IDENTIFICATION AND PURITY ANALYSIS OF SEED USING CONVOLUTIONAL NEURAL NETWORK

Ms. Kalaivani N
Department of Electronics andCommunication Engineering, Sri Krishna College Of Engineering And Technology Coimbatore, India kalaivani@skcet.ac.in

Deva Dharshini S
Department of Electronics andCommunication Engineering, Sri Krishna College Of Engineering And TechnologyCoimbatore, India 19euec033@skcet.ac.in

Divya S
Department of Electronics andCommunication Engineering, Sri Krishna College Of EngineeringAnd Technology Coimbatore, India, 19euec040@skcet.ac.in

Jeya Sundari K S
Department of Electronics andCommunication Engineering, Sri Krishna College Of EngineeringAnd Technology Coimbatore, India, 19euec060@skcet.ac.in

ABSTRACT

Agricultural research has gained traction in recent years. A farmer spends countless hours preparing for their crop, from researching and choosing the desired variety of seed, to make a successful crop and yield. And the reality is, all of the time, energy, and preparation can be immediately torpedoed if the seeds received and planted by the farmer are bad or defective, which is a problem a farmer, has no control over. There are many factors involved in agricultural production, and seed qualities are at the top of the list. Seed testing contributes to the reduction of agricultural production risk. Seed testing identifies quality faults in seeds prior to planting, saving time and money. Nowadays, virtually the entire world's population depends on seeds, or the plants they develop into, to carry out daily life. Our purpose is to decide the standards of a seed via means of identifying and testing physical purity, via a means of allowing the farming network to get exceptional seeds with the use of CNN.

KEYWORDS- Seeds, CNN, Convolution layer, pooling layer, Classification and purity test.

DOI Number: 10.14704/nq.2022.20.9.NQ44380 **Neuro Quantology 2022; 20(9):3296-3301**

I. INTRODUCTION

Seed testing is the cornerstone for all subsequent seed technology procedures. Seed testing is used to control quality parameters during seed handling. It is a quality measure that can be measured and guaranteed by seed performance. Our goal is to establish seed quality standards, including physical hygiene, so that farmers may obtain high-quality seed through CNN. Seed identification can be helped by deep learning and computer-aided viewing approaches. It was constructed, taught, and tested to recognize four different types of seeds in this study using a deep learning method and a computerized detection tool. Image acquisition is done under controlled lighting conditions. The image categorization procedure was utilized to cut down on the time it took to gather the training photos needed to train convolutional neural network models. Various solutions were also tested to find the best in-depth learning model with accuracy and training time. Approximately seven seeds were tested each for training and testing. Highly segmented accuracy (95%)

II. was achieved.LITERATURE SURVEY

Image processing, a computer-assisted theory, has long captivated experts. Image processing involves the removal of a feature, which is a difficult task and supports identifying the object and the separation. Machine learning strategies are employed for training the system to extract the elements using 900 cotton seeds. In seed separation cotton, the tree separation of decision and MLP has a 98% accuracy rate, and the division of the Nave Bayes has a 94% accuracy rate. According to the findings, the Nave Bayes separator makes 52 inaccurate classifications, while the decision and MLP tree separations each make 11 errors.

Agrawal and Dahiya used logistic regression to compare machine learning algorithms for categorizing different grain seeds. Using logistic regression, Linear Discriminant Analysis, k-Near Neighbor Separator, Tree Separator, Gaussian Naïve Bayes and vector support machine. This study reported performance levels for both lines as well as incompatible algorithms. The level of accuracy for the six algorithms is as follows: Logistic Regression level was 91%, Linear Discriminant Analysis level was 95%, k-Near Neighbour level was 87%, Tree Separator was 88%, Naïve Bayes rate was 88% and Support Vector Machine rate was 88.71%.

III. PROPOSED SYSTEM

BLOCK DIAGRAM

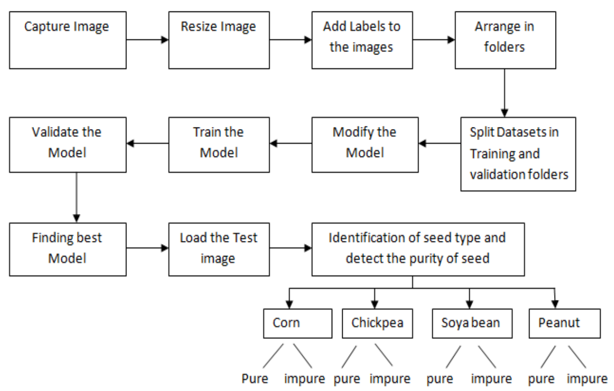


Fig: 1 Block Diagram of the desingned model

PROJECT METHODOLOGY

A new model based on CNN was established in the proposed work, which provides an efficient technique to identify and categorize seeds. The whole process consists of three levels:

- Database structure
- Training
- Testing.

The ultimate aim of this system is to distinguish healthy seed photos from newly germinated or bug-infested seed photographs. Therefore, all archives are divided into 4 sets:

- For training purposes, images of healthy seedlings
- For testing purposes, images of healthy seedlings
- For training purposes, images of unhealthy seedlings
- For testing purposes, images of unhealthy seedlings

Training images are treated as inputs for computer- assisted learning to understand incompatible with materials. Test images are then treated as controls to test the proposed algorithm.

CONVOLUTIONAL NEURAL NETWORK (CNN)

A Convolutional neural network is an algorithm which comes under deep learning which takes the image of the dataset and gives the necessity to each and every image of various classes and they are able to differentiate the one image from the other which belongs to the same class. The CNN is used mainly for image processing, classification and segmentation of the data.

CNN is a network designed to process structured data such as images. It consists of two stages. Namely,

1. Feature Extraction – consist of convolution and pooling layer

2. Classification – consist of dense neural network part

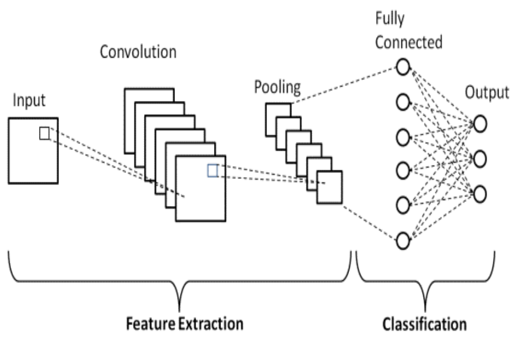


Fig: 2 Architecture of CNN

Data Set Formation:

For dataset formation, four different types of known seeds were chosen. The seeds chosen were corn, chickpea, peanut, soya bean. It's really well worth noting that the wide variety of pictures taken for every form of seed have to be huge a good way to keep away from elegant differences. There will be a bias effect on preference for a prominent set if one seed collection has more photos than the others.

Training:

The training phase represents the proposed system's training based on the data generated, in the dataset formation stage. For training, this phase is classified into two phases namely, the forward and backward phase.

- The forward phase is where the inputs are passed via the network.
- The backward phase is where updating of weights happens.

There are eight classes in the dataset. The training phase has some features such as image enhancement, decay level of learning and system testing. Generally, trained models result in better performance for a training set but fail once they are applied on to a test set.

Avoiding excess and improving the ability to do normal deep learning models, the process of adding an image was used. During image magnification, various images are made by embedding in various processing methods or combinations of many processing methods.

Testing:

The testing phase was started by applying a very well-trained system from the training phase. Dataset for testing was prepared. Unlike the training database, the images were randomly taken at various angles and positions by adjusting the light levels. This experimental database assisted in making the data model invisible to a real-world file. The size of the photos was changed to match the size of the suggested model. The images are fed into the proposed model after being resized, in order to guess the test data. Scores of the outcomes are collected during the model testing phase.

WORKING

• In the beginning, numerous varieties of seed were collected, with each type containing two sorts of seeds: healthy and unhealthy seeds.

• A seed is considered to be healthy when it is not damaged or has not started to germinate etc. In the case of faulty seed, the reverse is true. After the pre-processing of the image takes place in which a sample is captured and given as an input.

• Then, in feature extraction, the image is further processed in which some phases take place, like Convolution layer and pooling layer.

• In the convolution layer, various filters are applied to the original image. Pooling layer is used for dimensional reduction. Apooling layer is used for reducing spatial size.

• In the classification, a fully connected layer is present in which the features from the convolution and from the pooling layers are grouped.

• Features given by the fully connected layer are checked for identification of seed and purity analysis and, as a result, the type of seed and purity is known.

IV. RESULTS

In the proposed work, Anaconda python is used to update packages in the project, Tensorflow package is used for classification, understanding and in the prediction of the model .Jupyter Lab is the platformused. In this project, 320 images are used as a dataset. Out of which, for training, 256 images are used and for validation, 64 images are used.

Sample images:

(a) Corn Pure

(b) Corn impure

(c) Chickpea pure

(d) Chickpea impure

(e) Soya bean Pure

(f) Soya bean impure

(g) Peanut Pure

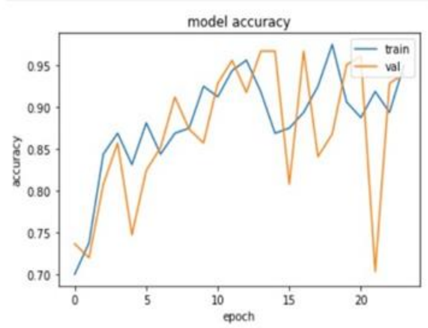
(h) Peanut impure

Fig 3 Images from the dataset

Images in fig 3 represent the training and testing dataset.

Following graphs explain the model's training, validationaccuracy and loss:

(a) Model's training and validation accuracy.



(b) Model's training and validation loss

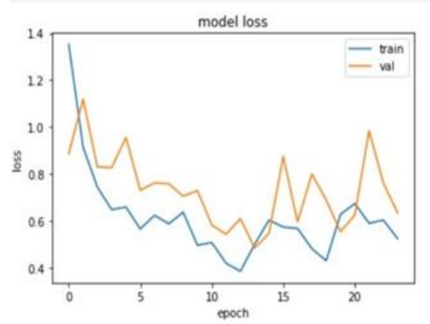


Fig: 4 Accuracy and loss of the model.

From image 4(a), we can obtain the accuracy of the testing and validation of the model. Initially, the accuracy for training of the model was as low as 0.6. On further increment in epoch values, it attains its maximum accuracy of 0.89. The low level accuracy of is due to lack of training. The model's functioning remains constant initially and then, there is a slow flow to improve model performance. At a certain point, it gains its maximum. After this point, the functioning of the model remained static. This is because of the use of more layers in CNN.

From image 4(b), we can obtain the losses from training and validation of the proposed model. Initially, the model loss value was very high, which was also during the starting stage of the training process as the model had been met with only limited data samples. It also shows at the loss of training and validation decreases as epoch increases. To increase the accuracy rate and to decrease the error rate, models should be given a greater number of images so that models are well trained and validated and the result is high model performance.

To enhance performance, the proposed system changed to be properly designed. To raise the accuracy and decrease errors, the proposed system changed to be made trained and validated.

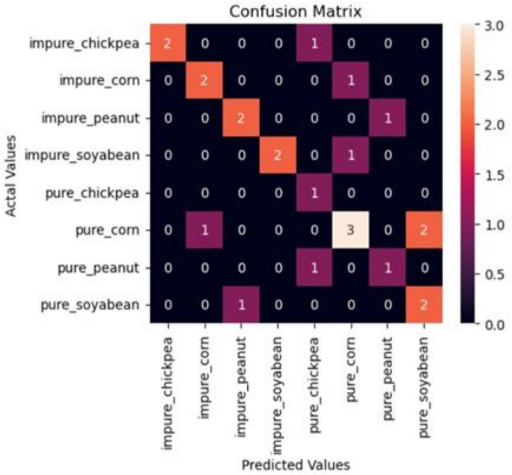


Fig 5 Confusion Matrix of the model

The above figure depicts the confusion matrix of the proposed model on the testing set.

Using the confusion matrix, the following values are calculated for the class pure corn is:

For the class corn pure ,the true positive value is 3,as only 3 images are predicted as corn pure whose actual image is also corn pure.

False Negative value is 3, as only 3 images correspond to the values, sum of rows except the value in True positive.

False Positive value is 2, as only 2 images correspond to the values of sum of columns except the value of True positive.

True Negative value is 16, as only 16 images and of given images is not corn pure and predicted as corn pure.

The above results represent the solid overall performance given for maximum of the classes within the model. The results from the testing phase are favorable and can be used for classification and analysis of seeds.

VII . FUTURESCOPE

In future, for this proposed works mobile-primarily based totally software the use of a bigger range of various seeds and in this mobile app voice input facilities can be given and also regional languages options can be added. This app will assist farmer with constrained know more about seeds.

VIII. REFERENCES

[1] Luan, Z.; Li, C.; Ding, S.; Wei, M.; Yang, Y. Sunflower Seed Sorting Based on Convolutional Neural Network. In Proceedings of the ICGIP 2019 Eleventh International Conference on Graphics and Image Processing, Hangzhou, China, 12–14 October 2019; Pan, Z., Wang, X., Eds.; SPIE: Bellingham, WA, USA.

[2] LeCun, Y.; Bengio, Y. Convolutional networks for images, speech, and time series. Handb. Brain Theory Neural Netw. 1995, 3361, 1995.

[3] Kuo, C.-C. J. (2016). Understanding Convolutional Neural Networks with A Mathematical Model. Journal of Visual Communication and Image Representation.

[4] Xinshao, W. Weed Seeds Classification Based on PCANet Deep Learning Baseline. In Proceedings of the 2015 Asia-Pacific Signal and Information Processing Association Annual Summitand Conference APSIPA ASC, Hong Kong, China.

[5] Kingma, D.P.; Ba, J.L. Adam: A Method for Stochastic Optimization. In Proceedings of the 3rd International Conference on Learning Representations, ICLR 2015—C, San Diego, CA, USA.

[6] G. E. Hinton, "Recti fied Linear Units Improve Restricted Boltzmann Machines Vinod Nair, " in Proceedings of the 27th International Conference on Machine Learning, Israel.

[7] Hawkins, D.M. The Problem of Overfitting. J. Chem. Inf. Comput. Sci. **2004**, 44, 1–12.

[8] Arnold, T.B. kerasR: R Interface to the Keras Deep Learning Library. J. Open Source Softw. **2017**, 2, 296.

[9] Barman, U.; Choudhury, R.D.; Sahu, D.; Barman, G.G. Comparison of convolution neuralnetworks forsmartphone image based real time classification of citrus leaf disease.

Comput. Electron. Agric. **2020**,177, 105661.

[10] Krizhevsky, A.; Sutskever, I.; Hinton,

G.E. ImageNet Classification with Deep Convolutional Neural Networks.In Proceedings of the Advances in Neural Information Processing Systems 25 (NIPS 2012), Lake Tahoe,NV, USA, 3–6 December 2012; pp. 1097–1105.

[11] Szegedy, C.; Vanhoucke, V.; Io_e, S.; Shlens, J.;Wojna, Z. Rethinking the Inception Architecture for ComputerVision. In Proceedings of the IEEE Computer Society Conference on ComputerVision and Pattern Recognition,Las Vegas, NV, USA, 27–30 June 2016; pp. 2818–2826.

[12] Agrawal, D.; Dahiya, P. Comparisons of classification algorithms on seeds dataset using machine learningalgorithm. Compusoft2018, 7, 2760–2765.

[13] Vlasov, A.V.; Fadeev, A.S. A machine learning approach for grain crop's seedclassification in purifyingseparation. J. Phys.Conf. Ser. **2017**, 803, 012177.

[14] Jamuna, K.S.; Karpagavalli, S.; Vijaya, M.S.; Revathi, P.; Gokilavani, S.; Madhiya, E. Classification of Seed Cotton Yield Based on the Growth Stages of Cotton Crop Using Machine Learning Techniques. In Proceedings of the ACE 2010–2010 International Conference on Advances in Computer Engineering, Bangalore, India, 20–21 June 2010; pp. 312–315.

[15] Santos, L.; Santos, F.N.; Dos Oliveira, P.M.; Shinde, P. Deep Learning Applications in Agriculture: A ShortReview. In Iberian Robotics Conference; Springer: Cham, Switzerland, 2020; pp. 139–151, ISBN 9783030359898.

[16] Santos, L.; Santos, F.N.; Dos Oliveira, P.M.; Shinde, P. Deep Learning Applications in Agriculture: A ShortReview. In Iberian Robotics Conference; Springer: Cham, Switzerland, 2020; pp. 139–151, ISBN 9783030359898.