Volume 66, Issue 2, 2022
Journal of Scientific Research
of
The Banaras Hindu University

A Convolution Neural Network-Based Wheat Grain Classification System

Seema Shedole[1], Sowmya B J[2], Neha Arthi V.P[3*]
[1,2,3]M S Ramaiah Institute of Technology (Affiliated to VTU) Bangalore, India
[1]seemas@msrit.edu
[2]sowmyabj@msrit.edu
[3]neha.arthi@gmail.com

Abstract: In most developing countries, agriculture is the primary economic sector. Seed categorization and evaluation could provide detailed information for seed germination, seed quality management, and impurity detection. Many academics are working on a range of seed image analysis systems using several categorization algorithms. India is the world's second-largest wheat grower. The importance of detecting the wheat quality cannot be overstated. Manually determining wheat quality necessitates skilled knowledge and is time demanding. The healthy and poor wheat growth is a complete indication of biological changes as well as a detecting factor for quality and yield. Hence detection of quality of Wheat Seed can be achieved through advanced technologies like Image Processing, machine learning, deep learning etc, our proposed paper includes various phases of implementation namely image processing techniques, and Deep neural network. In our model there are four classes considered they are Good grain, Damaged Grain, Broken Grain and Foreign Particles. where a convolution system is used to classify the wheat seeds and detect status of 4 distinct classes.

Keywords: Image Processing, Machine learning, Deep learning, CNN.

1. INTRODUCTION

Wheat is the world's most important grain, with yields second only to rice. It is a significant dietary crop and a significant raw material for industry. Computer vision is recognized to be a quite essential application in the fields of agriculture and engineering. Mostly for seedling growth monitoring, crop growth rate research, weed growth assessment, plant growth supervising, plant watering, and fruit, vegetable, and seed variety categorization are important. Seeds are the foundation of the agricultural process. When it comes to seed classes, there are numerous seeds in the industry that relate to the similar or distinct classes, or even distinct varieties within the similar class. Seed categorization is determined by farming fields, climate, and water system. Seed categorization and evaluation have provided more information [9]. Because the key important variables in the seed classification task are morphological aspects and color of seeds, seed examination and grading in the identification of various types of seeds are typically followed with reference to texture traits and color of seeds. In India, there are numerous wheat seed kinds, each with its own consumer and use. As a result, variety assurance is a crucial step preceding growing. Wheat classes are determined by highly skilled professionals who classify types based on their visual qualities. However, this is a qualitative technique with errors, such as differing input from various specialists, are not inevitable [10].

In the agriculture field we mainly concentrate on seeds. Without good quality of seeds there is no chance of producing or harvesting crops. Due to this increase in population growth, land for agriculture will reduce day by day, which in result decreases productivity. The production of the crop must be increased to balance the consumption rate with production rate. A couple of examinations have been gathering.

In the particularly delicate area of seed evaluation, computer vision is progressively being used. It is also a great tool for conducting experiments, monitoring, and categorization in a variety of different manufacturing sector domains. The usage of computer vision to classify the quality of wheat seeds, aids in producing a quick and precise process.

Image processing technologies can be used in a wide range of sectors to enhance the productivity of specific tasks. Image processing is used to include machine learning algorithms. In the sector of agriculture, image processing is commonly used to detect the kinds of numerous food crops and their quality. A machine vision system is a substitute for conventional organic product examination [4].
Previously, biological methods and chemical methods could be used for the identification of rice grain varieties and quality. But these methods are very expensive and time consuming and results are not accurate. On the other hand, machine vision or digital image processing is a nondestructive method. It is a very fast and inexpensive process compared to the biological methods.

A Convolutional Neural Network is a Deep Learning system that could take an image as input, allocate relevance to different

DOI: 10.37398/JSR.2022.660204

aspects in the image, and distinguish between them. CNN consists of several layers called: input layers, convolution layer, relu layer, pooling layer and fully connected layer are the layers of classification. Their central use is picture affirmation and request. A general depiction of CNNs model can be found in Figure 1. CNN can gain usefulness and precision visual imaginary examination using careful segment extractions.

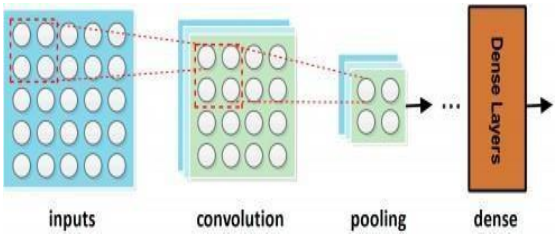

Figure 1. General Representation of CNN Model

The research work proposes an efficient approach for identifying the seed and classifying it based on deep learning CNN model and the application of symmetry. This works basically on CNN which mainly concentrates on transferring the knowledge across all the layers of the model. And classifying 4 different classes of wheat seeds.

Wheat grain is today the most important source of food on earth. Wheat is the staple food for most of the temperate and subtropical regions of the world. Apart from being sufficient, wheat quality is extremely crucial in wheat production. This has been neglected for years, and it has resulted in significant financial loss. Knowing the quality of wheat is important for farmers and for industrialists. However, analyzing and manual seed grading through naked eye and sorting through machines which consider only some features is a tedious task and time consuming and cost would be more and sometimes the results would not be accurate. So hence Sorting the seeds is very important at an early phase. So that precautionary measures can be undertaken to control the damage that occurs by not classifying the seed before sowing this can be achieved through leading edge technology Image processing and Machine Learning methods. Wheat is one of the most widely consumed cereal grains in the world. Seed grading improves seed's competitiveness. Manual grading were the farmers used to grade the seed through their naked eyes

which is very time consuming, requires a greater number of people just to grade the seed one after the other and the result won't be accurate. Image processing is a very important technology, and industrial need appears to be increasing every year. In a variety of domains and applications, image processing and machine learning have shown to be excellent tools for analysis. Since digital image processing is used in this project to detect seeds in wheat, it eliminates the traditional methods which were used in olden days and it removes human error.

The development of the agricultural sector has been given top attention by the state government to improve the rural economy. Because seed is such an important component of agriculture, ongoing efforts are being made to guarantee that farmers have access to high-quality seeds to maintain agricultural development. In the current circumstances, the need for quality seeds is so high that the

O.S.S.C. or any government body alone will not be able to supply the demand for quality seeds, which would be required to be filled by the private seed projects.

- To create a machine learning model that can accurately categorize a given image as good, damaged, broken, or foreign particle or cancerous cell.
- Testing this model with a variety of test sets and achieving expected prediction accuracy with a low error rate.
- To test the accuracy and efficiency of a machine and deep learning classifier method when applied to an image dataset of wheat seeds.
- Classify the seed by identifying the broken, good, foreign particles and other defective seeds.

The scope of the work includes the implementation of an application that would help farmers and industrialists to classify the wheat seeds using machine learning and image processing. Analyzing and manual seed grading through naked eye and sorting through machines which consider only some features is a tedious task and time consuming and cost would be more and sometimes the results would not be accurate. Hence classification of wheat seeds can be achieved through Machine learning and Image processing which is less time-consuming and cost effective, and the results obtained would be more accurate.

2. LITERATURE SURVEY:

Nabeel Ali Abdullah et. al., [1] aims to categorize 3 distinct wheat seeds into their appropriate classifications. There were two steps to the process. The first step involves applying image processing to the obtained images and extracting essential geometric parameters. The collected parameters are put into a neural network built using the back propagation supervised learning in the second stage. Three tests were carried out: the first used all the information, the second used noisy data, and the third used only a portion of the training data.

Using digital image processing and clustered random forest approaches, a novel classification methodology was proposed by Parminder Singh et. al., [2]. The digital image processing is used to calculate metrics including surface, perimeter, length, breadth, groove height, and asymmetry factor. In addition, another model is used to categorize wheat seeds depending on these attributes in a timely manner. The developed method aids the agricultural sector in seed categorization, defective seed isolation, and seed quality management relying on grading criteria. The study described here is employed in categories of

3 different kinds of wheat seeds. It could be used to categorize different varieties of wheat seeds created through the hybridization procedure consequently. In addition, the created approach may be applied to classify a variety of different seeds. Other metrics and machine learning techniques could help to increase the accuracy of the wheat seed categorization.

Raja Haroon Ajaz et. al., [3] performed seed categorization with the help of Weka tool. The information was gathered from a repository on the UCI webpage. Area, perimeter, compactness, length and width of kernel, asymmetry factor, and length of kernel groove are the characteristics of the seed utilized. They employed the Function, Bayes, Meta, and Lazy approaches in Weka classifier.

This research demonstrates the power and possibilities of image processing using well-trained multilayer neural network classifiers to identify uneven rice grain samples cultivated in various agricultural production areas around the nation. As per the business perspective, a machine vision built of proven neural network models can be utilized as a tool to deliver optimal and more unbiased rice quality assessments. Meesha Punn et. al., [4] carried out a systematic review of the image processing strategies utilized in classifying wheat grain integrity. They discovered that there has not been much effort done on identifying the category of Indian wheat varieties employing machine learning algorithms. Image processing algorithms collect features from digital pictures and use them to construct patterns when they are stored in the database. These patterns are fed into machine learning algorithms, which classify the objects into their appropriate categories. Pattern classifiers are a type of computer algorithm which is utilized to classify objects.

Abhijit M Taley et. al., [5] proposed a work that has been accomplished with the goal of providing the operator with the most optimal configuration of the grader machine to optimize the machine's efficiency; nevertheless, the operator gets it wrong, and a lot of good seed or grain is destroyed, leading to a loss to the grower or economy. The goal of the present study was to determine the ideal Grader Machine parameter that would produce the greatest quality seed with the highest production. The aim of this work is to adjust the Grader Machine's parameters to improve the quality of grading and the operating pace. The machine is currently set using a trial-and-error process depending on the operator's expertise and knowledge. P R Kumar et. al., [6] presents the findings of an investigation on the seed grading state of wheat seed accessible to growers for planting, with the goal of evaluating the extent to which wheat seed quality could be enhanced simply by grading and establishing the relationship among seed size and crop development. Seed classification is beneficial in the selection of high-quality seed. It is also a good idea to do it to ensure better crop development and increase seeding rate efficiency in the land. Grading seed, either automatically or manually, based on morphological characteristics such as size, mass, and color, is a common post-harvest management approach for all crops, however the grading requirements differ by crop. When the highest crops are desired, seedlings should be robust, and hence seed must be robust. A healthy seed is one of the most important components in the formation of a sturdy seedling.

S. K. Patil et. al., [7] examined the effects of vibration frequency, deck slope, and air speed on wheat seed separation. The weight of graded seed was identified as a seed grade quality. The most important factor in attaining good seed gradation was determined to be ventilation. Seed of enhanced variety with high physical and genetic integrity, high germination rate, and remarkable vigor, independent of seed-borne illness pests, need-based value added, extended life span, and good storability is referred to as quality seed. To meet the growing demand for high-quality seed, scientists must conduct scientific analyses of grading machines and research the characteristics that influence effective seed grading. WU Wei et. al., [8] applied deep learning methods to overcome the shortcomings of standard computer vision techniques. Wheat grain image datasets were gathered in three distinct kinds, six distinct backgrounds, and two distinct image collection devices with varying heights, orientations, and grain quantities. The model works well with a range of backgrounds, picture sizes, grain sizes, shooting orientations, and shooting altitudes, and varied grain congestion levels. It serves as a useful tool for detecting and counting wheat grain. This research serves as a foundation for future grain analysis and identification projects.

Jamuna et al. [11] used AI methods (for instance Innocent Bayes classifier, a decision tree classifier and MLP) to set up the model to incorporate extraction using an illustration of 900 seeds. They point by point that the decision tree classifier and MLP gave a comparable accuracy in orchestrating the seed, with a speed of 98.7%, and Naïve Bayes classifier had a precision speed of 94.22%. Their results show that Naïve Bayes classifier had the most vital mix-up rate, as it made incorrect game plan on different occasions, however the decision tree classifier and MLP made 11 wrong groupings each.

3. DESIGN AND IMPLEMENTATION

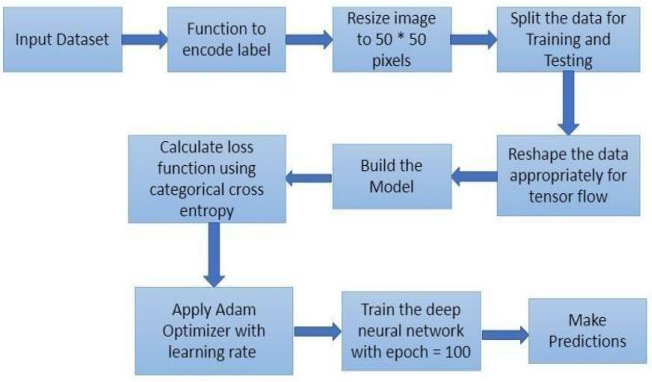

Figure 2: Proposed system for Grading of wheat seeds by detecting whether the seed is good, damaged, broken or contains a foreign matter.

Image classification is where a computer can analyze an image and identify the 'class' the image falls under. A class is essentially a label. Machines can recognize and extract information from images using deep learning. Image classification with deep learning involves convolutional neural network. The research work proposes an efficient approach for identifying the seed and classifying it based on deep learning CNN model and the application of symmetry.

Figure 2 explains the design of the proposed system. Initially wheat seeds have been selected for the preparation of the dataset from Kaggle and their own dataset has been taken. The chosen classes were damaged, Foreign, Good and Broken seeds. Once after downloading and real data the data set must be encoded using a function, encoding of the data is done for the target function which is nothing but the label, these labels are called dependent variables.

The dataset is preprocessed, preprocessing is nothing but converting the images into one proper format, in our case we are resizing the image to 50x50 pixel. Then the data will be fed to the training part. Before data is trained the dataset must be split to training part and testing part. (0% of the data is considered as training data and the remaining 10% is taken test data. Then it was reshaped using TensorFlow. TensorFlow is one of the python libraries mainly used in deep learning techniques.

The CNN model is then trained for the dataset with the classes defined. The model will be generated by using a training dataset w.r.t label. The trained model will have images which has been encoded. To identify loss "categorical cross entropy has been used. Multiclass classification is used mainly for loss function; the model will get created once the dataset is trained. To reduce the loss Adam optimizer is used with learning rate set to 0. 0001.It is a method that computes adaptive learning rate for each parameter which combines the advantage of both the methods i.e., gradient descendent with momentum and RMS prop algorithm. The model has been trained with 100 epochs based on the predefined batch size of 30. Once the model has been trained it is ready to be classified. It detects if the selected image is Good Grain, Damaged grain, broken wheat, or it is a foreign particle and displays it on the GUI.

wheat seeds have been selected for the preparation of the dataset from Kaggle and own dataset has been taken. The chosen classes were damaged, Foreign, Good and Broken as listed in Table 1. For experimentation purposes, Table 1 shows distribution of the dataset.

Table 1. Details of different grains and their image count

Classes of seeds	**No. of seeds**
Damaged Grain	756
Broken Grain	327
Foreign Particles	715
Good Grains	927

All images were captured from a distance away to get good images. This process has been taken at daytime to avoid change in texture of the grains.

4. RESULTS AND INFERENCES

When python cnn.py file is being executed, training is being performed that is feature extraction is been performed for training data with 100 epochs based on the predefined batch size was 30. Which give the accuracy, validation accuracy and validation loss and the model will be created which contains the features of the images.

The front page will contain 3 buttons Browse, Quit and Clear in which the farmer can use to start the application. Where Browse button is used to select the image that we want to test, quit button will terminate the page and Clear button used to clear the image which we select for testing. When clicked on the browser button one more button is created which is a get photo once clicked on the button get photo, Folders will be opened for us to test the image for classification, we can select the picture we want to classify.

Once we select the image, Image will be loaded on the screen and analyze button will be displayed below the image selected once clicked on that the classification of the image happens and give us the accurate result. The snapshot shows the good grain image where the predicted output showed the status has Good Grain.

In figure 6, when damaged grain was selected for testing and clicked on the analyze image button, the status of the image was displayed damaged.

In figure 7, when broken grain was selected for testing, the status of the image was displayed accurately for broken grain. Testing was done for the foreign particle and the result was accurate.

After training and validation, we were able to get the desired result. The model is trained using training data and then validated using test data.

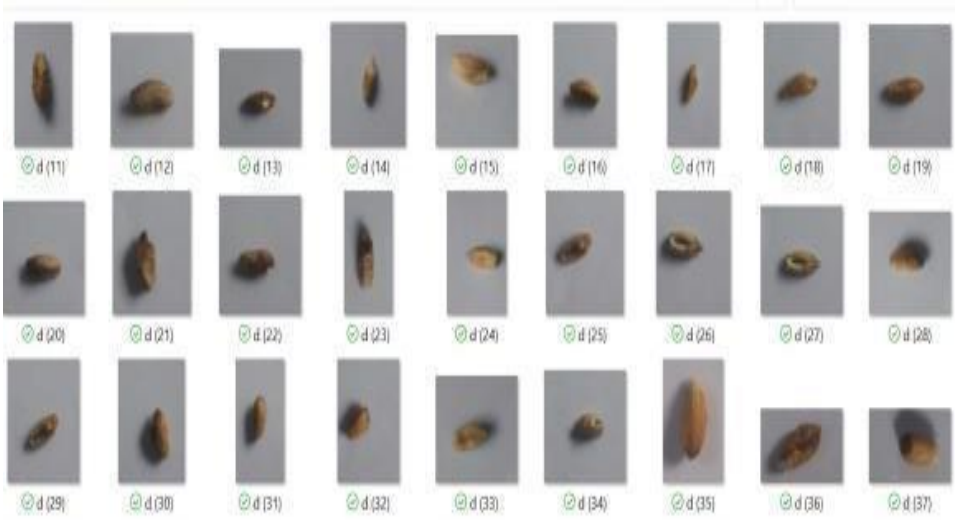

Figure 3: Samples of Seeds.

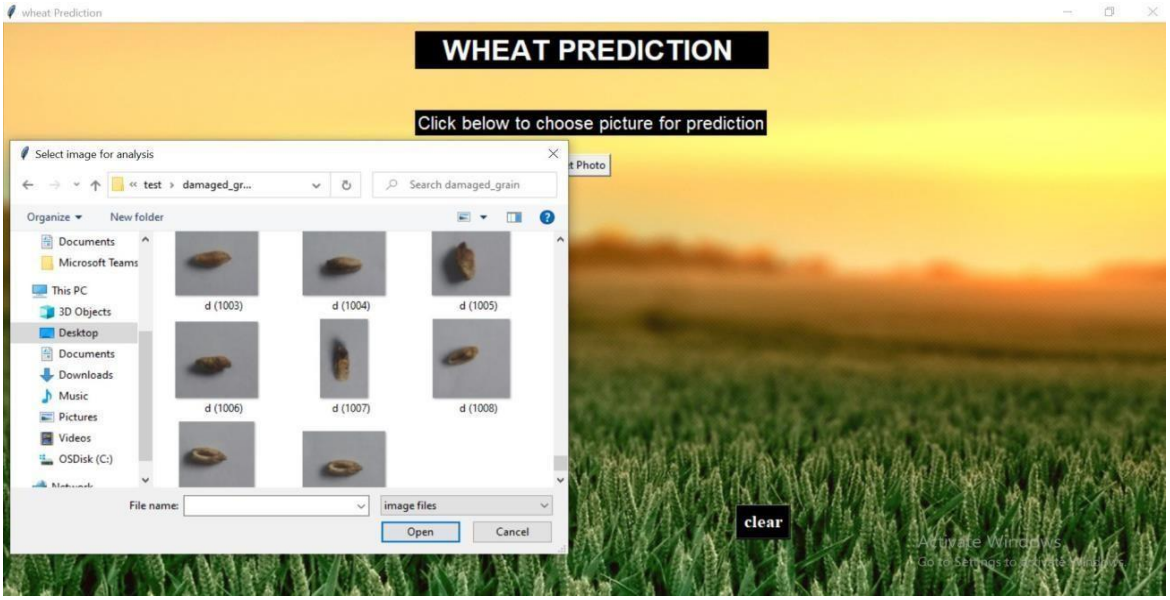

Figure 4: Shows the command to choose the appropriate image

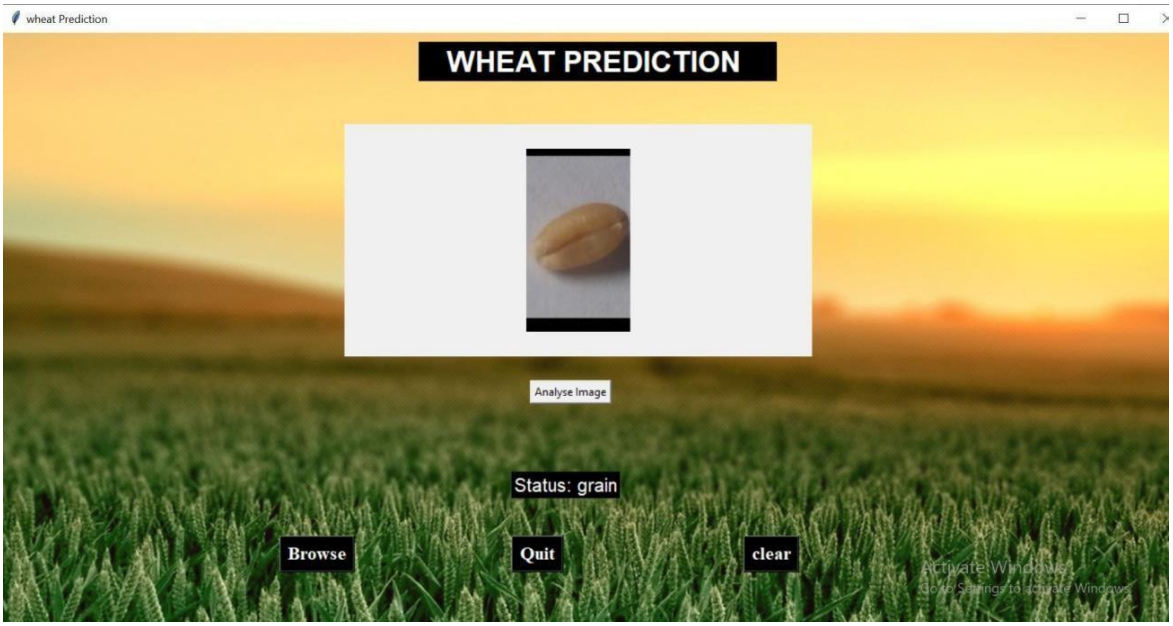

Figure 5: Classification result that is been obtained for test image – Seed is healthy.

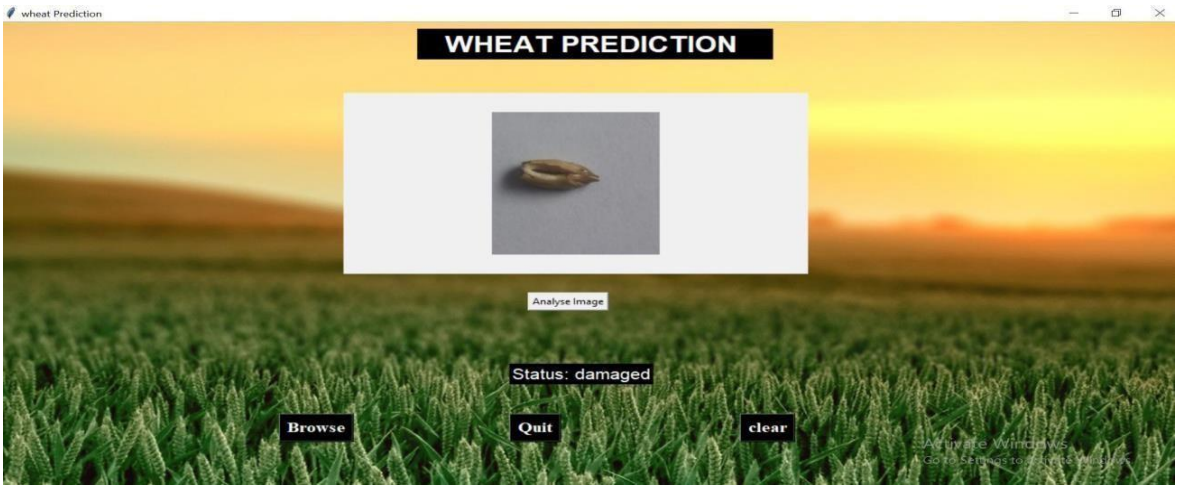

Figure 6: Classification result that is been obtained for test image – Seed is Damaged.

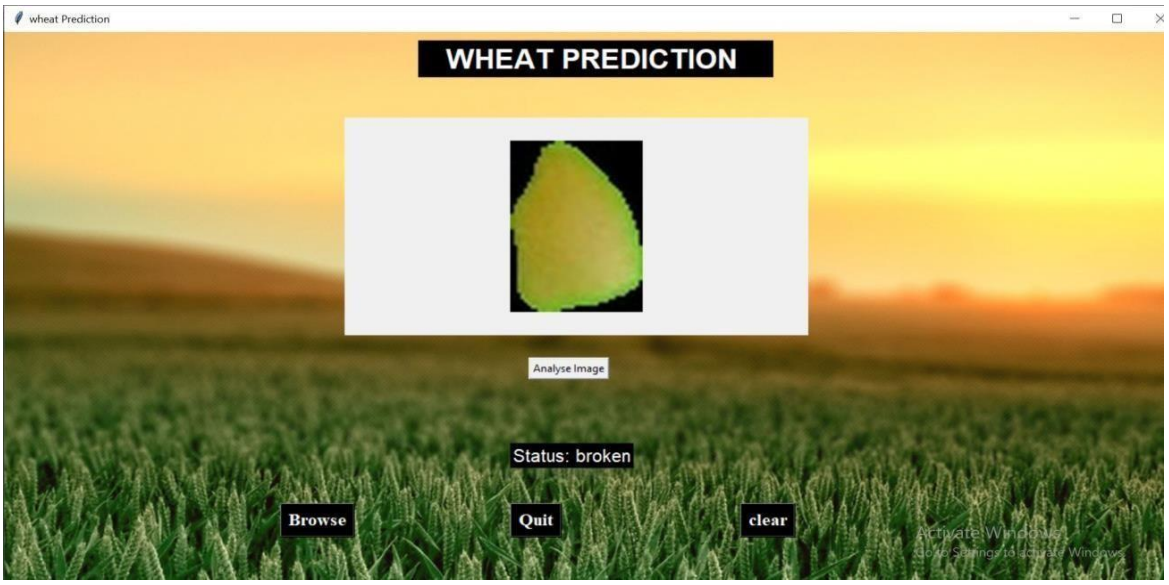

Figure 7: Classification result that is been obtained for test image – Seed is Broken.

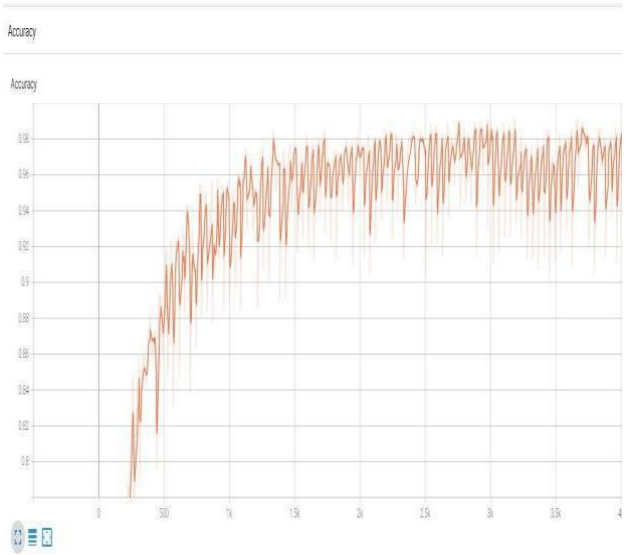

Figure 8: Training accuracy

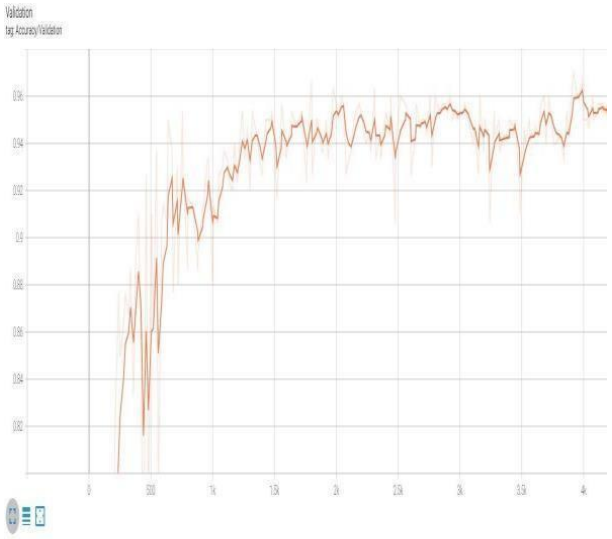

Figure 9: Validation accuracy

From figure 8 and 9, We can see that the training accuracy is 93% to 98 % which shows the high accuracy when compared to the other models used for classification of wheat and the validation accuracy is 94 % to 96% based on the training and validation accuracy.

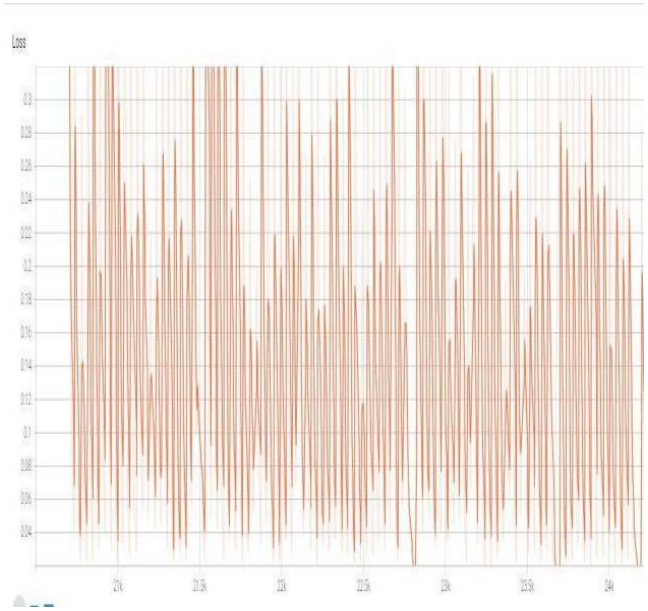

Figure 10: Model Loss

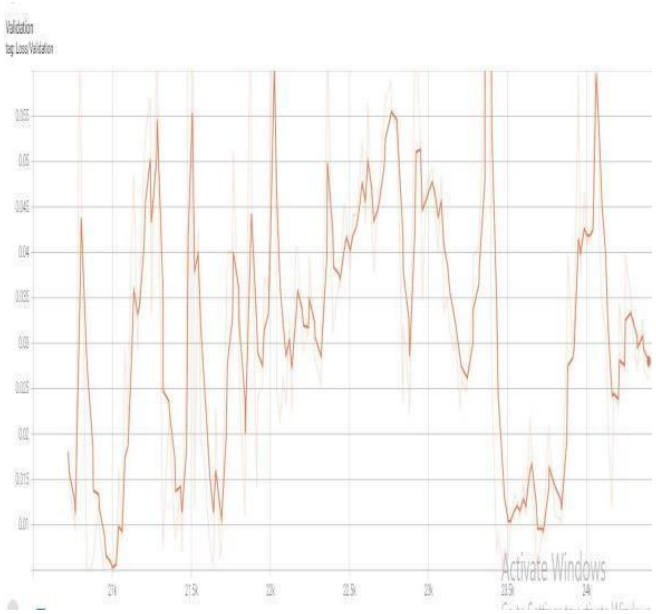

Figure 11: Validation loss

The Figure 10 shows the loss that happened during training which is below 0.3 and Figure 11 shows the loss for validation which is below 0.05. The training and validation loss is also less for our model.

5. Conclusion

The proposed system helps in identification of wheat analysis and grading of the wheat. The database obtained from the Kaggle and real data is properly segregated, and the different wheat species are identified and are renamed to form a proper database. Then, utilizing the training data, we train our classifier. We achieved good accuracy in predicting the classes of the different wheat grain. We employ Convolution Neural Network (CNN), which constitutes several layers that are utilized to make predictions. In our method we have 4 types of grains. The essential target was to recognize four classes of seed for enhancing wheat quality and supporting the wheat industrial chain's development and to prevent the yield losses and agricultural product quantity. Which in turn reduces the manual classification of wheat seed, reduces the amount of labor, and improves the quality of wheat seed. This work can be expanded by assembling more data, also to display the disease in the damaged wheat grain and its remedies, and detection can be done to various types of seeds.

The proposed system would help farmers or the cultivators in detecting damaged, broken, foreign particles and good grains at early phase using Image processing and Machine learning techniques, helps in taking preventive measures to control the loss before cultivating and this would significantly help farmers in terms of productivity, quality, quantity, low cost, efficiency, fast.

References

[1] Nabeel Ali Abdullah and Anas M. Quteishat, "Wheat Seeds Classification using Multi-Layer Perceptron Artificial Neural Network," *International Journal of Electronics Communication and Computer Engineering*, Vol. 6, pp. 306-309, 2015.

[2] Parminder Singh, Anand Nayyar, Simranjeet Singh and Avinash Kaur, "Classification of wheatseeds using image processing and fuzzy clustered random forest," *Int. J. Agricultural Resources, Governance and Ecology*, Vol. 16, pp. 23-156, 2020.

[3] Meesha Punn and Nidhi Bhalla, "Review on Classification of Wheat Grain Using Machine Algorithms," *International Journal of Scientific & Engineering Research*, Vol. 4, pp. 1248-1252,2013.

[4] Abhijit M Taley and Satish Bansod, "Performance analysis of Seed Grader through," International Journal of Engineering Research and Applications,Vol.2, pp.1135-1140, 2012.

[5] P R Kumar, C B Singh, Ashish Kumar Gupta and Kundan Kumar, "Quality status of wheat seed in Eastern India and the early vegetative growth of wheat (Triticum aestivum L.) in relation to size of seed,".

[6] S. K. Patil and S. V. Bansod, "Study the Effects of Some Parameters on Seed Grading in Gravity Separator," *International Journal of Engineering Research & Technology*, Vol. 3, pp. 2314-2328, 2014.

[7] WU Wei, YANG Tian-le, LI Rui, CHEN Chen, LIU Tao, ZHOU Kai, SUN Cheng-ming, LI Chun- yan, ZHU Xin-kai and GUO Wen-shan, "Detection and enumeration of wheat grains basedon a deep learning method under various scenarios and scales," *Journal of Integrative Agriculture*,pp. 1998–2008, 2020.

[8] G. Suseendran, E. Chandrasekaran, D. Akila and D. Balaganesh, "Automatic seed classificationby multi-layer neural network with spatial-feature extraction," *Journal of Critical Reviews*, Vol. 7, pp. 587-590, 2020.

[9] Radhika V Ronge and M M Sardeshmukh, "Indian wheat seed classification based on texture analysisusing ANN," *International Journal of Electrical, Electronics and Data communication, Vol 2, pp,96-97, 2014*

[10] 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.

[11] Agarwal and dhahiya Priyanka ; "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)

[12] Taylor, J.; Chiou, C.P.; Bond, L.J. A methodology for sorting haploid and diploid corn seed using terahertz time domain spectroscopy and machine learning. AIP Conf. Proc. 2019, 2102. [CrossRef]

[13] Lilik Sumaryanti, Aina Musdholifah, Sri Hartati: Digital Image Based Identification of RiceVariety Using Image Processing and Neural Network. (TELKOMNIKA Indonesian Journal of Electrical Engineering, ISSN: 2302-4046, Vol. 16, No. 1, October 2015).

[14] Kapadia VN*, Sasidharan N and Kalyanrao Patil: Seed Image Analysis and Its Application in Seed Science Research. (Juniper Publishers, Case Report, Volume 7 Issue 2, November 2017).

***