Detection and Ripeness Classification of Bananas Using Deep Learning Methods

Dinis Rocha Nuno Machado Tiago Miranda

I. Problem description

Ripening is the process in fruits and vegetables that causes these to become more palatable. In general, fruit becomes sweeter, less green, and softer as it ripens. These characteristics are desirable for most animals as eating unripe fruit can often lead to stomachache or stomach issues. Therefore, determining the ripening stages of each fruit is a skill developed by most animals. Recent advances in human civilization lead to the global availability of many different varieties of fruit and vegetables, inevitably contributing to the increase of food waste by missing the optimal consumption period of these items (from the client perspective as well as the manufacturer and reseller). This project intends to tackle this issue and aid the client or the seller to reduce food waste.

II. Why is the problem important?

Food waste is a major issue in today's society that is often understated and underestimated. One third of all food produced is lost or wasted (around 1.3 billion tonnes of food) costing the global economy close to 940 billion US dollars each year. Almost half of all fruit and vegetables produced are wasted (US citizens alone throw away 5 billion bananas each year).

Another reason to tackle this issue is that food waste is terrible for the climate. Up to 10% of global greenhouse gases comes from food that is produced, but not eaten. Wasting food is worse than total emissions from flying (1.9%), plastic production (3.8%) and oil extraction (3.8%). If food waste was a country, it would be the third biggest emitter of greenhouse gases after USA and China.

Therefore, for these and many more ethical and environmental reasons, food waste is a major societal problem that needs addressing [1] [2].

III. How is the problem addressed?

To address this problem, we preliminarily propose a computer vision solution that might serve industrial and consumer purposes. This approach is in two stages: first, the desired ripeness outcome is assessed; then, data collected by the user is employed to determine the time it will take for a specific banana to reach that outcome. Regarding the first stage, the assessment might be carried out in several different ways. In this work, we intend to focus on a consumer-centered approach, in which they are shown a sequence of pictures of bananas and will approve or discard each one according to their preference, perhaps in a Tinder-like fashion. As for the second stage, we will need computer vision and deep learning tools to produce robust results. Here, the media used as input data is a relevant aspect to consider. While image and video obtained with a smartphone will be the preferred data for a consumer application, an industrial setting may demand complementary information, for example in the form of infra-red sensors.

It is worth noting that a deep learning approach to estimating "time until a certain level of ripeness is reached" requires abundant data, not only of bananas in different stages of maturation together with time information, but also of the different environmental conditions they might be subject to. Furthermore, these very environmental conditions would be necessary for prediction, which could be unfeasible in a consumer application. We propose to tackle these issues in two ways: either by restricting our solution to a specific environment and estimating the time to maturation in these conditions, or by reframing our problem as "predicting the current ripeness stage": this would use tbe same architecture but would circumvent the data constraints.

IV. System arquitecture and main modules

The ripeness predictor, corresponding to the second stage outlined above, may be divided into three modules, as shown in Figure 1.

The process begins with the acquisition of image or video. Afterwards, it is necessary to locate the bananas in the frame, if any exist. This can be achieved with object detection tools, that identify and locate instances of one or more classes in an image. At the time of writing, deep learning is the preferred approach to object detection problems. In particular, the YOLO series of models [3], [4] is a fast and robust option to consider. It yields bounding boxes containing the instances found as well as confidence levels assigned to each one.

We observe that, this being an intermediate step of the process, the best way to approach it is determined by the needs of the ensuing procedure. Indeed, it may prove useful not only to locate the bananas but also to remove the background, i.e. to perform segmentation, so that the environment doesn't influence predictions. Fortunately, it is easy to incorporate this procedure in YOLOv8 [4], and we might pursue this option if needed.

We must account for the possibility that the object detection model we will use has not been exposed to bananas under certain maturation conditions. To overcome this, we have the option of fine-tuning the model [5], that is, re-training it on the datasets of different ripeness stages.

Several techniques were found to classify the ripeness stage of bananas using Machine Learning methods. Some by simply resizing the images and inputting them into a *CNN* [6] [7]. Others by converting the color space to *HSV* and then calculating the ratio of brown areas to the whole banana area and using this value and *HSV* values as inputs to a simple Neural Network [8]. The use of an *infrared* source could also provide additional data to the model [9]. Data augmentation techniques could also be implemented in order to balance the data and improve the robustness of the model. Another possible solution would be to combine several datasets, however some problems may arise because of the different labeling methods used. The exterior of a banana might not be the only indicator of ripeness, for example if stored in a fridge they might go brown while the interior remains unripe, so it might be interesting to try to use, for example, a measurement of its dielectric constant which is related to the ripeness level, however datasets for this kind of data might be limited [10] [11]. The use of attention based models such as transformers could also be interesting to explore, even though these are often considered to be less effective in classification tasks when compared to *CNN*'s.

An additional part of the project would be to first assess a user's preference regarding the ripeness of a banana, possibly using a decision tree, and then given a photo of a banana determine the time it will take for that banana to reach desired ripeness. This, however, will be challenging considering the lack of datasets suited for this time analysis. A possible solution would be for us to create our own dataset, but there's no guarantee that the results will be satisfying.

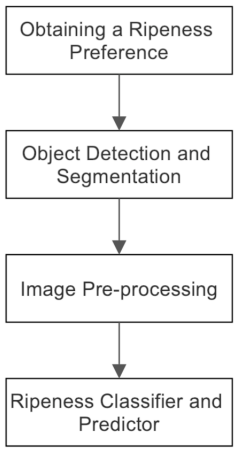

Fig. 1. Proposed workflow

REFERENCES

[1] https://www.ozharvest.org/food-waste-facts/

[2] https://www.chicagotribune.com/opinion/commentary/ct-opinion-food-waste-20210409-3k3llled4fbmlp3nwhiej3o354-story.html

[3] C. Yang, A. Bochkovskiy, H. M. Lia0 (2022). 'YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors'. arXiv:2207.02696.

[4] J. Terven, D. Cordova-Esparza (2023). 'A Comprehensive Review of YOLO: From YOLOv1 to YOLOv8 and Beyond'. arXiv:2304.00501.

[5] S. Rath (2022). 'Fine Tuning YOLO v7'. https://https://learnopencv.com/fine-tuning-yolov7-on-custom-dataset/

[6] N. Saranya, K. Srinivasan, S. K. Pravin Kumar (2021). 'Banana ripeness stage identification: a deep learning approach'. Journal of Ambient Intelligence and Humanized Computing. https://doi.org/10.1007/s12652-021-03267-w

[7] Y. Zhang, J. Lian, M. Fan and Y. Zheng (2018). 'Deep indicator for fine-grained classification of banana's ripening stages'. EURASIP Journal on Image and Video Processing. https://doi.org/10.1186/s13640-018-0284-8

[8] F. M. A. Mazen, A. A. Nashat (2018). 'Ripeness Classification of Bananas Using an Artificial Neural Network'. Arabian Journal for Science and Engineering. https://doi.org/10.1007/s13369-018-03695-5

[9] L. LUIZ, C. A. NASCIMENTO1, M. J. V. BELL, R. T. BATISTA, S. MERUVA, V. ANJOS (2021). Use of mid infrared spectroscopy to analyze the ripening of Brazilian bananas. Food Science and Technology. https://doi.org/10.1590/fst.74221

[10] M. Soltani, R. Alimardani, M. Omid (2011). Evaluating banana ripening status from measuring dielectric properties. Journal of Food Engineering. www.elsevier.com/locate/jfoodeng

[11] P. Baglat, A. Hayat , F. Mendonça, A. Gupta, S. S. Mostafa and F. Morgado-Dias (2023). Non-Destructive Banana Ripeness Detection Using Shallow and Deep Learning: A Systematic Review. https://doi.org/10.3390/s23020738