against which the study evaluated the performance of all three models.

III. System implementation (Methodology)

A subset of deep neural networks is CNN. The CNN model is represented in figure 3 as having an input layer, a convolution layer, a pooling layer, a fully connected layer, and an output layer. The photos are used as input to accurately classify the plant illness. The features from the photos are extracted using the convolution layer. The feature values are computed by the pooling layer using the retrieved features. The convolution and pooling layer can be further expanded to extract more details depending on the complexity of the images.

The output of earlier levels is combined by a completely linked layer into a single vector that may be utilized as the input for the following layer. The output layer categorizes the plant disease in the end.

Dataset: We use the New Plant Diseases dataset, which includes the following 38 are 70295 photos in the collection.

In the data set,63266 of these photos were utilized for training, while the remaining ones were used for testing.

We use the Soyabean seed dataset, includes the 5000 images.

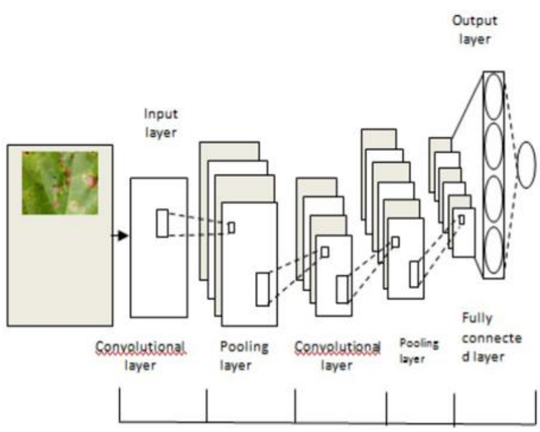
Belonging to 6 classes

In the data set 3867 images are used for training

3.1 Methodologies

3.1.1 Convolutional neural networks are a specific type of feed-forward artificial neural network whose neuronal connection topology is modelled after the visual cortex. Convolutional neural networks, or covnets, are just regular neural networks with shared parameters. The minuscule width, height, and depth of the learnable filters utilised in the convolutional layers are identical to those of the input volume.

CNN Architecture:



3.1.2 TensorFlow is a complete open source machine learning platform. Developers can easily design and deploy ML-powered apps using its vast, flexible ecosystem of tools, libraries, and community resources, while researchers can enhance the state-of-the-art in machine learning. TensorFlow provides a collection of steps with straightforward, high-level APIs for creating machine learning models in a range of languages for both beginners and experts.

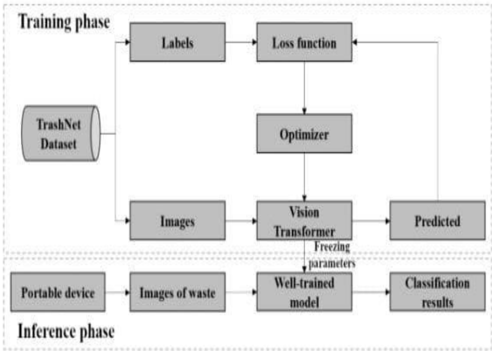
Developers can choose to deploy models on a variety of JavaScript platforms, including servers, the cloud, mobile and edge devices, browsers, and many others. This makes it considerably simpler for developers to move from model construction and training to deployment. All of the information is made available to programmers by TensorFlow using the Python language. TensorFlow is supported on Python versions 3.7 through 3.10; previous versions of Python may also work, but this is not guaranteed. TensorFlow applications are Python apps, while its nodes and tensors are Python objects. The transformation libraries that TensorFlow offers are written as high-performance C++ binaries. With the Keras library, a high level of tensor flow work is able to create nodes, layers, and links between them.

3.1.3 The ResNet50 variant of the network features convolution layers in each residual block, one before and one after the typical three by three convolution layer. The residual block of the ResNet suggested in [6] attempts to analyze the residual portion of the correct output. It uses the identity mapping shortcut connection to incorporate previously unrepresented portions of the network into the output. These 11 conv. layers "leave the three layer a bottleneck with decreased input/output dimensions" and "keep the same dimensions of the identification section and the residue half," which restrict and then amplify dimensions. We first use a convolutional layer and a pooling layer to get the rough properties of images in the ResNet50 model. The final residual block is coupled with a common pooling layer to downscale the characteristic matrix, a flatten layer to convert it to a vector, a dropout layer, and an absolutely linked layer to categorize the picture elements.

3.1.4 The Inception v3 model, which was unveiled in 2015, has 42 layers overall and a lower error rate than its predecessors. Significant modifications have been made to the Inception V3 model, including.

- Convolution Factorization Into Smaller Convolutions
- Spatial Factorizations into Asymetric Convolutions
- Usefulness of Auxiliary Classifiers
- Effective Decrease of Grid Size

3.1.5 A flowchart is designed that project shows the process flow of the entire project



a) Collection of datasets: Upon being captured, photographs undergo a meticulous categorization process wherein they are sorted into multiple distinct groups, including but not limited to In order to achieve optimal precision, comprehensive training of the model is of utmost importance. This involves initial labelling and sequencing of images, followed by segregation into two distinct sets, namely the training dataset and the testing dataset..

b) Image pre-processing involves the application of a diverse range of operations to images at the lowest possible level of abstraction. Its primary objective is to mitigate undesirable distortions and augment pertinent image information essential for subsequent processing. Effective image processing is a vital step towards obtaining the desired outcome. It enables the performance of various actions such as batch-size adjustment, rescaling, label generation, image size modification, shear range control, zoom range calibration, and more.

c) Training Data: Machine learning involves a common objective of researching and developing algorithms that leverage prior successes to make diverse predictions on a given dataset. To calibrate the model's parameters, a training dataset is initially employed as a reference. This dataset serves as an exemplar and is used to fit the model's parameters.

d) Data for testing: The data utilized for software system testing is commonly referred to as test data. Test data is distinguished by a specific identification and can be provided by either automated testing technologies or testers themselves. The primary purpose of test data is to enable regression analysis, as identical data can be repeatedly utilized for testing.

e) Model Evaluation: Cross-validation and hold-out are two widely used techniques in data science to evaluate the performance of a model. In order to prevent overfitting, a test set is employed to assess the model's performance.

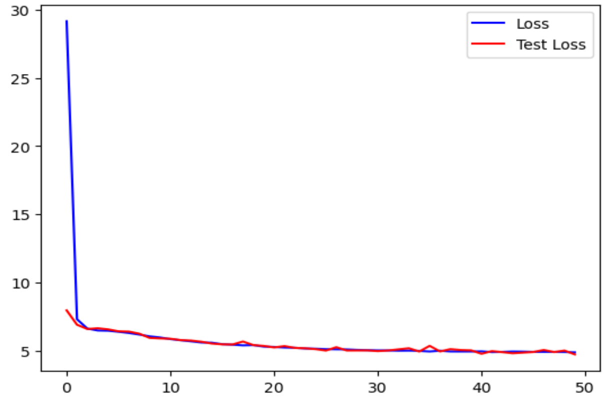
IV. EXPERIMENTS & RESULTS

The proposed solutions aim to utilize various machine learning techniques to achieve higher accuracy in identifying leaf diseases and seed quality, while also reducing labor requirements. The model accurately classifies the image into the appropriate category, enabling efficient disease and quality assessment.

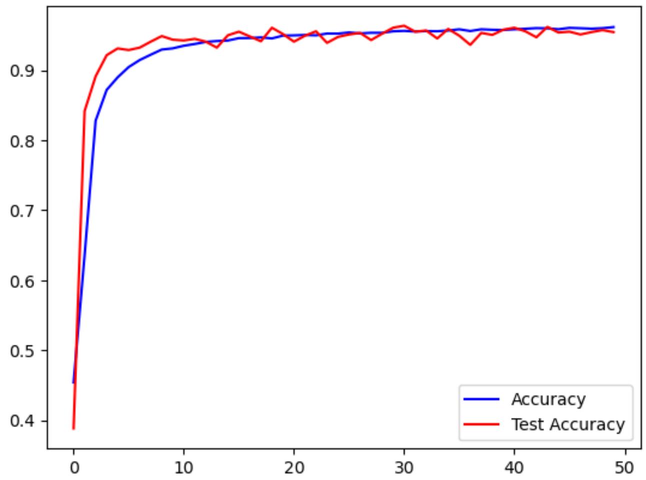
We defined the model named with some hyperparameters
We added few hidden layers for better accuracy and output

Finally we compiled and fit the dataset into the model and trained it with 50 EPOCHS

We plotted a graph between the training loss function and validation loss function.



We plotted a graph between the training accuracy and validation accuracy to understand the bias.



We plotted a graph to find the correlation between learning rate and loss.

V. EVALUATION METRICS

We plotted a confusion matrix to know the accuracy percentage of validation data. Basically, this matrix shows how much percentage of data is predicted correctly in the validation data set and it also gives how much validation data is not predicted correctly.

VI. CONCLUSION

To summarize, our proposed system utilizes machine learning algorithms for seed quality assessment and leaf disease detection, enabling the separation of different waste components. This technology has the potential to reduce pollution and illness while also minimizing the need for human intervention. Our model achieved a remarkable accuracy of 97% for the leaf module and demonstrated excellent performance when evaluated against the leaf dataset. The proposed method allows for the creation of an improved web application to classify seeds and diagnose leaf diseases.

.

VII. FUTURE WORK

The accuracy of the system can be increased in the future if more images are uploaded to the dataset. By adjusting some of the employed parameters, we will tend to improve our system so that it can classify more leaf diseases items. Machine learning techniques enable the classification of seed and leaf within the context of labeling. To increase accuracy rates, more data is required. Without utilizing any data augmentation techniques, we have obtained good classification success in the context of our suggested model. According to studies, a more accurate model can be produced if the images collected from other parts of the spectrum like infrared and x-ray. The images from various parts of the spectrum can help improve accuracy by helping extract more features.

Acknowledgement

Without the outstanding help of my supervisors, Professors Dr.B.V.D.S.Sekhar and Dr. K. Kishore Raju sir, I am grateful for the valuable support and guidance provided by the team at Sagi Rama Krishnam Raju College throughout the project. Without their timely recommendations and insights, as well as the opportunity to gather data from various sources facilitated by the organization's personnel, it would have been difficult to complete the work on schedule. Their contributions were indispensable to the successful outcome of the project.

References

[1] Salome Hema Chitra,S.Suguna and S.Naganandini Sujatha, "A Survey on Image Analysis Techniques in Agricultural Product", Indian Journal of Science and Technology.

[2] Timothy J. Herrman, Extension State Leader Grain Science and Industry Carl Reed "Soybean Grading Procedures", Kansas State University.

[3] D. Wang ,F. E. Dowell,M. S. Ram &W. T. Schapaugh. Classification of Fungal-Damaged Soybean Seeds Using Near-Infrared Spectroscopy.

[4] Sudhir Gupta, Smitha Patil, Asha Kulkami and Zoheb Hukkeri "Classification of Soybean seeds by Color Image Analysis" 5th

International Conference of the Asian Federation For Information Technology in Agriculture (AFITA).

[5] Ma, F.; Wang, J.; Liu, C.; Lu, X.; Chen, W.; Chen, C.; Yang, J.; Zheng, L. Discrimination of kernel quality characteristics for sunflower seeds based on multispectral imaging approach.

[6] A. Bhatia, A. Chug, and A. Prakash Singh, "Application of extreme learning machine in plant disease prediction for highly imbalanced dataset," Journal of Statistics and Management Systems, vol. 23, no. 6, pp. 1059–1068, 2020.

[7] S. Verma, A. Chug, and A. P. Singh, "Application of convolutional neural networks for evaluation of disease severity in tomato plant," Journal of Discrete Mathematical Sciences and Cryptography, vol. 23, no. 1, pp. 273–282, 2020.

[8] J. Chen, J. Chen, D. Zhang, Y. A. Nanehkaran, and Y. Sun, "A cognitive vision method for the detection of plant disease images," Machine Vision and Applications, vol. 32, no. 1, pp. 1–18, 2021.

[9] https://www.kaggle.com/datasets/asheniranga/leaf-disease-dataset-combination

[10] **https://www.mdpi.com/journal/sensors**

[11] **https://scholar.google.com/**