Predicting Apple Sugar Sweetness by Applying Apple Images on Hybrid Machine Learning: Machine Learning and Deep Learning

Heejun Park
College of Science & Technology
Yonsei University
Wonju, South Korea
parkie0517@yonsei.ac.kr

Hanbi Kim
Department of Computer Science
Chungbuk National University
Cheongju, South Korea
hanbikim20@g.cbnu.ac.kr

Jeongho Ha
College of Engineering
Jeju National University
Jeju, South Korea
hjh4212@naver.com

Seokkhyeon Heo
Blockchain Security
Jeju National University
Jeju, South Korea
gj4535@gmail.com

Taeyun Kim
College of Economics & Commerce
Jeju National University
Jeju, South Korea
persimm0ncrack@gmail.com

***Abstract*— Given the history of consumer experiences, it is difficult to select the best apples in terms of sugar level or sweetness. In addition, it is challenging for farmers to predict the optimal harvesting time based only on their experience. To solve this problem this research aims to develop a Hybrid Machine Learning(HML) model. HML is a fusion of Machine Learning(ML) and Deep Learning(DL). ML model is trained using features such as the color of apples. The other is a DL model which uses images of apples to train. Both models will give a numerical Brix value as an output. The predictions will then be used to give an estimate of how sweet an apple is. While previous research intends to find just one model from either ML or DL. The purpose of this research is to use HML model to increase the prediction accuracy of apple sugar levels in a non-destructive manner. In this project, the best performance of ~% accuracy was obtained by combining the two models. Shortly, this research could be implemented into smartphones as an application.**

Keywords—Hybrid Machine Learning, fruit sweetness classification, multiple linear regression, Convolution Neural Network

I. INTRODUCTION

The United States is the second largest country that grows apples in the world after China (4,336,000 tons), according to the United States Department of Agriculture [1]. People in the US can buy apples easily due to the bountiful supply of crops. Yet what is a critical factor when purchasing fruits? There are many elements in fruits, such as shape, color, and acidity. However, sweetness is the most significant element as it is a key factor in the flavor of fruits. It is important not only for consumers but for farmers as well. High sugar-level apples give economic benefits and market competitiveness to farmers. Such apples have high demand from consumers. But how are the apples harvested? People commonly think farmers would pick a few random apples and check the sugar level before harvesting. In reality, farmers harvest apples as long as they are ripe without checking the sweetness level. All the apples harvested by the farmers are sent to the Agricultural Products Processing Center (APC). APC has machines to measure the sugar level of the apples. The problem is that farmers harvest apples without knowing the Brix value, i.e., not checking if they are good enough to be sold for commercial value. The paper's proposed system benefits farmers by providing the sugar levels for their apples. Farmers would have a contract on good terms by utilizing the estimated value based on apple sweetness. Moreover, it helps raise the selling rate of apple by performing the method. For customers, knowing the information on sugar levels before buying would lead to more satisfying purchases. However, the reality is that most consumers purchase apples without knowing their sweetness level, which would reduce the chances of making a satisfactory purchase.

There are two universal ways to measure sugar levels: The destructive sugar levels measurement method (DM) and the Non-destructive sugar levels measurement method (NDM). DM measures sugar levels by testing the juice from cut fruits. It is a more affordable method than NDM. A disadvantage is that cut fruits lose their value to be sold. NDM displays sugar levels by detecting the reflected wavelength from the apples that were irradiated to near-infrared. It measures the sweetness of apples without damaging the fruits and shows the sugar levels immediately. However, the method is expensive, and the accuracy of the outputs is not stable. Therefore, this paper suggests an NDM, based on the combination of deep learning and machine learning by the data from images. It removes the disadvantages and maximizes the advantages of both methods. It is time and cost-effective and reduces the amount of labor.

This research uses 200 Pixie Crunch apples to obtain the data. Data such as pictures and Brix values train the HML models. By photographing the top, bottom, front, rear, right and left directions of an apple, the authors collected 6 different images. The reddest side of an apple is set as the front side. Starting from the front side, the authors took pictures by rotating an apple 90 degrees clockwise. When measuring the sugar concentration in a fruit, Brix is a degree commonly used as a unit of measurement. However, the Brix value alone cannot indicate the absolute sweet taste for all apple varieties. Complex information, such as its acidity, the amount of obtained sunshine, and other factors affect the sweetness of apples. However, when comparing the sweetness of the same kind, the apple with a higher Brix value is sweeter. Thus, Brix readings help estimate the sugar levels in a targeted fruit. According to the apple maturity guidelines of Michigan State University, the definition of apple sweetness is broken down into four parts based on Brix levels. It classified apple sweetness for all apple varieties in Table 1 below [2].

Table 1: Sweetness of all apple varieties (Unit: BRIX)

	LOW	FAIR	GOOD	EXCELLENT
ALL APPLE VARIETIES	UNDER 11	11	12	OVER 13

The Brix distribution for the measured apples in this research was between 9.70 °Bx to 16.75 °Bx, with an average of 13.20 °Bx. Most of the Brix values were between 12 °Bx to 13 °Bx. Referring to Table 1, this research classified the apple sugar level in Table 2 below.

Table 2 : Sweetness of Pixie crunch (Unit: BRIX)

	LOW	MEDIUM	HIGH
PIXIE CRUNCH	UNDER 12	12	OVER 13

The research aims to draw the Brix level from apple images by developing a HML model that combines ML and DL. This technology can be implemented into smartphones as an application. From further development, the application will have various features. It is used to measure the sugar levels of multiple apples at once. At the same time, it can be used to pick out rotten apples that tend to have a higher Brix value. In addition, this technology will be applied to Unmanned Aerial Vehicle (UAV) research to detect far-distance apples using a drone. Lastly, this technology has the potential to expand across the fruit that uses the Brix value.

II. Literature Review

Prior to starting research on predicting apple sweetness using HML, the author found previous related research.

A. Analysis of Apple Colors and Sugar Contents Using Linear Regression[3]

Kim's research utilized linear regression (LR) to analyze the relationship between color space (RGB, HSV, La*b* colors) and sugar content. In the paper, Kim found out that high-sugar content apples had a more yellow surface. Kim's research used LR to determine the causal relationship between the independent and dependent variables. Kim's research only used the picture of the top view of an apple. This method has a problem as just one side does not represent the whole surface of an apple. To compensate for the problem, this paper develops a LR model that represents the correlation using 6 different viewpoints of an apple.

B. Predicting Fruit's Sweetness Using Artificial Intelligence—Case Study: Orange[4]

The purpose of Mustafa's research was to find the relationship between the RGB values and the sweetness of the orange by using the Orange Data Mining Tool (ODMT). This paper applied various scoring methods. The methods showed, oranges with higher red values in RGB, had a higher concentration of sugar. In the case of apples, yellow and red are important for determining sugar levels. Mustafa's research used 5 models that include ML and DL models. The accuracy of ML models was relatively higher than DL models since only 50 oranges were used for training. Therefore, the small number of samples did not bring good results for DL models. In apple research, data augmentation has been applied to expand a small amount of data to increase the accuracy of the DL model.

C. A Non-Invasive Method to Classify the Sweetness Levels of Apples[5]

Lee's research collected a total of 130 apples to generate the dataset. The author acquired 6 different images from one apple. The reddest side of an apple is set as the front side. The picture of the top, bottom, front, rear, right and left sides of an apple were taken for the datasets. This method influenced the way of collecting data. Lee's research used New Zealand Fuji, Chile Fuji, and American Envy Fuji apples to classify sugar levels. Most of the sweetness values were between 14 °Bx to 16 °Bx. Thus, Lee's research classified the definition of apple sugar level in Table 3 below. The way Lee's research defines the classes of apple sugar levels is similar to this paper. Through this, the author confirmed that the Pixie Crunch classifications are reasonable.

D. Apple Sweetness Measurement by image processing technique[6]

Ittatirut's research proposed the Apple Sweetness Measurement System (ASMS) that shows the sweetness value of an apple. The system gets the image of a flesh part of an apple, then returns the sweetness value to its Graphical User Interface (GUI). Ittatirut et al. use MATLAB, which is helpful in image processing. The module in the referenced paper used a neural network technique with the TRAINSCG algorithm for sweetness determination. Ittatirut resized the images in half. Similar to Ittatirut's paper, the HML in this project uses cropped apple images. Different from Ittatirut's research, the HML use the image of the surface of an apple. The process of the HML is cost-effective and convenient.

E. Determination of Chokanan Mango Sweetness Using Non-Destructive Image Processing Technique[7]

The purpose of Bejo's research is to propose a model that determines the sweetness of Chokanan mango in a non-destructive way. Bejo *et al.* captured the images of the mango with the Keyence Machine Vision System. Bejo used the Pearson correlation to find the relationship between the color components and sweetness. Among the color components, the hue had the highest value of the Pearson correlation. Bejo's model is developed based on the hue using LR analysis. Therefore, the Pearson correlation will be used to find a feature that has the most signification relationship with Pixie Crunch's sweetness. For that reason, the LR model will be developed based on the feature.

REFERENCES

[1] H U.S. Department of Agriculture. World Production, Markets, and Trade Report (2022, Jun, 14), *Fresh Apples, Grapes, and Pears: World Markets and Trade.* [Online]. Available: https://apps.fas.usda.gov/psdonline/circulars/fruit.pdf

[2] Schwallier. P, "Checking apple maturity: What to look for," Michigan State Univ. Extension, Aug. 28, 2012. Accessed: Oct. 6, 2022. [Online]. Available: https://www.canr.msu.edu/news/checking_apple_maturity_what_to_look_for

[3] Kim, Seon-Jong, "Analysis of Apple Colors and Sugar Contents Using Linear Regression," in *Proc. The Journal of the Convergence on Culture Technology (JCCT)* Vol. 8, No. 1, pp. 201-207, Jan. 2022. [Online]. Available: https://koreascience.kr/article/JAKO202208148789847.view?orgId=anpor&hide=breadcrumb,journalinfo

[4] M. A. J. Al-Sammarraie *et al.*, "Predicting Fruit's Sweetness Using Artificial Intelligence—Case Study: Orange," in *Proc. Applied Sciences*, vol. 12, no. 16, p. 8233, Aug. 2022, [Online]. Available: https://doi.org/10.3390/app12168233

[5] Chu-Hui Lee and Jhih-Chen Jhou. "A Non-Invasive Method to Classify the Sweetness Levels of Apples," in *Proc. 2021 5th International C*

onference on Artificial Intelligence and Virtual Reality (AIVR) (AIVR 2021) Association for Computing Machinery pp. 128–134, July 2021. [Online]. Available: *https://doi.org/10.1145/3480433.3480453*

[6] T. Ittatirut, A. Lekhalawan *et al*., "Apple Sweetness Measurement by i mage processing technique, " in *Proc. 2016 Fifth ICT Int'l Student P roject Conf. (ICT-IS; PC*), pp. 182-185, May 2016, [Online]. Availab le: https://ieeexplore.ieee.org/document/751926

[7] S. K. Siti Khairunniza Bejo *et al*., "Determination of Chokanan Mang o Sweetness (Mangifera indica) Using Non-Destructive Image Proces sing Technique," in *Proc, Australian Journal of Crop Science (AICS)*, Jan. 2014, 8. 475-480, [Online]. Available: http://www.cropj.com/bej o_8_4_2014_475_480.pdf