International Journal of Advanced Culture Technology Vol.9 No.1 190-195 (2021)
DOI https://doi.org/10.17703/IJACT.2021.9.1.190

IJACT 21-3-26

A Study on Fruit Quality Identification Using YOLO V2 Algorithm

Sang-Hyun Lee[1]

[1]*Assistant Professor., Department of Computer Engineering, Honam University, Korea*
leesang64@honam.ac.kr

Abstract

Currently, one of the fields leading the 4th industrial revolution is the image recognition field of artificial intelligence, which is showing good results in many fields. In this paper, using is a YOLO V2 model, which is one of the image recognition models, we intend to classify and select into three types according to the characteristics of fruits. To this end, it was designed to proceed the number of iterations of learning 9000 counts based on 640 mandarin image data of 3 classes. For model evaluation, normal, rotten, and unripe mandarin oranges were used based on images. We as a result of the experiment, the accuracy of the learning model was different depending on the number of learning. Normal mandarin oranges showed the highest at 60.5% in 9000 repetition learning, and unripe mandarin oranges also showed the highest at 61.8% in 9000 repetition learning. Lastly, rotten tangerines showed the highest accuracy at 86.0% in 7000 iterations. It will be very helpful if the results of this study are used for fruit farms in rural areas where labor is scarce.

Keywords: *YOLOV2, Region Of Interest, Bound Box, R-CNN, Artificial Intelligence.*

1. INTRODUCTION

Currently, the labor force of domestic farmers is suffering from aging, and technologies to help underprivileged workers by utilizing artificial intelligence, one of the fourth industrial revolutions, are in dire need. In the past, experts in the field of agriculture inspected crops through direct visits or counseling on pests to determine pests, but recently, research on image processing technology has also been actively conducted.

Representatively, a fruit quality system using a computer vision system (CVS) for the quality of fruits harvested from orchards is being studied. Computer vision systems have been widely studied and applied because they can repeatedly perform high-precision inspections by replacing manual inspectors in applications such as defect inspection, classification, and recognition [1].

Among the recent artificial intelligence models, the Convolution Neural Network (CNN) model is the most used image processing algorithm. Convolution Neural Network (CNN) [2] is a technology that mimics the structure of the human optic nerve, and automatically learns all the features necessary for recognition, from image processing to self-recognition, merchandising, and object recognition, while effectively learning shape variation. It is an absorbable algorithm [3]. And the YOLO model is an algorithm that estimates the type and location of an object by looking at an image once and calculates the object probability for several bounding boxes through a single network [4]. In particular, as a technology that can automatically generate the title of a document that can be represented by compressing the content of the document, researches are actively being conducted to generate a document summary using a recurrent neural network (RNN) [5].

Therefore, this paper attempts to classify the quality of fruits by using the YOLOV2 model, which is the

Manuscript received: January 26, 2021 / revised: February 5, 2021 / accepted: February 15, 2021
Corresponding Author: leesang64@honam.ac.kr
Tel: +82-62-940-5285, Fax:+82-62-940-5285
Assistant Professor, Dept. of Computer Engineering, Honam University, Korea

technology that is receiving the most attention among the image recognition fields according to the recent development of artificial intelligence technology [6]. This is because YOLOV2 (YOLO9000), the second version of the YOLO family, was adopted to solve the problem of tangerine status, and the YOLO model is faster and has superior performance compared to other models in terms of speed.

2. RESEARCH CONTENT

2.1 YOLO V2 Model

The YOLO system stands for “You Only Look Once” [7], and is an algorithm that has strengths in real-time processing speed in the field of object detection. For Faster R-CNN [8], which has been used in the past, R-CNN-series detection networks first select a candidate group for ROI-Region of Interest: Region of Interest that is likely to have an object in the image. The small image parts of the ROIs selected as candidates are classified by a Classification network and a bound box is found. The Region Proposal Network proposes regions in the image where objects are likely to exist.

Until now, the detection system changed the classifier or localizer to suit the purpose to perform detection, and then divided the image into a number of small images and compared it with the original model. In doing so, areas with high scores are detected and judged. However, in the case of YOLO, one neural network is applied to the entire image. The neural network divides the image into several sections and calculates the probability that the object to be found is correct for each section. It helps to find the object to be detected by giving different weights to the bounding box according to the predicted probability.

R-CNN (Region-based Convolutional Neural Network) generates rectangular windows of various sizes by selective search. After creating a rectangular area of region proposal that can roughly classify objects through the created window, this area is classified as a classifier, redundantly removed by post-processing, and re-scored by comparing with other objects again. Therefore, as a drawback of R-CNN, the task of training each of these components is very complex, and it takes a lot of time to perform detection.

To overcome these shortcomings, YOLO handles the single regression problem, regression and the target point for the object in the bounding box as a probability by calculating the coordinates in pixels of where the object is in the image. Because of the advantage of being able to process images in real time at high speed, we decided that it was the best algorithm to detect objects with similar fruit shapes and applied YOLO technology.

Image preprocessing was performed through YOLO_MARK to learn the data set in YoloV2 applied in this paper. Here, YOLO_MARK is a program that designates the object to be learned in the image.

2.2 YoloV2 based tangerine classification

The development environment in this paper was Windows 10, CPU: Intel Pentium G4560, GPU: Nvidia Gefore GTX 1050, and software for learning was OpenCV(4.1.0), CUDA(11.1), cuDNN(8.0.4). And Darknet53.conv.74 was used as the learning weight.

2.2.1 The Dataset

In order to create a tangerine detection model selected as a sample model, the dataset was obtained by crawling search images related to tangerine from Google. The basic classification values of the data set for learning are classified into three types (mandarin orange, unripe mandarin orange, and rotten mandarin orange).

The size of the training data set is from 175x289 to 1300 x 866, and there are 1-12 fruits in the picture, and the size of the fruits is extracted from 6 x 20 to 95 x 45. The images were classified manually using Yolo_mark, with 183 rotten mandarins and 457 mandarin photos. As shown in Figure 1, parts that are difficult to detect due to background or leaves are excluded from training data.

Figure 1. Dataset (sample)

2.2.2 YOLOV2 Architecture

In this paper, we propose a detection model that detects the tangerine image coming in the input value and classifies the status (tangerine, unripe tangerine, rotten tangerine), and uses the Yolo algorithm, Darknet, and Yolo_mark for learning.

In order to proceed with the work of taking note of the boundary box coordinates in the windows environment, apply Visual Studio to OpenCV, save using Yolo make, and modify the filter and class of the .cfg file according to the learning values [6]. In Table 1 and Figure 2, when learning is performed for each image, the convolutional filter and maxpool learn while reducing or increasing the size of the incoming image.

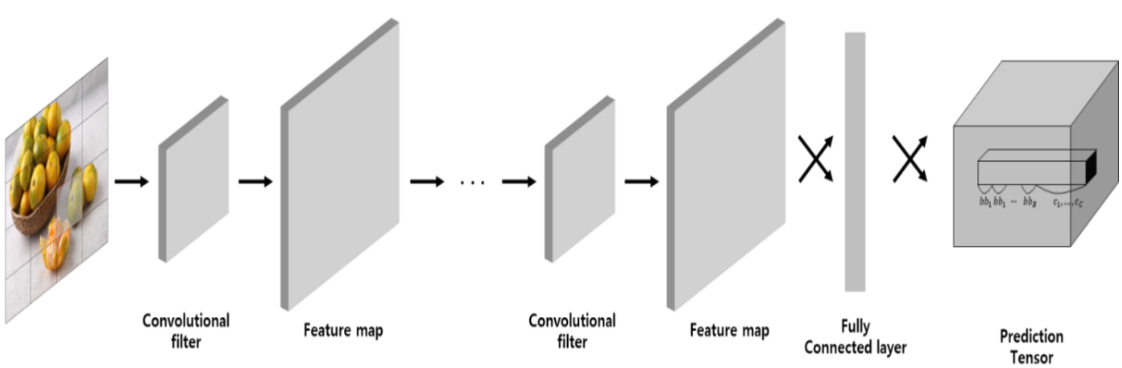

Figure 2. YOLO Network Architecture

Convolutional with a Size/Stride value of 1 x 1 reduces the number of channels in a specific map in half. The number of learning of the detection model is 9000 times, and the subsequent learning ends the learning because the change value of Loss is insignificant. Loss value can be checked in Figure 3.

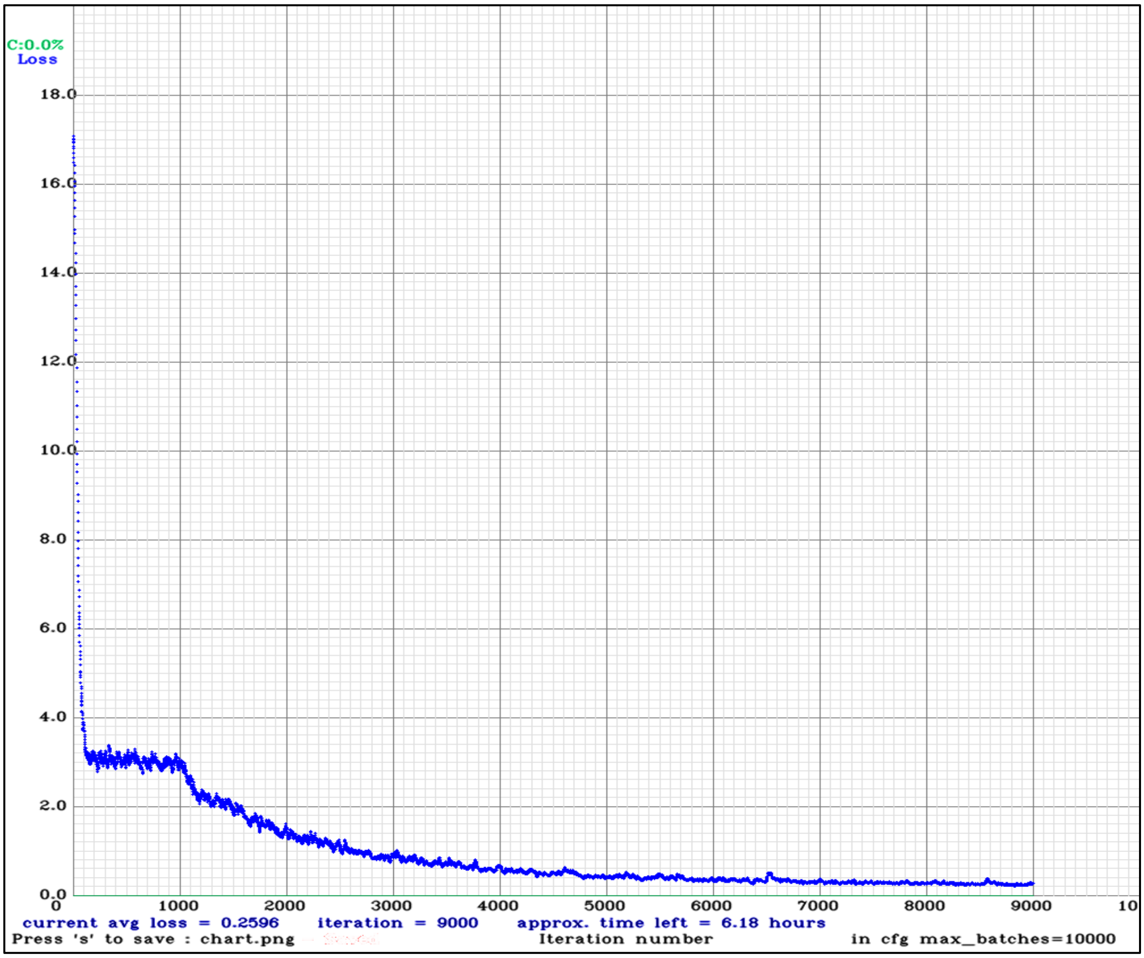

Figure 3. Loss value

3. RESULT

In this paper, the value learned in the console is calculated by executing Figure 4 to show the result of the proposed detection model, and the test result of the number of learning times 9000 is shown as a visualized figure by drawing a bounding box on the orange image.

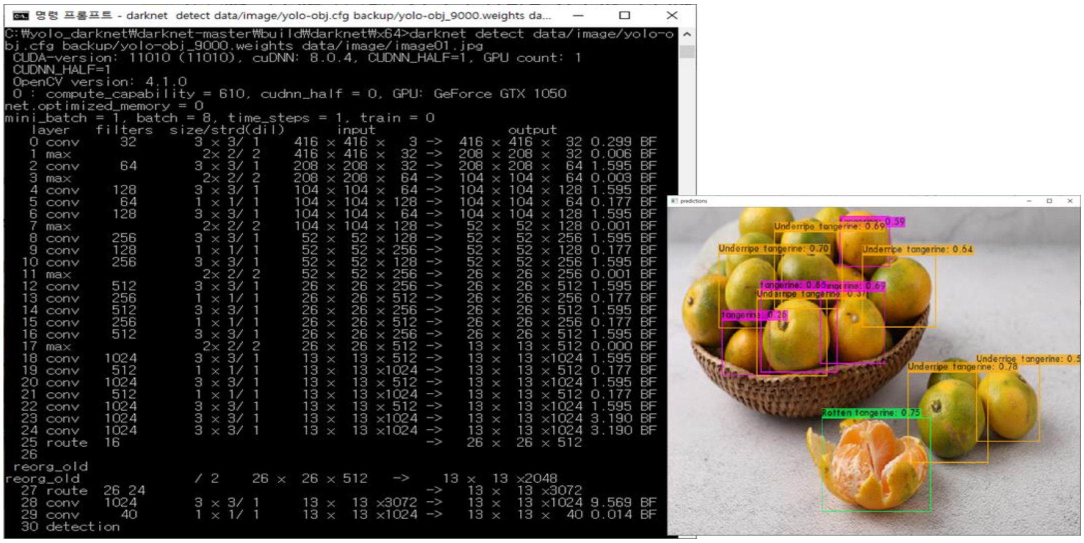

Figure 4. Console / Test result of 9000 learning times [9]

The check results according to the next number of learning are shown in Table 1.

In Table 1, it was found that the number of learning counts of 2,000 or less was not significant in accuracy, and in the number of learning of 3,000 to 5,000 counts, normal tangerine (65.0~53.7%), Unripe tangerine (36.0~40.5%), and Rotten tangerine (79.0~64.0) %), the number of images detected was low, but data other than tangerine was detected. At 6000 training counts, the accuracy and precision of tangerine were very high with normal tangerine (83.6%), unripe tangerine (60.2%), and rotten tangerine (80.0%). The accuracy and precision of 7,000 to 8,000 are poor with normal mandarin oranges (43.0 to 40.7%) and unripe mandarin oranges (51.7 to 56.4%) compared to 6,000 learning counts, but the accuracy of rotten mandarins is 7,000 counts (86.0%). The detection rate was very high. The detection accuracy of normal tangerines was high (60.5%) in the number of learning 9,000 counts, and the detection accuracy of unripe tangerines was high (61.8%). However, the detection accuracy of rotten mandarin oranges was higher in 7,000 learning counts (86.0%) than in 9,000 learning counts (75.0%). Therefore, after confirming the best learning rate, it will be possible to increase the detection rate using the result.

Table 1. Main parameters

Division	Number of learning	Normal tangerine		Unripe tangerine		Rotten tangerine	
		Count	accuracy	Count	accuracy	Count	accuracy
1	1000	0	0%	0	0%	0	0%
2	2000	0	0%	7	50.7%	1	0%
3	3000	7	65.0%	3	36.0%	1	79.0%
4	4000	6	68.7%	6	44.8%	1	64.0%
5	5000	6	53.7%	6	40.5%	1	64.0%
6	6000	5	83.6%	5	60.2%	1	80.0%
7	7000	3	43.0%	6	51.7%	1	86.0%
8	8000	4	40.7%	8	56.4%	1	80.0%
9	9000	4	60.5%	6	61.8%	1	75.0%

4. CONCLUSION

This paper attempts to classify the quality of fruits using the YOLOV2 model, which is the technology that is drawing attention among the image recognition fields with the recent development of artificial intelligence technology. YOLOV2 (YOLO9000), the second version of the YOLO family, was adopted to solve the problem of tangerine status, and the YOLO model is faster and has excellent performance compared to other models in terms of speed.

We are the purpose of the study in this paper is to study a detection model that distinguishes the quality and condition of fruits, and to proceed with learning using YoloV2. The accuracy of tangerines was very high at 6000 learning counts of the number of learning, and the accuracy of unripe tangerines was 9000 counts, and 7,000 counts of learning rotten tangerines was the best. And if you compare the training of 6000 to 7000 with each other, the accuracy of rotten tangerine has little change, but there is a difference in accuracy between tangerine and unripe tangerine. As a result of this study, we were able to find the ranking with the highest accuracy in terms of learning rate.

REFERENCES

[1] M. Baigvand, A. Banakar, S. Minaei, J. Khodaei, and N. B. Khazaei, Machine vision system for grading of dried figs, Computers and Electronics in Agriculture, Vol. 119, pp. 158-165, 2015. https://doi.org/10.1016/j.compag.2015.10.019

[2] Face It – The Artificially Intelligent Hairstylis t, https://software.intel.com/en-us/articles/face-it-the-artificially-intelligent-hairstylist, Jun. 2018.

[3] G. S. Juan, M. G. Jose D, S. O. Emilio, M. S. Marcelino, M. B. Rafael, and B. Jose, Detecting rottenness caused by Penicillium genus fungi in citrus fruits using machine learning techniques, Expert systems with applications, Vol. 39, No. 1, pp. 780-785, 2012. https://doi.org/10.1016/j.eswa.2011.07.073

[4] Jun-Woo Park, Ji-Hun Kim, Jin-Seo Lee and Young-Duk Kim, "The identification and propagation of CCTV Fire situation in Road Using Deep Learning-Based YOLO," Information and Control Symposium, pp. 530-531, October 2020.

[5] Sung-Min Cho and Wooseng Kim, "Automatic Document Title Generation with RNN and Reinforcement Learning," Journal of Information Technology Applications & Management, Vol. 27, No. 1, pp. 49-58, 2020. https://doi.org/10.21219/jitam.2020.27.1.049

[6] I. Sa, Z. Ge, F. Dayoub, B. Upcroft, T. Perez, and C. McCool, DeepFruits: A fruit detection system using deep neural networks, Sensors 2016, Vol. 16, No. 8, 2016. https://doi.org/10.3390/s16081222

[7] J. Redmon, S. Divvala, R. Girshick, A. Farhadi, "You only look once: Unified, real-time object detection", Proceedings of the IEEE conference on computer vision and pattern recognition, pp 779-788, 2016. https://doi: 10.1109/CVPR.2016.91

[8] R. Girshick, J. Donahue, T. Darrell, J. Malik, "Rich feature hierarchies for accurate object detection and semantic segmentation", Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580-587, 2014. https://doi:10.1109/CVPR.2014.81
http://item.gmarket.co.kr/Item?goodscode=1926792948

[9] Yoo, Suk Won. "Object Recognition Using Comparison of External Boundary." International Journal of Advanced Culture Technology Vol. 7, No. 3 pp. 134–42, 2019. https://doi.org/10.17703/IJACT.2019.7.3.134