ISSN: 2458-8989

Natural and Engineering Sciences

Supplement, 2019, 4(3): 100-106

Detection of Defective Hazelnuts by Image Processing and Machine Learning

Oğuzhan KIVRAK[1], Mustafa Zahid Gürbüz[2]

[1]Bandırma Vocational School Computer Programming, Bandirma Onyedi Eylul University, Turkey

[2]Department of Computer Engineering, Dogus University, Turkey

Emails: okivrak@bandirma.edu.tr, zgurbuz@dogus.edu.tr

Abstract

Hazelnut, is an oily food, that contains nutrients which is important for human health. The quality of the hazelnuts can be varied by internal and external factors such as the temperature of the environment, relative humidity of the environment, harvesting, drying and storage conditions, pesticide and mold growth. After harvesting, a machine (patoz machine) is used to separate the other shell of hazelnut. The patoz can mix the poor-quality hazelnut into the solid hazelnut, damage the shell and also discard impurities such as iron and stones during extraction. The average amount of impurities in raw hazelnuts at the time of 40 kg/ton. The average transaction for a factory is 200 tons per day, which can result in significant financial losses. The aim of this project is to separate intact hazelnuts from damaged or imperfect hazelnuts and impurities by using image processing and artificial intelligence. 1000 number of photos of hazelnut that obtained from the patoz machine were taken. They uploaded to the system. The system used supervised learning method. In this paper, the obtained results are very satisfactory

Keyword(s): Hazelnut, image processing, machine learning

Introduction

Hazelnut, which uses by food industries approximately 90% of the worldwide, is one of the most wanted raw materials for the bakery, chocolate and confectionery (Moscetti, et al., 2015). Hazelnut is an important export items which is about 70 percent of the World Trade, covered by our country (Gonenc, et al.,2006). So that, the quality of hazelnut is very important for our country. But the quality of the hazelnuts can be varied by internal and external factors such as;

- the temperature of the environment,
- relative humidity of the environment,
- harvesting, drying,
- storage conditions and
- pesticide and mold growth.

Using damaged or imperfect hazelnuts in production line in factory may cause poor quality of output product (crushed hazelnuts, hazelnuts cream). Also using mold hazelnuts in products is not healthy for human. So that the factories have to choose good quality hazelnuts for their output products. For this selection, they pay extra money for extra workers and expensive machines. Because of this, our motivation of this project is to reduce the cost of workers, to improve the quality of nuts and to prevent buying imperfect hazelnuts from higher price.

For damaged or imperfect hazelnuts, we will define the process of hazelnuts before buying from factories. First the farmers started to collect hazelnuts from tress on August every year. After collecting, hazelnuts are laid to dry in the sun on the grasses. When the green shell of hazelnuts changes color to brown, patoz machine separates the green shell of hazelnut. But the patoz machine can mix the poor-quality hazelnut into the solid hazelnut, damage the shell and also discard impurities such as iron and stones during extraction. After this process farmers also laid the hazelnut to dry in the sun and most of the time choose impurities such as iron and stones during extraction. Nowadays these choosing step skips by the farmers and this cause the poor quality. So that the aim of this project is to separate intact hazelnuts from damaged or imperfect hazelnuts and impurities by using image processing and artificial intelligence.

Material and Methods

Image Tagging

In machine learning, especially for supervised methods, dataset generation is one of the most part. Image process dataset should contain specific images for the field of study. For specific studies it is not available to find ready to go datasets. Researches prepares their datasets by themselves. In this study for dataset, 1000 photos of hazelnuts have been taken by a mobile phone for image tagging. Image tagging is a process to label an image by selecting in a rectangular area which is labeled by a semantic tag. So, the algorithm able to understand the content of image. For image tagging with Yolo v3, labelImg(github) project which is an open project, is used. In study, three trial has been done.

In first trial, the photos have taken with 12 megapixel and the hazelnuts are combined each other. In the second trail, the photos have taken with 48 megapixel and the hazelnuts are combined each other. In the third trail, the photos have taken with 48 megapixel and the hazelnuts has spaces between each other. In each trial, researcher tagged each photo as below Figure 1.

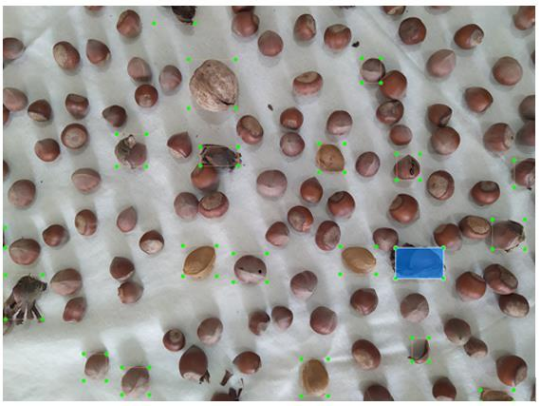

Figure 1 – Image tagging

Object Detection & Image Processing

Object detection refers to the capability of computer and software systems to locate objects in an image and identify each object. It has been widely used in many fields of practice such as face detection, vehicle detection, pedestrian counting etc. In object detection research Yolo algorithm is widely used. Yolo contributors has developed to the third generation of Yolo v3. In our paper, V3 algorithm has been used for object detection because V3 algorithm is improved in recognition accuracy and speed (Zhang, etc., 2019).

Image processing is a process for changing the nature of an image for improving its pictorial information for human interpretation or rendering it more suitable for autonomous machine perception (McAndrew,2016, p:1)

In study, Facebook Artificial Intelligence tool (Pytorch) and Darknet which is an open source neural network framework written in C and CUDA.

Measurement

In object detection, mean average precision (mAP) is the most common method for measuring the accuracy of the result (Kim and Lee, 2013). Beitzel and friends has described the formulation of mAP (Beitzel et. al., 2009). First of all, in this method, average precision must be calculated. For this calculation, precision and recall function must be used that results are in the range of 0 and 1.

Precision function measures the accuracy of predictions. Recall function is the ratio of achievement of positive findings over all positives. Precision and recall function declares as below:

$$precision = \frac{true\ positive}{true\ positive + false\ positive} \quad (1)$$

$$recall = \frac{true\ positive}{true\ positive + false\ negative} \quad (2)$$

To calculate the true positive or false positive, we used intersection over union (IoU) method. In this method, real tagged object and predicted rectangle checked whether overlap

percentage. If the percentage overlaps more than a threshold, it is classified as true positive. In this paper, the threshold is 50%. IoU formulation is below:

$$IoU = \frac{area\ of\ overlap}{area\ of\ union} \tag{3}$$

Figure 1 – IoU sample

Average Precision (AP) is determined by finding the area under the precision over recall curve.

$$AP = \int_0^1 p(r)\, dr \tag{4}$$

Mean average precision is the mean average of all average precisions over documents. Its formulation is below:

$$mAP = \frac{\sum\ AP}{Document\ size} \tag{5}$$

Results and Discussion

In this study, three datasets had been used. First one is taken by standard mobile phone camera with 12 MP resolution. The hazelnuts are crowded and intersects the hazelnuts with each other.

In figure 2, the sample picture and result are shown. mAP result is under 1% in this attempt and it classified as failed.

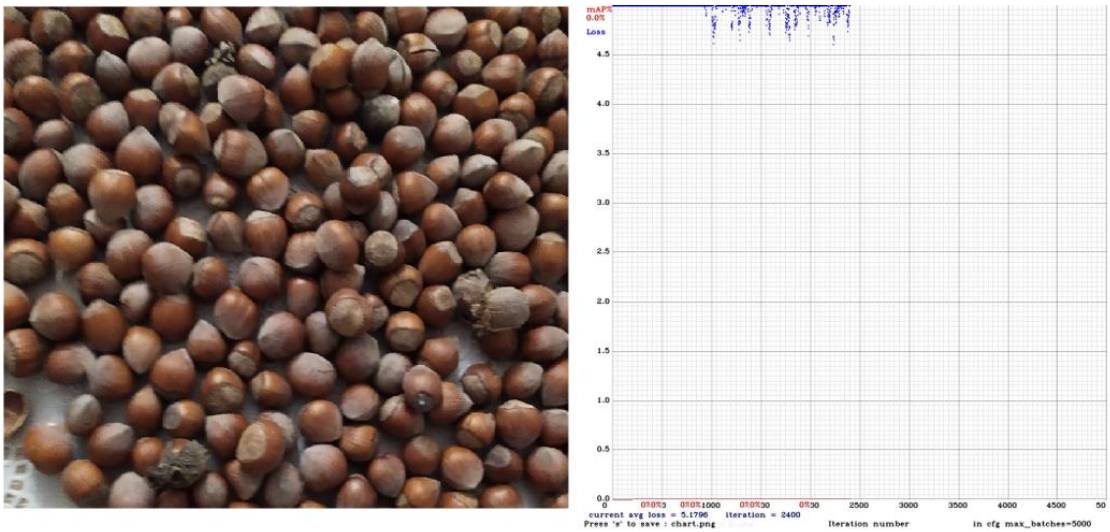

Figure 2 – First Attempt of detection. mAP is below 1%

The second one is taken by standard mobile phone camera with 48 MP resolution. The hazelnuts are crowded and intersects the hazelnuts with each other as in the first dataset. In figure 3, the sample picture and result is shown. mAP result is about 8%. it is classified as can be improved.

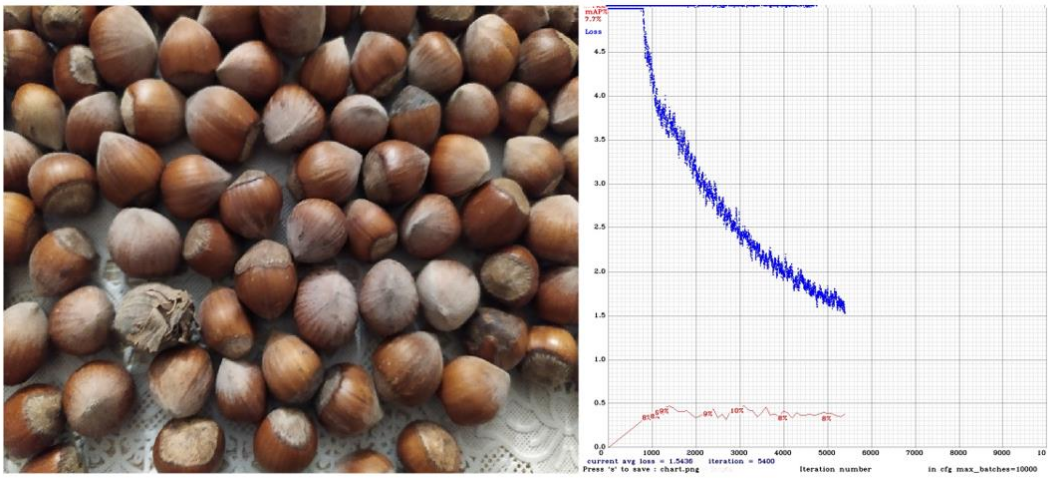

Figure 3 – Second attempt of detection. mAP is about 8%

The final dataset is taken by standard mobile phone camera with 48 MP resolution. But this time, the hazelnuts are separated and have spaces between each other opposite to other datasets. In figure 4, the sample picture and result is shown. mAP result is above 40%. it is classified as success in our research.

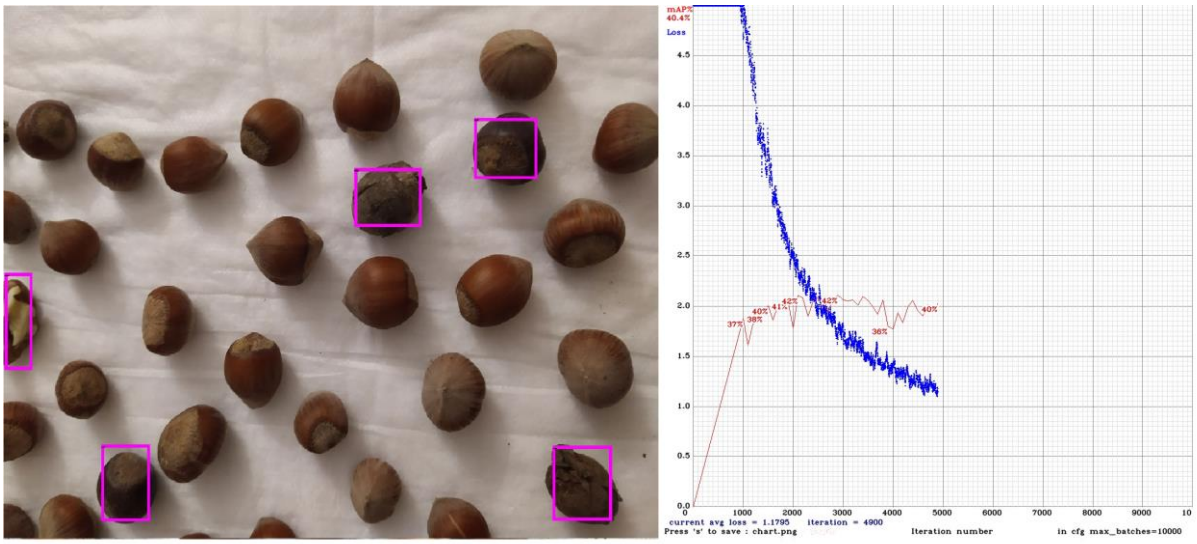

Figure 4 – Final result of detection. mAP is above 40%

At the final dataset, the photos of hazelnuts have some shadows. it may reduce the performance. In the future studies, the photo quality can be improved to get a more accurate result.

Conclusion

The aim of this project is to separate intact hazelnuts from damaged or imperfect hazelnuts and impurities by using image processing and artificial intelligence. 1000 number of photos of hazelnut that obtained from the patoz machine were taken. Three datasets are prepared for this experiment. Each dataset has the same hazelnuts with different resolution and layout. The applicable result is obtained by third dataset which has high resolution and separated layout. As further study, the quality of photos can be improved to get higher mAP value. Another object detection algorithm can be applied for comparing the performance of each method.

Conflict of Interest: The authors declare that they have no conflict of interest.
Note: This paper is presented in the International Conference on Artificial Intelligence towards Industry 4.0 held on November 14 - 16, 2019 at Iskenderun Technical University, Iskenderun, Turkey.

References

Gonenc, S.,Tanrıvermiş, H.,Bülbül, M. (2006). Economic Assessment of Hazelnut Production and the Importance of Supply Management Approaches in Turkey. Journal of Agriculture and Rural Development in the Tropics and Subtropics, 107(1), 19–3.

Moscetti, R., Saeys, W., Keresztes, J. C., Goodarzi, M., Cecchini, M., Danilo, M., Massantini, R. (2015). Hazelnut Quality Sorting Using High Dynamic Range Short-Wave Infrared Hyperspectral Imaging. Food and Bioprocess Technology, Volume 8, Issue 7, 1593–1604.

Zhang, X., Zhang, Li, D. (2019). Transmission Line Abnormal Target Detection Based on Machine Learning Yolo V3. International Conference on Advanced Mechatronic Systems (ICAMechS). DOI: 10.1109/ICAMechS.2019.8861617

Github : https://github.com/tzutalin/labelImg , 2019.

McAndrew, A. (2016). A Computational Introduction to Digital Image Process Second Editon. © 2016 by Taylor & Francis Group, LLC, International Standard Book Number-13: 978-1-4822-4733-6 (eBook - PDF)

Kim, I., Lee, C. (2013). An Efficient Gradient-based Approach to Optimizing Average Precision Through Maximal Figure-of-Merit Learning, J Sign Process Syst., 74:285–295, DOI 10.1007/s11265-013-0748-0

Beitzel S. M., Jensen E.C., Frieder O. (2009). MAP. In: LIU L., ÖZSU M.T. (eds) Encyclopedia of Database Systems. Springer, Boston, MA.