Volume 9, Special Issue No. 1, February 2018

ISBN: 978-93-5300-455-2

International Journal of Advanced Research in Computer Science
(ISSN: 0976-5697)
RESEARCH PAPER

ICRACC 2018
(23rd–24th February 2018)

Available Online at www.ijarcs.info

A Method for Classifying the Germination of Green Gram Image using Neural Networks

Harish S Gujjar
Asst. Professor and Head, Dept. of Computer Science, SSA GFGC Ballari, Karnataka, India
gujjarh@gmail.com

ABSTRACT:

This paper deals with a computer vision system based on machine learning techniques in the field of image processing and germination of the green gram is spontaneously assessed the rate of germination. The germination test is most important and trusted method to determine the speed and successfulness of germination. It gives us the important information regarding the successful of converting the germinated seed into plant. On a whole all green grams are not able to germinate. In this paper we use Artificial Neural Network (ANN) which uses Multilayer Perception structures are used. In between 30 ^{0}C to 40^{0}C almost all the seed germination takes place. On an average only 90% of the seeds can germinate. This work is able to classify accurately of 96% of germinated seeds.The Green gram samples are collected from APMC in Ballari districts of Karnataka for the growing year 2018.

Keywords: artificial neural networks, seeds, germination, green gram, image processing

INTRODUCTION

To increase the growth and supply of money is based on the quality of the seed which is the most fundamental part of any agriculture. Earlier, the detection of many physical and interpretations of seeds was focused, But today a large variety of techniques are available for seed testing green gram seed assessment and judgment which correlate well with certain vigor and germination parameters (McDonald 1998). In the present context seed testing is carried out in scientific experiment by expert humans. The experiments are sketched to assess the quality of the seed. Large number of tests is carried out. For example, the highest germination capability can be found out by the germination test or successful of germination of seed. The germination speed of certain seed bulk seeds is an important sign which represents the seed performances in the farm and it is indicated as percentage (for example 89% germination pace means out of 100 , 89 seeds are expected to germinate subjected to the good growth atmosphere). The above measure is important for estimating maximal seeding speed and to estimate whether a certain seed lot has the ability to process a quality crop.

From then to now the manual counting is tedious and takes lot of time and human resource, the efficiency of the process can be boosted by very large number of ways. On the basis of Machine learning and Image processing and with the help of Computer vision system, we can design a machine which tests the approximation of germination of seeds.

Analysis of images was introduced in the area of seed technology by Howarth and Stanwood (1994) those are created a database of color images to identify the variation of both environmental and genetic phenomena.

In the field of seed Identification and classification in the field of image processing it has a very good result (Uchigasaki et al. 2000, Granitto et al. 2002) and judging the germination (McDonald et al. 1998). Dell' Aquila et al. (2000) he has taken analysis of image to signalized the objective of seeds of White cabbage while Geneve and Kester (2001) assessed size of the seeding following germination by digital image analysis which are assesed by computer by using scanner (Ducournau et al. 2004). Urena et al. (2001) suggested a Machine vision system which included a process of gathering data by using a system with a technique of logic based fuzzy to assess the quality of germination. Ducournau et al. (2004) produced a Machine vision system modeled to count the number of radical tips which were appeared on the seeds below temperature, lighting and measurement of moisture in air conditions. A mechanical process system which works on the algorithm which has an ability to count the seeds which are germinated and give the average germination time on the point of difference between the two successive pictures.

An image processing and computer vision uses famous and renowned software called MATALAB. A digital camera of Nikon D810 FX 36.3MP Digital SLR (Single Lens Reflex) Camera sigma 18 -200 mm lens zooming. The camera is fixed to a iron stand having movement towards vertical which gives a firm support. The camera was fixed at a distance of 450 mm. The images were having a resolution of 3873 X 2593 pixels, a horizontal and vertical resolution of 300 dpi respectively having a bit dept of 24. A 210 mm diameter white light of 22 Watts fluorescent tube was fixed which glown at a voltage of 230 volts was placed over the cell culture dish having the seed samples placed on a white tissue paper. A 270 mm diameter semi spherical steel utensil is used to cover bulb to get the diffused light to eliminate the impact of external factors (Figure 1). The images thus obtained from the above are transferred from digital camera to PC (quad core processor having 4 GB RAM) through a USB cable.

IMAGE AQUISITION

The samples of green grams are bought from the APMC from the areas of ballari districts of Karnataka, 2018 as the growing year. At the start of the experiment, the green grams are preserved for 30 days in a controlled air and temperature at around 4 C°, the seed is moisturized to balance the relative humidity at 50%. The 700 seed samples are randomly chosen from the bag. Further, a white filter paper is put to a cell cultured dish which is moisturized with a 3ml of pure water. The optimal contrast between the seed was minimized by using the white filter paper. An approximately 25 seeds were spread over the top of the misted filter paper in every dish and was positioned at a relatively similar distances. The dish was inculpated with cover. The germination of seeds were carried out in an well directed condition and in very low light at 20 to 30 °C and relative humidity at 75% in an incubator invented by Jacobsen. In a day 8 hours were used to illuminate the seeds.

A cool white fluorescent light of 750 flux was used to provided. The seizing of the images was carried out by a digital camera with Nikon D810 FX 36.3MP Digital SLR (Single Lens Reflex) Camera sigma 18 -200 mm lens zooming. The camera is fixed to a iron stand having movement towards vertical which gives a firm support. The camera was fixed at a distance of 450 mm. The images were having a resolution of 3873 X 2593 pixels, a horizontal and vertical resolution of 300 dpi respectively having a bit dept of 24. A 210 mm diameter white light of 22 Watts fluorescent tube was fixed which glows at a voltage of 230 volts was placed over the cell culture dish having the seed samples placed on a white tissue paper. A 270 mm diameter semi spherical steel utensil is used to cover bulb to get the diffused light to eliminate the impact of external factors(Figure 1). The images thus obtained from the above are transferred from digital camera to PC (quad core processor having 4 GB RAM) through a USB cable.

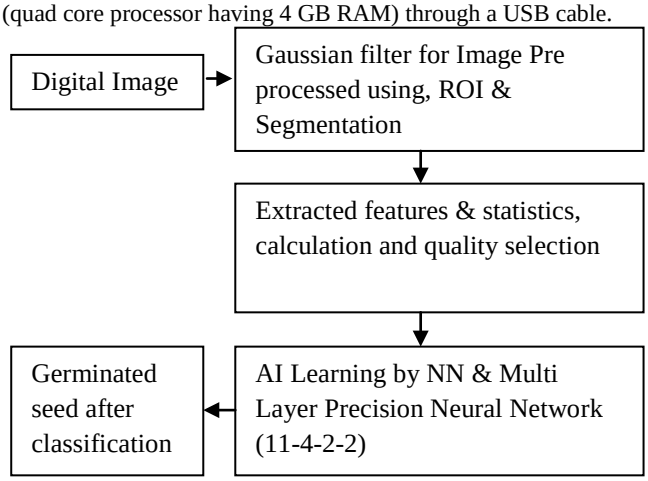


Figure1: Proposed methodology.

PRE PROCESSING OF AN IMAGE

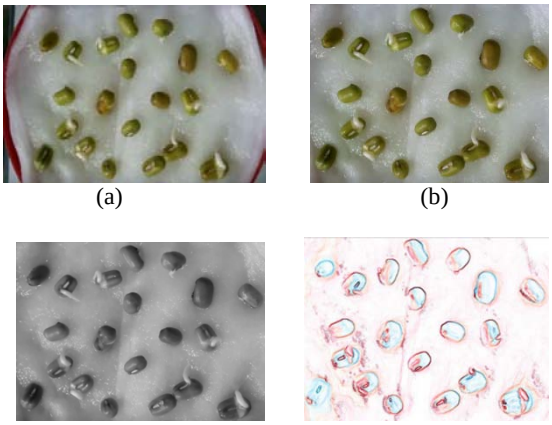
Firstly, original RGB images were used to extract the features with the help of MATLAB which helps in image processing (Figure 2a). The individual images thus received were cropped by a fixed known radius in the region of interest (ROI) (Figure 2b). To efficiently manipulate all the images, by cropping of images the size of the image was reduced which is helpful in manipulations. The original image has a size of 3873 X 2593 pixels was diminished to 1854 X 1836 pixels. Secondly, the images were smoothed with a σ limited at 2 which uses a Gaussian filter (Eq. 1) drawn by Rasband (2008):

$$\varphi_{\mu,\sigma^2}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{(x-\mu^2)}{2\sigma^2}} \qquad (1)$$

Where, the pixel intensity is represented as *X*, mean is indicated μ , Standard deviation is notified as σ, variance is signified as σ^2 =3.18 and **e**=2.718. The conversion of the image from color space RGB to an 8-bit image of a gray scale (Figure 2c) and was ultimately converted by thresholding into a binary image (Figure 2d). The predefined testing of the thresholding were carried out by applying the lower limit as 85 and the highest limit as 255grayscales.

The maximum and minimum size of a particle was set to 1500 to 14,000 pixels respectively. By doing this we can just ignore the smaller and larger areas which cannot be considered as seeds. Lastly the seed was extracted from the background and was named by a device itself. The external tracing of the boundaries of the seeds was considered as yellow (Figure 2e).

The minimum particle size was set to 1500 pixels and the maximum to 14.000 pixels. With this we additionally avoided the smaller and the bigger areas that could not be accounted as seeds. Finally, presented seeds were separated from the background and automatically labeled with the integer. The external perimeter of the seed was traced in yellow .



(a) (b) (c) (d)

Figure2: (a) Original color image of RGB, (b) Image is cropped, (c) 8-bit grey scale image, (d) Binary image after thresholding.

After the Thresholding and Labeling of the seeds the upcoming step is to examine the particles which were labeled. The detailed description of the steps in image processing is shown in Table 1.

The culmination of the different features which was extracted together from different Petri dish was tabulated, where row represents the seed and the column represents the individual parameter of a single seed. For the examination by expert a 28 different images and 25 different seeds was bestowed for germinated and non germinated classification over a 700 seeds. These results were also included in the table.

PROPOSED METHODOLOGY

Further the analysis was carried out by software which uses machine learning called WEKA (Waikato Environment for Knowledge Analysis), A formatted file is created using *csv* (comma-separated values) and send once the feature characteristics are generated. WEKA is a data mining and machine learning tool which consists of large number of algorithms for operating by means of a program, categorization, Measurement of relation between the mean and variance, grouping the number of things of same kind, amalgamation rules and representation of objects.

The information identification from main data was taken from main streams of agricultural, to do the above the University of Waikato, New Zealand proposed WEKA and further it was used for other fields also. Initially it was available i Java version and was translated to programming in MATALAB.

The amalgamation of the two tools was primarily described by Mayo et al. (2007) in this method feature vectors were used for classification of months by using species for drawing out of the each month of the image. Further choosing and grading was carried out by WEKA. The assessment process used info grain attribute evaluator which was used to estimate by attributing the grade by grading and rating by a single assessment. The edge filtering and small mean from the outer length of the closed boundary is the largest result of classification which is the gross value of each and every pixel in the classified region of pixels (Table 1).

Parameter	Description
Perimeter	The end to end are of the outside area is called a Perimeter.
Kurtosis	The measurement of the degree of curved arck is known as Kurtosis.
Max	The highest value of grey level within the selected boundary.
Skewness	The lack of equality of distribution of measure of the degree is considered as Skewness.
Standard Deviation	The average grey value is obtained by the grey values of standard deviation.
Major	Major is the primary axis of the best fitting ellipse.
Area	A four sided are of pixels is called the selected area.
Mode	The largely repeated value in the area of grey mode level. It refers to the top most point in the histograph.
Median	The median value is the average of two middle values of pixels in the particular area.
Minor	Minor is the subsidiary axis of the prime suited ellipse.
Mean	It is the central value of grey inside the selected area. It is the summation of the gray values in the wanted area which is divided by the number of pixels.

Table 1: Measured Parameters.

The 10 sets of data were used for learning of modules of classification, each having 70 occurrences. Then, 9 sets of these were used for training and the testing was carried out on the remaining sets (training seeds of 630 and testing of 70 seeds for each execution. The above process was successive used for 10 times on the structure thus designed.

CLASSIFICATION MODEL

The best method for categorization is Artificial Neural networks (ANN) with multilayer perceptions model (MPL) was used and compared with the human made counting. The back propagation algorithm model was used for training. At 0.3 was the value set for rate of learning and 0.2 as momentum rate. The input layer has 11 neurons and where as the output layer has w neurons as the number of features and classes was 11 and 2 respectively. Inside the ANN there are large number of hidden layers and neurons which were trained at the speed of 100 to 2000 epochs.

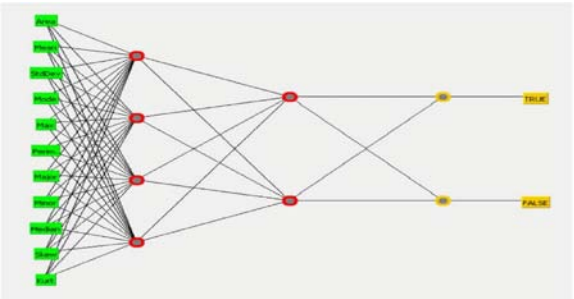


Figure3: Model of Artificial Neural Network

PERFORMANCE EVALUATION

To judge the performance of the ANN, we have cauterization accuracy, correctness, recollect and F-calculate which are extracted from confusion of matrix. Accuracy of classification is the process of the judgment of the accuracy of the model of neural network in comparison made with the human beings and their correctness over 25 data sets. The calculation is carried out by the number of accurately categorized occurrences which were slashed by the number of occurrences.

$$\text{Accuracy} = \frac{\sum(TN,TP)}{\sum(TP,FP,TN,FN)} * 100\ \% \qquad (2)$$

where true positive is abbreviated as TP, True Negative is abbreviated as TN, FP & FN is unfolded as false positive false negative respectively. The summation of TP+TN+FP+FN is the total number of instances occurred in the test set TN + TP is the accurate count cauterization of instances (Witten and Frank 2005).

In this paper TP is constitute as real seeds of germination which were also assumed as germinated seeds. TN is not constitute as real seeds of germination which were also assumed as non germinated seeds. The actual not germinated seeds are notified as FP and FN are assumed as not germinated and were actually germinated.

Precision is the amount of instances predicted positively which are positive which were really assumed as positively among the total which is estimated as follows :

$$p = \frac{TP}{\sum(TP,FP)} \qquad (3)$$

where, TP is abbreviated as true positive and FP is abbreviated as false positive. A FP happens when the class is incorrectly estimated as positive when it is actually obstructive (Witten and Frank 2005).

model	Accuracy		Precision	
	Mean	Std. dev	Mean	Std. dev.
ANN	95.99	3.17	0.995	0.0344

model	Recall		F-measure	
	Mean	Std. dev.	Mean	Std. dev.
ANN	0.9955	0.0200	0.959	0.0264

Table 2: Calculated Performance for cauterization

Recall is the summation of the TP and TF which is explained as true positive to summation. It is sometimes named as sensitivity in different areas. It is measured as follows:

$$\text{recall} = \frac{TP}{\sum(TP,FN)} \qquad (4)$$

FN is incurred when approximation is incurred as negative when it is really positive (Baeza and Riberio 1999, Witten and Frank 2005).

F-measure is stated as the recall and harmonic mean of precision. It is mesured as follows:

$$\text{F-measure} = \frac{2\,(Precision\ *recall\ )}{(\ Precision\ +Recall\ )} \qquad (5)$$

When both precision and recall have high values the net result is also has high values and it is decide as the fine method of agreement between them (Baeza and Riberio 1999).

REASULT AND DISCUSSION

The finest correctness of 95.99 % (Table 2) was extracted with two hidden layers that included four and two hidden neurons respectively which included a training set of 500. By algorithms we obtained the pace of TP at round 97% with 3 % FN, 93 % TP

and 7% FP respectively. We obtained the recall at 0.9955 and F measure at 0.959 respectively by using the ANN.

The importance of our model has gained over the research of (Dell' Aquila et al. 2000), person observed the germinated seeds the method of difference in seed position with respect to XY point of seed on superimposition of two successive images of before and after germination. A single image was used to extract 11 different features to estimate the germination of the seed. A single image is also advantageous because there may be a chance of fungal infection caused by the removal of the incubator by Jacobsen.

The very important dissimilarity between our research and the jossen et al (2010) is that the automation was entirely carried out without the human interventions for exchanging the data between various software areas.

CONCLUSION

In this paper, a automatic judgment between the seeds germination speed was implemented by using the image processing by computer vision and machine learning techniques. The outcome indicates that the accuracy of using ANN is very high (96%). The model thus developed has the capability of cauterization of germinated seed which has out raided other methods. It has also illustrated the great signs of time consumption in automated process and the human process in cauterizing the germination of seed quality.

REFERENCES

1. Hiroyuki Nonogaki, Seed biology updates - Highlights and new discoveries in seed dormancy and germination research, International Journal of Frontiers in Plant science, Volume 8, Apr 11, 2017.
2. Sangeeta Mishra, Effect of Temperature and Light on the seed germination of *Sida cordifolia* International Journal of Scientific and Research Publications, Volume 6, Issue 1, January 2016, ISSN 2250-3153.
3. Dushyant, K. S., Showkat A. G., Gurpreet S., Rajneesh K. A. and Rajendra S. 2015. Reproductive Phenology of Sida cordifolia L. Ann. Bio. Sci. 3 (1):10-15.
4. Jossen R, Kodde V, Willems L, Ligterink W, Van der Plas L, Hilhorst H. Germinator: a software package for high-throughput scoring and curve itting of Arabidopsis seed germination. Plant J. 2010; 62(1):148-159.
5. Rasband WS. Image. US National Institutes of Health, Bethesda, Maryland, USA, 2008.
6. Mayo M. Watson AT. Automatic species identification of live moths. Knowledge-Based Systems, 2007;20:195-202.
7. Witten, I., and Frank, E. (2005).Data Mining: Practical Machine Learning Tools Techniques (2nd Ed.).
8. Morgan Kaufmann. Ducournau S, Feutry A, Plainchault P, Revollon P, Vigouroux B, Wagner MH. An image acquisition system for automated monitoring of the germination rate of sunflower seeds. Comp. Electron. Agricult. 2004;44:189-202.
9. Granitto PM, Navone HD, VerdesPF, Ceccato HA. Weed seeds identification by machine vision. Comp. Electron. Agricult. 2002;33:91-103.
10. Geneve RL, Kester ST. Evaluation of seedling size following germination using computer-aided analysis of digital images from flat-bed scanner. Hortscience 2001;36(6):1117-1120.
11. Urena R, Ro driguez F in B erenguel M. A machine vision system for seeds germination quality evaluation using fuzzy logic. Comp. Electron. Agricult., 2001,32:1-20.
12. Dell'Aquila A, van Eck JW, van der Heidjen GWAM. The application of image analysis in monitoring the imbibitions process of white cabbage *(Brassica oleracea* L.) seeds. Seed Sci. Res. 2000;10:163-169.
13. Uchigasaki M, Serata K, Miyamoto S. An automated machine vision system for classification of seeds using color features. J. Agricult. Struct. 2000;30(4):325-332.
14. Baeza R, Ribeiro B. Modern Information Retrieval. Addison-Wesley Longman Publishing Co., Boston, 1999: 544 p.
15. McDonald MB. Seed quality assessment. Seed Sci. Res.1998;8:265-275.
16. Howarth MS, Stanwood PC. Extracting 3-D information using 2-D images of seeds. Comp. Electron. Agricult. 1994;10:175-188.

Harish S. Gujjar is working as an Assistant professor and Head in department of computer science, SSA GFGC, Ballari, Karnataka, India. He did his M.C.A (Master of Computer Application) from Vishveswaraiah Technical University, Belgaum, Karnataka, India. M. Phil (Computer Science) from Allagappa university, Karaikudi, Tamilnadu, India. And pursuing Ph. D (Computer Science) from Bharathiar University, Coimbatore , Tamilnadu, India.