Available online at www.sciencedirect.com

ScienceDirect

Procedia Computer Science 170 (2020) 1017–1022

Procedia Computer Science

www.elsevier.com/locate/procedia

The 10th International Symposium on Frontiers in Ambient and Mobile Systems (FAMS)
April 6-9, 2020, Warsaw, Poland

Prediction of the Corn Grains Yield through Artificial Intelligence

Noel Varela[a], Jesus Silva[b]*, Omar Bonerge Pineda[c], Danelys Cabrera[d]

[a,d] *Unversidad de la Costa, Barranquilla, Colombia*
[b] *Facultad de Negocios, Universidad Peruana de Ciencias Aplicadas, Lima, Peru.*
[c] *Universidad Tecnológica Centroamericana (UNITEC), San Pedro Sula, Honduras*

Abstract

Currently, the determination of the quality of the cereals is done manually by grain classifier experts prior to the marketing stage. In this paper we present a web software tool that allows determining the quality level of a corn sample automatically from an image of it. Image processing algorithms were implemented to correct distortions caused mainly by the capture process. The K-Means classification algorithm was used and a function was developed to calculate the hectolitre weight in relation to the sample area. The results obtained by the application for grades 1 and 2, are close to those measured by the experts. However, those for grade 3 have not been similar since the subsamples selected were not representative.

Keywords: Cereal Quality, Image Processing, Web Tool.

1. Introduction

Cereal quality is defined as the set of defects that make a batch of grains unfit for consumption [1]. These defects may occur for various reasons, such as unfavorable climatic conditions, soil degradation, biological adversity and disadvantages arising from agronomic practices [2]. Grains deteriorate on the plant, during harvesting or in

* Corresponding author. Tel.: +51-975857503
E-mail address: jesussilvaUPC@gmail.com

10.1016/j.procs.2020.03.080

warehouses when they are not given proper care. Alterations in the natural state and health of the grains are classified into several categories: damaged or faulty grains, chopped grains, presence of foreign material, commercially objectionable odors, presence of undesirable seeds, broken grains and/or cracked grains, among the most important. The above-mentioned defects, known as quality determinations, are those that must be analyzed during the quality assessment process [3][4], according to the quantity or intensity in which each of them is present in the sample.

On the other hand, producers do not know the quality and/or condition of their grains until the marketing process is developed and their value is fixed. This situation reduces the possibility of taking corrective actions, such as adjusting machinery at the time of harvest, in order to minimize, as much as possible, external factors that may affect or diminish in some way the quality of the grains [5].

Analyzing the above situations, this paper proposes to develop a tool that allows to determine the quality level of a corn sample automatically, from the image of a sample. In this way, producers, collectors and exporters will be able to obtain in a fast, precise and efficient way all the information of the measurements of each of the quality determining factors, simply by sending an image of the sample of their corn harvest from a computer or a mobile device.

2. Grain Quality Evaluation

According to [6], there can be three grades: Grade 1 indicates superior quality. Grade 2 determines intermediate quality, and Grade 3 determines inferior quality. Consequently, goods with this grade will receive a discount on the base purchase price. To determine the grade of a good, each quality item is classified in the corresponding grade, and then, all items determining grade are analyzed as a whole. Thus, the item that is rated the worst will be the one that determines the grade of the goods. Consequently, for a good to be Grade 1, all items must be in grade 1, so the mere fact that an item is not in that grade will determine another grade. The quality items that determine the grade for corn are the hectolitre weight [7], damaged grains (sprouted, fermented, rotten, burnt or greenish), foreign matter and broken grains [8].

3. Proposed Method

The main objective of this work is based on the design and implementation of an evaluation mechanism for each of the corn quality items in an automatic way. The stages involved in image processing are described below.

3.1 *Image Capture*

To start this process, a 50 g. corn sample should be dispersed on a flat surface, trying to separate the objects from the sample to avoid overlapping. Great care must be taken at this stage, as the process of homogenization of the sample as defined by the quality standard is quite rigorous, and failure to comply with it may result in a non-representative sample, such as some small grains remaining at the bottom of the container [9].

The surface should be a light blue A3-size sheet with the edge of a printed rectangle centered on the cardboard. In this way, the sample is contained in the rectangle. The plane of focus, which is the plane where the cardboard is placed to disperse the sample, should be perpendicular to the axis of the camera lens.

With regard to lighting, if the environment has good solar lighting, shadows are generated to complicate the process of detecting the contours of objects and interfere with the intensity of color in these areas. Therefore, to diffuse the light, soften the shadows and make a good image capture, the camera's flash is used in a closed environment with low light [10].

Figure 1 shows the different variants mentioned above to obtain a photograph of a corn sample with the chosen reference system.

3.2 Image pre-processing

This stage is intended to eliminate those noises or elements in the image that may distort the process of identifying objects corresponding to the corn sample. The steps involved in this stage include the classic image enhancement processes such as: image thresholding, geometric or elastic transformations, morphological processes and binarization [11][12]. The result generated by this stage is shown in Figure 2.

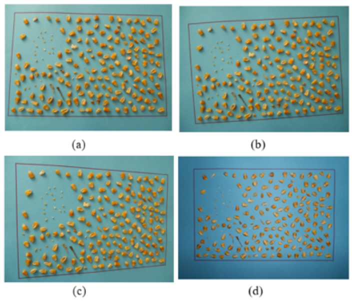
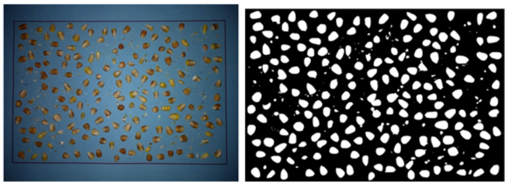


Fig. 1. Alternatives of different captures: a) Image with solar illumination. b) Image with solar illumination where the rectangle presents a small rotation in relation to the obtained picture. c) Image with solar illumination where the angle of the camera is modified slightly and this one does not remain perpendicular to the bottom. d) Image with flash of the camera in a closed environment with little light.

For defining the threshold, the application allows to select one or more fragments of the image that determine the its background. This value is determined by the minimum and maximum value of hue, saturation and brightness (HSV) [13] of each pixel belonging to the selected fragment(s) associated with the image background.



(a) (b)

Fig. 2. Image pre-processing: a) Input image. b) Output image.

3.3 Image Segmentation

The objective of the segmentation is based on partitioning the image into significant regions. The purpose is to obtain a list of objects on which the characteristics can be measured to determine which type each one belongs to. At this stage, through the application of filters and known algorithms, it is possible to separate the elements of interest from the sample. Later, the result of this stage will be used in the following ones, for its description, recognition and interpretation [14].

3.4 Characteristics Extraction

Once the image is segmented and the contour of the objects is extracted, the characteristics of each one are then extracted. Although there are several methods for selecting them, in [15], the geometric characteristics that reveal the shape of each of the objects were analyzed. Morphological and texture descriptors were selected to represent the shape and color observed by the experts in order to carry out their analysis. These characteristics are detailed in [16][17].

3.5 Identification of the Different Types of Objects

In the previous stage, the most determining features were defined to identify the different types of objects found in a sample. In order to differentiate each of these types of objects, it is necessary to define the average value that an object can acquire to belong to a certain type. For this analysis, the characteristic features of each type identified in a sample are determined [18].

The method used in [8] consists of defining the features or characteristics that indicate whether or not the object belongs to the class. The characteristic value of the feature for a given class, ideally, allows to differentiate that class from another in that feature. For example, one can differentiate a "Broken Grain" from a "Corn" by knowing its diameter. On the other hand, in Chamico seeds, which have a circular shape, the value of circularity and aspect ratio is close to 1. For an object to belong to the Chamico class the distance of Battacharya [19] between the HSV histograms of the class and the HSV histograms of the object must be greater than a threshold value close to 1.

3.6 Classification by Object

With the basic idea of the K-Means classification method [20] and some modifications (according to the need of the problem to be solved), the steps of the algorithm that allows to classify the objects according to the set of determining characteristics of each type are specified below:

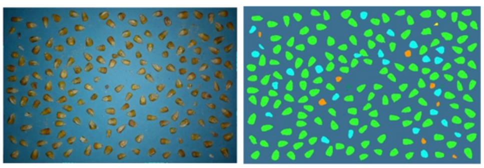
Step 1. Initialization: The set of objects to be classified is defined (which are the objects belonging to the sample), the number of groups (made up of each of the types: Corn, Broken Grain, Damaged Grain, Shaman and Foreign Matter) and a centroid for each group. For each group, the centroid is determined as the set of values associated with those features. This value is calculated as the average between the minimum and maximum values for each feature, except that the centroid corresponding to the Foreign Matter type group is not defined with any value since it is the default group. Objects that do not belong to any of the other groups are assigned there.

Step 2. Classification: For each object, its distance to each centroid is calculated, except for the Foreign Matter type, and the closest one is determined. As long as the object's value does not exceed the centroid value by a predefined percentage, the object is included in the group related to that centroid. Otherwise, it is determined that the object corresponds to the Foreign Matter type group. The distance to each centroid is calculated as the sum for each feature of the distance (difference) between the value of the centroid and the object associated with that feature.

Step 3. Centroid calculation: For each group generated in the previous step, its centroid is recalculated [21].

Step 4. Convergence condition: Converges when there is no exchange of objects between the groups. If the convergence condition is not met, steps two, three and four of the algorithms are repeated [5].

As a result of this algorithm, it is obtained the type of object to which each of the objects in the sample belongs, see figure 3.



(a) (b)

Fig. 3. K-Means classification of the objects in a sample: a) Input image. b) Output image

Figure 3-b shows the classification of each of the objects detected in the sample. The objects classified as yellow corn are painted in light green, the objects belonging to the colored corn in turquoise, the damaged grains in orange, the broken grains in dark green, the Chamico seeds in magenta and the objects corresponding to foreign matter can be seen in yellow.

3.7 Determining Quality Grade and Discounts

In order to evaluate the quality of the corn sample, it is necessary to know the percentage of damaged grains, broken grains and foreign matter, the quantity of Chamico seeds, the percentage of humidity and the hectolitre weight [21]. To obtain the first three, the percentage of the area occupied in relation to the total area of the objects in the sample is calculated, since this is the only way to determine this magnitude from an image. With respect to the amount of Chamico seeds, only the number of objects associated with the Chamico class is counted, and the count is adjusted to the duplicate of this class. This is because the quantity of objects associated with this class corresponds to a sample of 50 g.

4. Application Development

The architecture used for the development of the web application was the Spring Framework Model-View-Controller (MVC) [21] for the separation between the different layers of the application. Hibernate [14] was used for the persistence and MySQL [18] as a database management system that allows the object-relational mapping of the database. For the management, Maven [19], and Apache Tomcat [20] were used for the deployment of the application. The web tool developed allows the classification of corn grains and can be accessed from the web through a digital picture of the corn sample. By way of example, Figure 4 shows the classification carried out automatically by the developed application.

Fig. 4. Web application - Ranking result.

5. Results

Table 2 shows the comparative results measured by the specialist and those obtained by the application. These results correspond to the five samples analyzed by the experts of the Chamber of Commerce [20]. It can be observed in Table 1 that the results obtained by the application are similar to those analyzed by the specialist. Consequently, the statistical results of the analyzed tests were studied.

Table 1. Comparative results between the grain sorting experts and the web application developed.

Sample	Grade		Damaged Grade		Strange matter		Broken grain		Hectolitre weight	
N°	Specialist	App	Specialist	App.	Specialist	App.	Specialist	App.	Specialist	App.
1	1	1	3.14	2.24	0.19	0.34	0.09	0.10	78.13	77.35
2	1	1	1.25	1.36	0.06	0.06	0.70	0.20	78.85	76.25
3	1	1	1.84	1.75	0.07	0.01	1.32	0.05	78.25	76.54
4	2	2	1.14	0.69	0.20	0.18	0.48	0.20	73.25	73.69
5	2	2	5.36	4.74	0.22	0.15	0.14	0.08	74.86	74.36
6	2	2	1.47	1.64	0.13	0.40	0.56	0.56	73.74	73.74
7	2	2	1.62	1.20	0.14	0.10	0.79	0.30	74.47	73.06
8	2	2	1.74	2.36	0.05	0.00	2.14	0.80	76.95	75.74
9	3	3	3.32	2.74	0.15	0.55	4.62	4.21	74.35	71.36
10	3	3	1.25	6.36	0.44	0.16	5.00	0.74	76.01	73.14
11	3	3	1.14	4.47	0.20	1.30	4.10	0.33	74.47	73.95

6. Conclusions

The application developed extended the existing desktop application [14] for grain classification to a web environment, making it possible to evaluate sample quality from any device with Internet access. In this sense, it is not strange that the technological revolution has also reached the agricultural sector with more and more professionals from different disciplines that are grouped together to develop applications, aiming at facilitating their tasks. In addition, this application includes an unsupervised classification mechanism such as K-Means, improving previous results. On the basis of the analysis of the test carried out and the interaction with the professionals of the area, it was possible to improve the tool by adding the functionality of the estimation of the hectolitre weight in a satisfactory way.

References

[1] Ramírez, J. (2010). Rendimiento y calidad de cinco gramíneas en el Valle del Cauto. Tesis en opción al grado de Doctor en Ciencias Veterinarias. Instituto de Ciencias Agrícolas. La Habana. Cuba.

[2] SenseFly. (2014). El dron para la agricultura de precisión. Disponible en: https://www.sensefly.com/fileadmin/user_upload/sensefly/documents/bro chures/eBee_Ag_es.pdf [Consultado el 29 de septiembre del 2015].

[3] Cruz, M. C., Rodríguez, L. C. y Vi, R. G. (2013). Evaluación agronómica de cuatro nuevas variedades de pastos. Revista de Producción Animal, 25(1).

[4] Erenturk, K., Erenturk, S. yTabil, L. G. (2004). A comparative study for the estimation of dynamical drying behavior of Echinacea angustifolia: regression analysis and neural network. Computers and Electronics in Agriculture, 45(1-3).

[5] Hernández, D., Carballo, M. y Reyes, F. (2000). Reflexiones sobre el uso de los pastos en la producción sostenible de leche y carne de res en el trópico. Pastos y Forrajes,23(4).

[6] Hernández, R. M., Pérez, V. R. y Caraballo, E. A. H. (2012). Predicción del rendimiento de un cultivo de plátano mediante redes neuronales artificiales de regresión generalizada. Publicaciones en Ciencias y Tecnología, 6(1).

[7] López, A. M., Adolfo, A., Guido, J. P. y Ortega, A. C. (2006). Software de Predicción de la Producción Forrajera. Disponible en: http://www.unne.edu.ar/unnevieja/Web/cyt/cyt/2001/8-Exactas/E-002.pdf [Consultado el 29 de septiembre del 2015].

[8] Martín, B. yMolina, A. S. (2001). Redes neuronales y sistemas borrosos. 2ªed.España: Alfaomega. Ra-Ma.

[9] Lezama, O. B. P., Izquierdo, N. V., Fernández, D. P., Dorta, R. L. G., Viloria, A., Marín, L. R.: Models of Multivariate Regression for Labor Accidents in Different Production Sectors: Comparative Study. In International Conference on Data Mining and Big Data, Springer, Cham, 10942 (1), 43-52 (2018).

[10] Suárez, J. A., Beatón, P. A., Escalona, R. F., Montero, O. P.: Energy, environment and development in Cuba. Renewable and Sustainable Energy Reviews, 16(5), 2724-2731 (2012).

[11] Sala, S., Ciuffo, B., Nijkamp, P.: A systemic framework for sustainability assessment. Ecological Economics 119 (1), 314-325 (2015).

[12] Singh, R. K., Murty, H. R., Gupta, S. K., Dikshit, A. K.: An overview of sustainability assessment methodologies. Ecological indicators 9(2), 189-212 (2009).

[13] Sánchez L., Vásquez C., Viloria A., Rodríguez Potes L. (2018) Greenhouse Gases Emissions and Electric Power Generation in Latin American Countries in the Period 2006–2013. In: Tan Y., Shi Y., Tang Q. (eds) Data Mining and Big Data. DMBD 2018. Lecture Notes in Computer Science, vol 10943. Springer, Cham

[14] Aitkenhead, M. J., Dalgetty, I. A., Mullins, C. E. y Strachan, N. J. C. (2003). Weed and crop discrimination using image analysis and artificial intelligence methods. Computers and Electronics in Agriculture, 39(3).

[15] Varela, N., Fernandez, D., Pineda, O., Viloria, A.: Selection of the best regression model to explain the variables that influence labor accident case electrical company. Journal of Engineering and Applied Sciences 12 (1), 2956-2962 (2017).

[16] Olivera, Y., Machado, R. y Pozo, P. P. (2006). Características botánicas y agronómicas de especies forrajeras importantes del género Brachiaria. Pastos y Forrajes, 29(1).

[17] Bezerra, B. G.; Da Silva, B. B., Bezerra, J. R. C., Brandão, Z. N.: Evapotranspiração real obtida através da relação entre o coeficiente dual de cultura da FAO-56 e o NDVI. Revista Brasileira de Meteorologia 25 (3), 404 – 414 (2010).

[18] Diaz-Balteiro, L., González-Pachón, J., Romero, C.: Forest management with multiple criteria and multiple stakeholders: An application to two public forests in Spain. Scandinavian Journal of Forest Research 24(1), 87-93 (2009).

[19] Hák, T., Janoušková, S., Moldan, B.: Sustainable Development Goals: A need for relevant indicators. Ecological Indicators 60 (1), 565-573 (2016).

[20] Lampayan, R. M., Rejesus, R. M., Singleton, G. R., Bouman, B. A.: Adoption and economics of alternate wetting and drying water management for irrigated lowland rice. Field Crops Research 170 (1), 95-108 (2015).

[21] Delgado. A., Blanco, F. M.: Modelo Multicriterio Para El Análisis De Alternativas De Financiamiento De Productores De Arroz En El Estado Portuguesa, Venezuela. AGROALIMENTARIA 28 (1), 35-48 (2009).