International Journal of Engineering and Advanced Technology (IJEAT)
ISSN: 2249-8958 (Online), Volume-11 Issue-1, October 2021

Classification of Grains and Quality Analysis using Deep Learning

Priyanshu Shrivastava, Karan Singh, Ashish Pancham

Abstract: ***There are various varieties of Rice and lentils. Price fabrication and adulteration have been some of the various issues faced by the consumers, farmers and wholesale retailers. Traditional methods for Identification of these similar types of grains and their quality analysis are crude and inaccurate. Methods were tried to implemented earlier but due to financial inability and low efficiency, they weren't successful. To overcome this problem, the project proposes a method that uses a machine learning technique for identification and quality analysis of these grains. Rice and Lentils which have the maximum consumption have been selected. Lentils are designated into classes based on colors. The technique of determining the elegance of a lentil is with the aid of seed coat shade. Red lentils can be confirmed through the cotyledon coloration. Lentil types may also have a huge variety of seed coat colors from inexperienced, red, speckled inexperienced, black and tan. The cotyledon colour may be red, yellow or inexperienced. The size and color of every Indian Lentil type (i.e. Red, Green, and Yellow, Black, White) are decided to be large or Medium or small, then size and colour end up part of the grade name. An smart machine is used to perceive the kind of Indian lentils from bulk samples. The proposed machine allows kernel length and coloration size using picture processing techniques. These Lentil size measurements, when combined with color attributes of the sample, classify three lentil varieties commonly grown in India with the highest accuracy. Rice is one of most consumed grains in India so its quality is of utmost importance. In this project, we identify and grade five types of rice and grade them with the help of their distinguished features such as size, color, shape, and surface. The project works in three phases viz., Feature Extraction, Training, and Testing. Various rice grain has a different shape, size, surface and various lentils come in different colors, Hence the feature that will be extracted is texture and colors. The method of regression will be adopted for the grading mechanism where the output will be in terms of percentage purity. The methodology for the extraction of the feature will be GLCM and Edge Detection where for supervised learning SVM and Back Propagation will be utilized. The project provides an efficient replacement for the traditional grading mechanism and standardizes the pricing of farm products based on their quality only.***

Keywords: Resnet, Fastai.

I. INTRODUCTION

Agriculture is the most crucial quarter of the Indian Economy. The Indian agriculture zone accounts for 18 percent of India's gross domestic product (GDP) and provides employment to 50% of the international locations body of workers. India is the world's largest producer of pulses, rice, wheat, spices, and spice products. The agricultural market nevertheless is still in bad form in rural India. In the absence of sound advertising centers, the farmers should rely on nearby investors and middlemen for the disposal in their farm produce that's bought at a throw-away charge. In most cases, these farmers are forced, under socio-economic conditions, to carry on the distress sale of their produce. In most small villages, the farmers sell their produce to the moneylender from whom they usually borrow money. According to an estimated 85 percent of wheat and 75 percent of oilseeds in Uttar Pradesh, 90 percent of Jute in West Bengal, 70 percent of oilseeds and 35 percent of cotton in Punjab is sold by farmers in the village itself. Such a scenario arises due to the incapacity of the terrible farmers to anticipate long after harvesting their vegetation. In order to fulfill his commitments and pay his debt, the negative farmer is forced to promote the produce at anything fee is offered to him. The Rural Credit Survey Report rightly remarked that the producers, in popular, promote their produce at an destructive vicinity and at an negative time and usually they get detrimental terms. According to the data provided by the Department of Economics and Statics (DES) the production of food grains for the year 2013-2014 is 264 million tons which are increased when compared to (2012-2013) 257million tons. This is a good symptom for the Indian economy from the agriculture sector. India remains among the main three as far as the production of different agricultural things like paddy, wheat, pulses, groundnut, rapeseeds, natural products, vegetables, sugarcane, tea, jute, cotton, tobacco leaves and so on. On the alternative hand, at the marketing front, Indian agribusiness is as yet confronting the troubles, as an instance, low stage of enterprise area reconciliation and integration, availability of reliable and handy information wanted by way of farmers on unique problems in farming.

Indian is an agriculture-based country, where more than 50% of the population depends on agriculture. This structure the main source of income. The commitment of agribusiness within the countrywide income in India is all of the greater, finally, it is said that agriculture in India is a backbone for the Indian Economy.

The contribution of agriculture in the initial two a long time in the direction of the whole country wide output is between forty-eight% and 60%. In the year 2001-2002, this contribution declined to just round 26%.

Manuscript received on October 18, 2021.
Revised Manuscript received on October 27, 2021.
Manuscript published on October 30, 2021.

*Correspondence Author

Priyanshu Shrivastava*, Department of Electronics and Telecommunication Engineering, Mumbai (M.H), India. Email: pshrivastava403@hotmail.com

Karan Singh, Department of Electronics and Telecommunication Engineering, Mumbai (M.H), India. E-mail: karansingh2687@gmail.com

Ashish Pancham, Department of Computer Engineering, Mumbai (M.H), India. Email: panchamashish@gmail.com

The combination Share of Agriculture and Allied Sectors, such as agribusiness, domesticated animals, and ranger provider and fishery sub-segments as some distance because the price of GDP is 13. Nine percent for the duration of 2013- 14 at 2004-05 costs. Agricultural exports constitute a 5th of the full exports of the us of a.

In perspective of the overwhelming role of the Agricultural Sector, collecting and help of Agricultural Statistics anticipate first rate importance. Most of the Indians are directly or circuitously depending on agriculture. Some are immediately connected to farming and some other people are worried in doing commercial enterprise with those items. India has the potential to provide food grains that could make a sizable difference within the Indian Economy.

In the absence of an organized marketing structure, private traders and middlemen dominate the advertising and buying and selling of agricultural produce. The remuneration of the services provided by the middlemen increases the load on the consumer, although the producer does not derive similar benefits. Many market surveys have revealed that middlemen take away about 48 percent of the price of rice, 52 percent of the price of groundnuts and 60 percent of the price of potatoes offered by consumers.

In order to store the farmer from the clutches of the money creditors and the middlemen, the authorities have come out with regulated markets. These markets generally introduce a gadget of competitive shopping for, assist in removing malpractices, make sure using standardized weights and measures and evolve appropriate machinery for settlement of disputes thereby making sure that the manufacturers are not subjected to exploitation and receive remunerative fees.

Even after the implementation of the regulations imposed by the government, there are hardly any efficient means for identification and grading of grains. For an efficient flow of the proposed regulation a proper identification and grading system is required. This project aims to be an efficient replacement for the traditional grading mechanism. It will be a tool for standardization of price for the products.

II. PROBLEM DEFINATION

There is no efficient tool available in the market to grade the grains, rice, lentils or the other products of the farm. The traditional method which is being practiced since decades is by human inspection. The human inspection comes with a lot of variables depending on factors such as knowledge, experience, alter motives, etc. This leads to unfair grounds for doing business. Hence, there is a need for developing an integrated tool for not only identifying the farm products but also grade them efficiently. This will help in maintaining the quality and provide a fair price for the products.
We propose to develop an integrated tool to identify and grade farm products using Image Processing and Neural Networks.

A. Objective & Motivation

As a rule, there's hardly ever any grading of the commodities to be advertised. Which is why the purchaser has little to none confidence in the quality of the products.

The British Government passed the Agricultural Produce (Grading and Marketing) Act in 1937 to solve this problem. But nothing really has happened. According to the Act, licenses are issued on a selective basis to dependable traders, under the supervision of the workers of Government.

The graded commodities are ultimately surpassed directly to the market below the label of “AGMARK”. However, due to the financial inability of farmers, the method could not be implemented on a large scale.

India, being an agrarian country, it is imperative that people in agriculture sector should be high on spirits. In order to maintain their spirits, the minimum requirement is to provide them the right price for their hard work. Hence, we are building a standardized tool that will grade the products efficiently and correspondingly provide the right price.

B. Scope

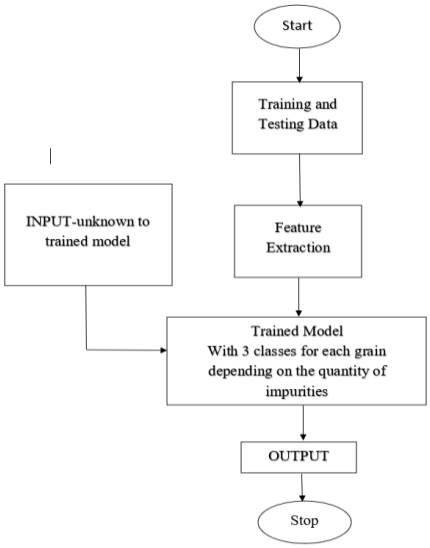
- Due to a lack of grading and standardization, the rural farmers are not privy to the effect of grading and standardization in their merchandise. Due to this unawareness, the cheating functionaries within the marketplace take undue gain of farmers. They may declare any produce as low-grade and quote a low price for such stocks, which in turn creates a big loss to farmers.
- Advancement in image processing and Machine Learning can make grading of the products completely automatic.
- By comparing the features of ideal products with the feature of given products an efficient grading tool can be developed. This project can be further used to check the condition of the crops and detect the disease if any, also it can provide the corresponding remedies.
- With more research, the health of the crops can be predicted which will, in turn, increase the yield.
- Furthermore, we can implement a web application where the users can get themselves registered. Farmers and retail sellers can upload pictures of their crops on hand and get a rating of the quality grade of their crop and according decide the selling price of their crop. The Web Application itself can get government determinant prices of crops using API, thus the sellers can have the idea about the price of their crop depending upon the quality of the crop.
- The consumers and buyers can upload live pictures of grains and lentils they are looking to buy and check if the seller’s price is reasonable for that particular crop.
- Registration details, login details and activity data of all users can be stored using RDBMS and SQL.
- The Web Application can collab with a fitness-based application to provide fitness related information (such as protein, carbs, fat, vitamins, etc.) for that particular picture uploaded crop, grain, lentils.

C. Application

- Farmers struggling to get the right prices for their crops can upload the picture of their crops and get the actual government determinant prices for their crops based on the quality of their crop
- Consumers and buyers can upload the picture of the grains and lentils they are looking to buy and check if the seller's price is reasonable for that particular crop.

III. PROPOSED WORK AND METHODOLOGY

Initially the untrained model will be feed with testing and training inputs after which the Model will train itself from the feature extracted. We shall understand the working of which model is can be chosen and its working in future topics After the model has been trained 3 classes will be created for each type of grain depending upon the quantity of impurities. These classes will help to determine type and quality of the grain. An Input unknown to trained model is given which will give us the desired output.



A. Creating Data Set

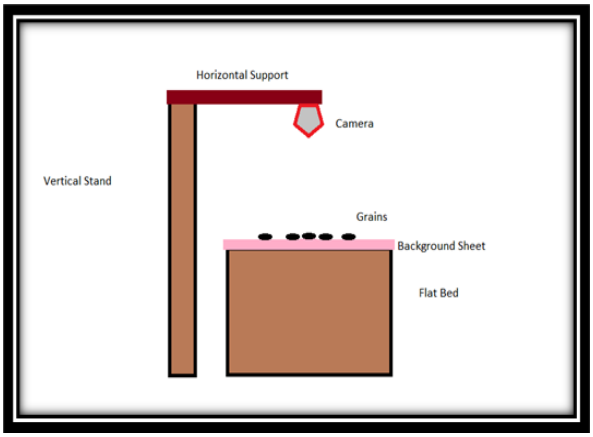
Due to the unavailability of the Data Set, a completely new data set will be created which will be project-oriented. There are 5 types of rice and 5 types of lentils to be identified and graded where each type will be allocated 100 photos which makes it a total of 1000 photographs. Out of the 100 photos allocated to every type 60 photos will be reserved for the training phase and remaining for the testing phase. The above task is given a timeline of 2 weeks.

The images captured will be of higher resolution which will affect the processing speed. In order to solve the issue and compare the efficiency of the output of Higher and Lower resolution images resizing of the image to lower resolution is necessary. The resizing task is allocated a timeline of 1 week

The inner working of the project consists of two-phase identification and grading. Performing various image processing methods for identification and various Machine Learning techniques for grading in order to get the highest efficiency based upon the Literature survey, it is allocated a timeline of one month. The project is proposed to be developed on Python where it will be used for developing codes, GUI, and debugging of the code. The prototype is expected to be ready within four months of the allocated time frame for the project and the remaining time will be utilized for improving the efficiency and making the product industry-ready.

B. Dataset Creation

The Dataset was created in a controlled environment for five different types of rice and five different types of Lentils. The different type of rice that were included for the dataset creation are Brown Basmati rice, Kolam Rice, Parmal Rice, White Basmati Rice & Wild Rice whereas in case of lentils Masoor dal, Moong dal, Toor dal, Udat dal(black), Udat dal(white) were used. The setup for creation of a dataset is as shown below in fig Canon EOS 90D was used to click the images for the dataset.



As seen above a vertical stand in combination with horizontal support was used to incline the camera such that the vision of the camera with the flat bed is perpendicular. Background was chosen either black or white depending upon the contrast of the grains i.e., for white basmati black background chosen whereas for black udat dal white background was chosen. Grains were scattered randomly from each batch of a grain on the background sheet. For grading purpose, like mentioned previously grains were scattered randomly but this time it included added impurities such as flakes & stones. Depending upon the added impurity the classes were labelled as High-Quality Grain, Moderate-Quality Grain, Low-Quality Grain.

C. Classification

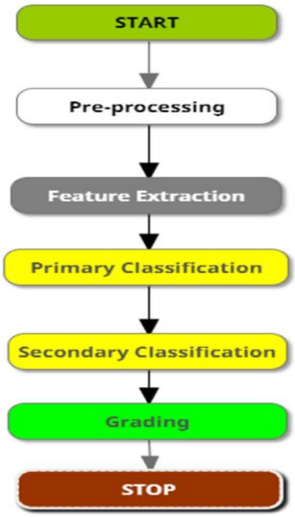
Initially the classification was done on Matlab with 83% accuracy but due to the low accuracy on classification purpose and limited number of extracted features the project was switched on python where the number of extracted features were comparatively very high.

D. Classification on Matlab

The project is divided in five phases which is mentioned below:

1. Developing Data Set
2. Pre-processing.
3. Contour Detection
4. Feature Extraction.
5. Classification and Grading

Published By:
Blue Eyes Intelligence Engineering and Sciences Publication (BEIESP)



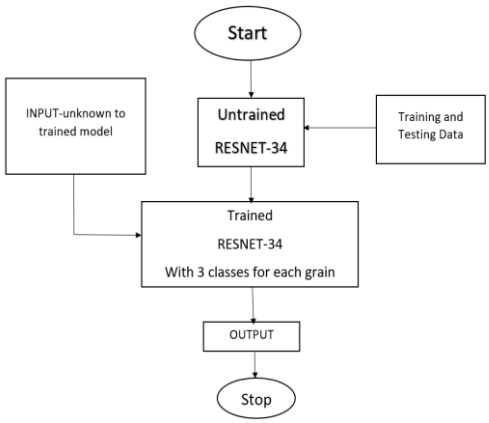
The image is first feed to the System where it performs thresholding by using either using Binarization or Otsu's thresholding. The noise in the image is removed using Gaussian noise filters. The contour is then detected by using Watershed algorithm. Concept of 8-chain connectivity is used for obtaining features which then is formulated for classification and Grading. Classification is divided as Primary and Secondary where in the Primary the system will classify between rice and lentils and in secondary it will classify between either type of rice or type of lentil.

E. Classification on Python

The project is divided in five phases which is mentioned below:

1. Developing Data Set
2. Separation of Data Set into Train and Test Set
3. Train and Test RESNET model for Classification and Grading
4. Input (unknown to the trained model)
5. Classified and Graded Output

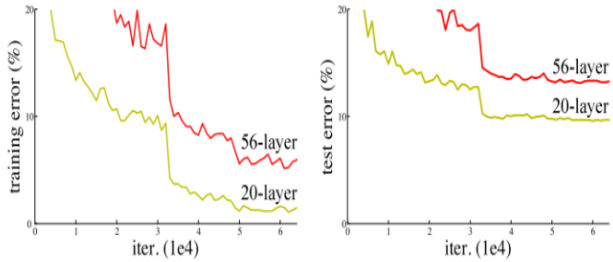
1. Flowchart



Initially the untrained model will be feed with testing and training inputs after which the RESNET-34 will train itself from the feature extracted. We shall understand the working of RESNET-34 in future topics After the model has been trained 3 classes will be created for each type of grain depending upon the quantity of impurities. These classes will help to determine type and quality of the grain. An Input unknown to trained model is given which will give us the desired output

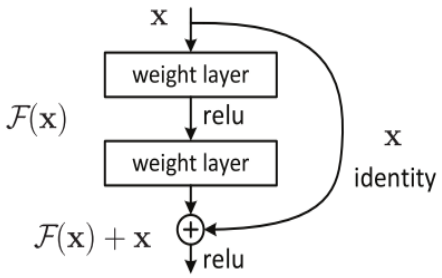
F. Resnet

RESNET is an open-sourced Residual Neural Network model developed by Stanford University in collaboration with ImageNet. It works on the principle of Deep Convolutional Neural Network DCNN with additional layers upon layers added depending upon the application and efficiency required. RESNET was the successor to the AlexNet which only used 8 neural network layers of which 5 were Convolutional and remaining fully connected. Much of the success of Deep Neural Networks has been accredited to these additional layers.[10] However, increasing network depth does not work by simply stacking layers together. Deep networks are tough to educate due to the notorious vanishing gradient problem — as the gradient is back-propagated to in advance layers, repeated multiplication might also make the gradient infinitively small. As a result, as the network goes deeper, its performance gets saturated or even begins degrading unexpectedly.



increasing network depth leads to worse performance

The core idea of RESNET is introducing a so-called "identity shortcut connection" that skips one or more layers, as shown in the following figure:



a residual block

G. Grading on Python

The Grading on Python will be performed on the same trained model which was used for classification.

Here there will be a greater number of classes because each class created in the classification process is further sub divided into three more class which are labelled as High-Quality Grain, Moderate-Quality Grain, Low-Quality Grain classes. For example, in White Basmati there will be three more classes namely High-Quality Grain White Basmati, Moderate-Quality Grain White Basmati, Low-Quality Grain White Basmati. Here the High, Moderate, Low Quality indicates the number of impurities added in the grain samples. Higher the Quality lower the number of Impurities. This makes the confusion matrix of the size 30x30.

IV. RESULT

Before you begin to format your paper, first write and save the content as a separate text file. Complete all content and organizational editing before formatting. Please note sections A-D below for more information on proofreading, spelling and grammar. Keep your text and graphic files separate until after the text has been formatted and styled. Do not use hard tabs, and limit use of hard returns to only one return at the end of a paragraph.
Do not add any kind of pagination anywhere in the paper. Do not number text heads-the template will do that for you.

A. Results from Execution on Matlab

Here the testing was performed on White Basmati rice and Toor Dal.8-Chain connectivity was implemented in Matlab and features such as Perimeter, Area, Solidity, filled area, Centroid was obtained. Circularity is a logic formulated from the above features so as to perform Primary classification i.e. between rice and lentils.

Circularity = 〖(Perimeter)〗^2/(4*π*(filled area))

Samples used:


Results on processing:

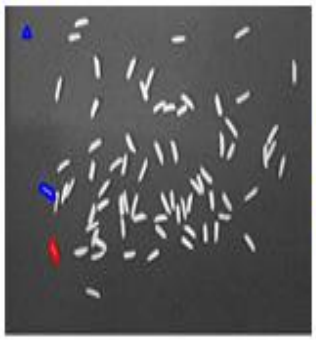
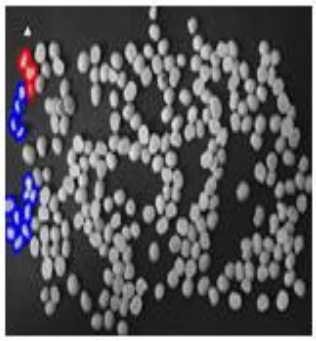
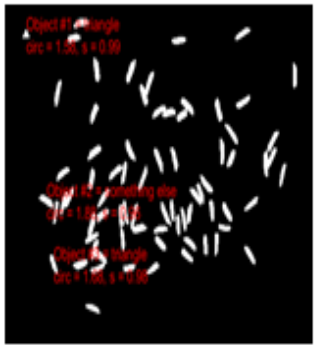
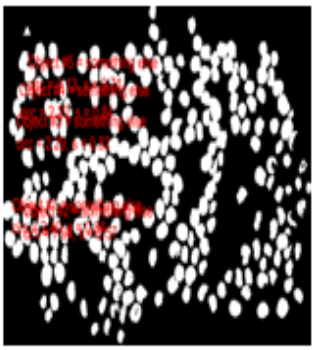





Table i. Extracted Features for Rice

RICE					
	AREA	FILLED AREA	SOLIDITY	PERIMETER	CIRCULARITY
	16308	16308	0.960	620.257	1.877
	12322	12322	0.97	509.669	1.677
	11376	11376	0.940	546.946	2.09
	16182	16182	0.967	620.46	1.89
	13323	13325	0.9	556.09	1.84
	13874	13874	0.95	525.941	1.586
	15590	15590	0.970	575.888	1.691
	16227	16228	0.959	622.084	1.89
	12494	12494	0.96	559.401	1.993
	16160	16160	0.962	586.408	1.69
	14317	14317	0.976	527.603	1.54
	14844	14844	0.963	557.902	1.668
	17002	17002	0.963	589.399	1.62
	15065	15065	0.97	543.684	1.561
	13892	13892	0.977	569.005	1.854
	14683	14683	0.945	596.828	1.93
	12182	12182	0.972	535.666	1.87
	13192	13192	0.9733	550.81	1.8
	12574	12574	0.966	503.669	1.60
	16155	16155	0.9486	579.273	1.6
AVG	14388	14388.25	0.962	563.84915	1.77

Table ii. Extracted Features for Lentils

LENTILS					
	AREA	FILLED AREA	SOLIDITY	PERIMETER	CIRCULARITY
	16308	16308	0.960	620.257	1.87
	12322	12322	0.9757	509.669	1.67
	11376	11376	0.940	546.946	2.0
	16182	16182	0.96	620.46	1.89
	13323	13325	0.94	556.09	1.84
	13874	13874	0.953	525.941	1.58
	15590	15590	0.97	575.888	1.69
	16227	16228	0.95	622.084	1.8
	12494	12494	0.967	559.401	1.99
	16160	16160	0.962	586.408	1.69
	14317	14317	0.976	527.603	1.54
	14844	14844	0.963	557.902	1.66
	17002	17002	0.963	589.399	1.62
	15065	15065	0.973	543.684	1.56
	13892	13892	0.977	569.005	1.85
	14683	14683	0.945	596.828	1.93
	12182	12182	0.972	535.666	1.87
	13192	13192	0.973	550.81	1.83
	12574	12574	0.966	503.669	1.6
	16155	16155	0.948	579.273	1.6
AVG	14388.1	14388.25	0.962	563.84	1.77

When 8-Chain Codes were implemented on white basmati and red lentils it was observed that circularity can be a differentiating factor since the average circularity of Lentils was found to be 1.7 and that was rice was 1.14.

A threshold of 1.56 was set for classification purpose. But the above threshold was able to successfully identify either the grains as rice or lentils but failed to determine which type of lentil or rice. Hence due to the limited extracted feature and failure to separate the overlapped grains from each other the classification accuracy was limited to 83%

B. Results from Execution on Python

The RESNET50 trained model was able to classify with an accuracy of 98%. The iteration performed by the trained model on the data set is mentioned in the below table.

epoch	train_loss	valid_loss	accuracy	time
0	2.278574	0.500827	0.788235	05:11
1	1.159382	0.070934	0.976471	04:56
2	0.771946	0.050056	0.988235	04:46
3	0.576763	0.047852	0.988235	04:46

Figure i. Accuracy of RESNET-50

But it was observed that the model was confused between parmal rice and white basmati rice as evident from the below confusion matrix By reducing the layers in the RESNET model we can improve the performance characteristics such as reduced training time and higher accuracy. By reducing the layers, the complexity in CNN is reduced as we have a smaller number of classes which require less feature to be examined. The RESNET34 model gives near to perfect performance characteristics as it can be observed below.

```
learn.fit_one_cycle(4)
```

epoch	train_loss	valid_loss	accuracy	time
0	1.473831	0.029502	0.992806	06:39
1	0.666485	0.049584	0.978417	05:37
2	0.401792	0.012975	1.000000	05:56
3	0.242544	0.009053	1.000000	05:35

Figure ii. Accuracy of RESNET-34

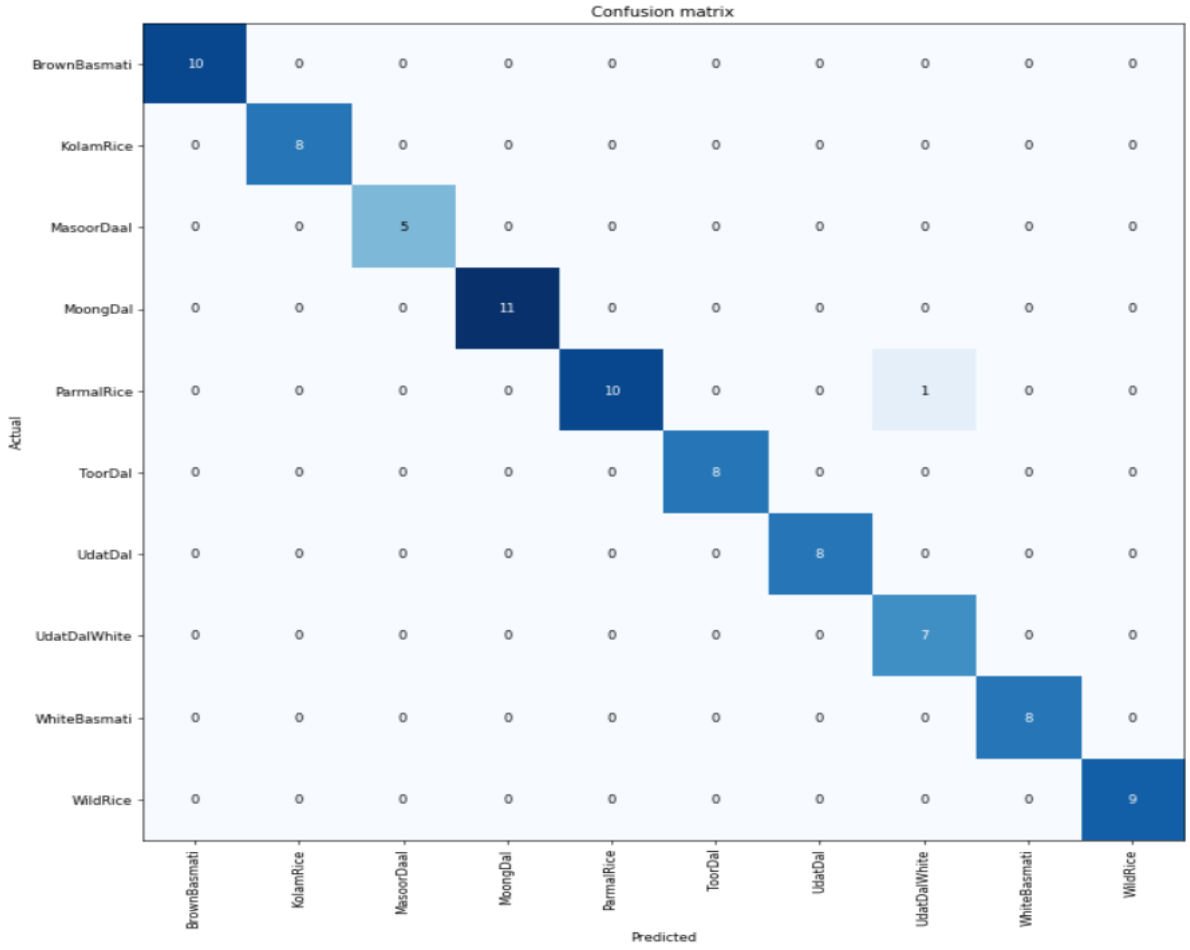


Figure iii. Random results from implementation on resnet-50

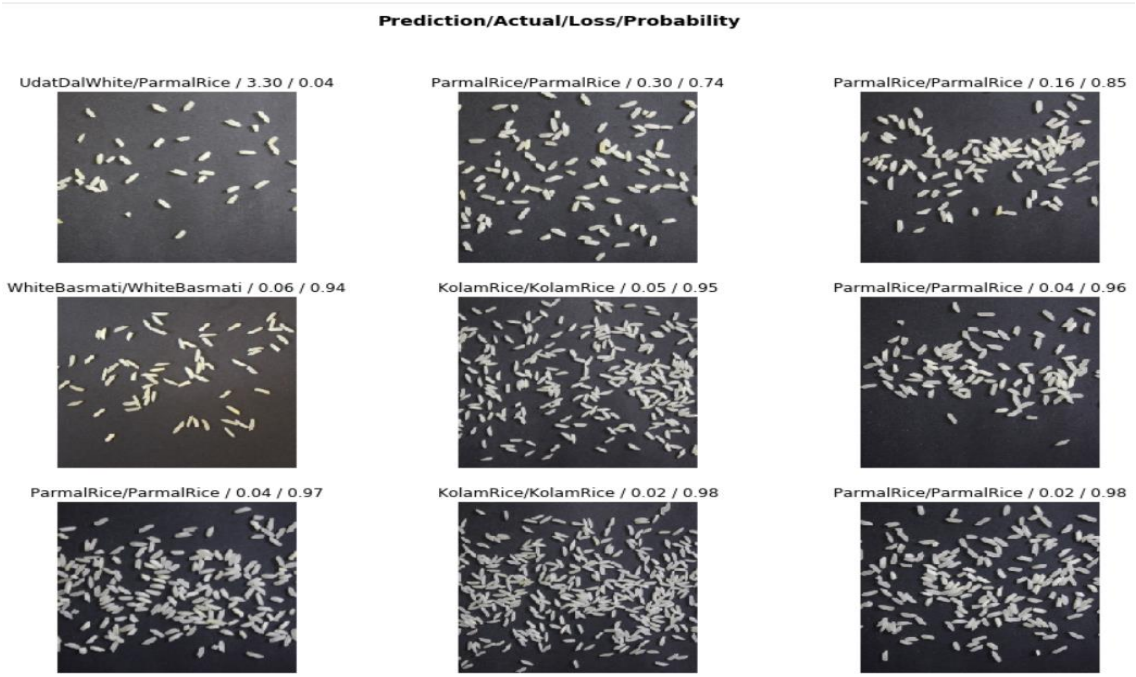


Figure iv. Confusion Matrix

V. CONCLUSION

With a really optimum accuracy and high predictability of the type of grain and its quality, it will enhance quality improvement and also fair sale of the food crops produced by farmers. Implementing the classification on both MATLAB as well as Python where it was observed python had additional tools and libraries for better visualization so the project advanced on Python. With the help of open sourced RESNET model we implemented classification of 10 classes of grains and it was inferred that RESNET34 model had better Performance Characteristics than RESNET50. With the help of the pre-trained deep neural resnet model, high accuracy of 98%-100% was achieved. Not sticking to just 10 classes (10 different types of food grain) we can add more food grains or even crops, in general, to make classification and grading more comprehensive.

REFERENCES

1. "India at a glance | FAO in India | Food, and Agriculture Organization of the United Nations." [Online]. Available: http://www.fao.org/india/fao-in-india/india-at-aglance/en/. [Accessed: 15-Nov-2019].
2. N. Nagoda and L. Ranathunga, "Rice Sample Segmentation and Classification Using Image Processing and Support Vector Machine," 2018 13th Int. Conf. Ind. Inf. Syst. ICIIS 2018 - Proc., no. 978, pp. 179–184, 2019.
3. P. Wijerathna and L. Ranathunga, "Rice Category Identification using Heuristic Feature Guided Machine Thresholding (Binarization or Otsu's Thresholding) Use the concept of chain codes to obtain the connectivity of the edges and find the physical features and use HUE color extraction to find the color features of the image Apply Noise Filters Use watershed algorithm to find the contour of the grains Based on the results of features extracted formulate a logic to classify and grade the images Vision Approach," 2018 13th Int. Conf. Ind. Inf. Syst. ICIIS 2018 - Proc., no. 978, pp. 185–190, 2019.
4. M. J. Asif, T. Shahbaz, S. Tahir Hussain Rizvi, and S. Iqbal, "Rice Grain Identification and Quality Analysis using Image Processing based on Principal Component Analysis," RAEE 2018 - Int. Symp. Recent Adv. Electr. Eng., pp. 1–6, 2019.
5. D. Vishnu, G. Mukherjee, and A. Chatterjee, "A computer vision approach for grade identification of rice bran," Proc. - 2017 3rd IEEE Int. Conf. Res. Comput. Intell. Commun. Networks, ICRCICN 2017, vol. 2017-Decem, pp. 10–14, 2017.
6. N. V G, "An Intelligent system for identification of Indian Lentil types using Artificial Neural Network (BPNN)," IOSR J. Comput. Eng., vol. 15, no. 5, pp. 54–60, 2013.
7. "Agricultural Marketing in India: Defects and their Remedial Measures." [Online]. Available: http://www.economicsdiscussion.net/agriculture/marketing/agricultural-marketing-in-india-defects-and-their-remedialmeasures/12854. [Accessed: 15-Nov-2019].
8. P. Mohanaiah, P. Sathyanarayana, and L. Gurukumar, "Image Texture Feature Extraction Using GLCM Approach," Int. J. Sci. Res. Publ., 2013. [9] B. S. Prajapati, V. K. Dabhi, and H. B. Prajapati, "A survey on detection and classification of cotton leaf diseases," Int. Conf. Electr. Electron. Optim. Tech. ICEEOT 2016, pp. 2499–2506, 2016.

AUTHORS PROFILE

Priyanshu Shrivastava is a B.E Electronics and Telecommunication Engineering 2020 graduate. Deep Learning specialization by DeepLearning.AI, have completed Mathematics of Machine learning specialization by Imperial College of London. The areas of interests are Machine Learning, Data Analytics, Deep Learning. E-mail: pshrivastava403@hotmail.com

Karan Singh is a B.E Electronics and Telecommunication Engineering 2020 graduate. Currently an Associate Software Engineer. Developing Application and performing Account provisioning for multiple applications of the client. Independently performing user's request automation using Python and Machine Learning which will significantly increase the Team's efficiency. The areas of interests are Machine Learning, Data Analytics, Deep Learning, Cloud Computing. E-mail: karansingh2687@gmail.com

Ashish Pancham is a B.E Computer Engineering 2020 graduate. Courses in Stack Development, React, Angular, NLP. The areas of interests are Machine Learning, Data Analytics, Deep Learning, Full Stack Development, Cloud Computing, E-mail: panchamashish@gmail.com