Machine Learning: Determining Fruit Ripeness from Visual and Auditory Data

A Technical Report submitted to the Department of Computer Science

Presented to the Faculty of the School of Engineering and Applied Science
University of Virginia • Charlottesville, Virginia

In Partial Fulfillment of the Requirements for the Degree
Bachelor of Science, School of Engineering

William Tan
Spring, 2023

On my honor as a University Student, I have neither given nor received unauthorized aid on this assignment as defined by the Honor Guidelines for Thesis-Related Assignments

Ashish Venkat, Department of Computer Science

Machine Learning: Determining Fruit Ripeness from Visual and Auditory Data

CS4991 Capstone Report, 2023

William Tan
Computer Science
The University of Virginia
School of Engineering and Applied Science
Charlottesville, Virginia USA
wht7vk@virginia.edu

ABSTRACT

When it comes to telling whether a fruit is ripe or not, people often resort to unreliable familial wisdom or folk science to decide, resulting in inaccurate choices and dissatisfaction. A machine learning approach can provide both accurate and precise determinations on ripeness. We can achieve this by developing a mobile phone application that can take advantage of the phone's camera and microphone. We can then consider the appearance of a particular fruit via computer vision and process the sound of a user knocking upon the fruit via signal processing, and render a determination on ripeness. We anticipate that our application and approach would be accurate for most types of fruits, save for those that can vary greatly in size, hardness, and color. This would allow for more customers to be better informed upon the ripeness of these kinds of fruit. Additional work may be needed in order to accommodate both a wider variety of fruit and vegetables as well as a greater amount of variation in each individual type.

1. INTRODUCTION

In the case of some fruits, the ripeness cannot be determined at first glance or at first touch. Ripeness can be simple to determine in some cases. The more yellow and spotty a banana is, the riper it is. Similarly, the ripeness of fruits such as mangoes, peaches, and apples can be determined by a quick visual examination. However, other kinds of fruits, particularly those with hard rinds like a melon or a coconut, need to be opened to judge whether they are ripe or not, which could mean that a customer did not get the most satisfactory fruit for their money or, in the worst case, wasted their money entirely.

In an attempt to accommodate the potential of waste, customers have developed folk methods that aim to determine the ripeness of a fruit without needing to purchase and open it. These practices, such as checking for marks and examining the colors of rinds, which can border on the superstitious, can be of questionable accuracy. In addition, while some qualities of fruit do correlate with the fruit's ripeness, the tests to examine such qualities are vulnerable to subjectivity from the customer and therefore may not be useful.

One widely-practiced test of a fruit is of this nature. This test, in which a customer knocks upon the rind of a fruit and listens for any particularities in the sound, is used for many fruits, such as melons. As these fruits mature, their internal chemistry changes, which alters the sound they produce. This change can then be used to judge ripeness. Expanding upon this, we can bolster this method with the greater consistency and accuracy of software instead of having to depend upon more fallible human senses.

2. RELATED WORKS

Previous efforts to classify fruits' ripeness have been primarily focused on a single, specific fruit or a highly specific subset of fruits. For instance, visual classifiers were generally limited to fruits with appearances that were generally correlated with their ripeness, such as bananas, mangoes, and cherries [1]. Classifiers that depended on the sound of knocking on the fruit were also limited to a single fruit, such as pineapple [3], coconut [6], durian [7], and especially watermelon [10]. Both approaches have advantages and disadvantages in their respective fields: fruits such as watermelon and pineapple rarely give visual indicators of ripeness, while fruits such as apples and mangoes produce highly variable sounds when knocked upon. Both are discussed below.

2.1. Visual Classifiers

Naturally, a large proportion of efforts to classify ripeness depending on image processing concentrated on fruits with visibly different appearances depending on their stage of ripeness [1]. Fruits such as bananas, cherries, and mangoes would have their images and ripeness used to train convolutional neural networks (CNNs) or deep neural networks (DNNs). Such algorithms were highly accurate, with accuracy ratings of over 90%. In some fruits that did not obviously change their appearance, such as Durian [7], other alterations, such as stem shape, could be captured and utilized by classifying algorithms to accurately determine ripeness. However, when it came to fruits that did not change much in appearance, such as tomatoes, visual classifiers significantly suffered in performance [1]. An alternative approach would have to be used.

2.2 Auditory Classifiers

When it comes to classifying the ripeness of fruits that do not change significantly in appearance when ripening, such as watermelons [4], coconuts [5], or pineapples [3], most classification efforts focus on the sounds they produce. While it is possible for the ripeness of such a fruit to be determined via an image-processing approach, such an approach yields an accuracy of only 81-84% [4][8]. In contrast, an approach that uses audio data from knocking upon the fruit can yield a much higher degree of accuracy [10]. While Zeng, et al. (2014) developed a support vector machine (SVM) method of classification that is almost 90% accurate, more recent methods once again utilizing neural networks have boasted accuracy ratings above 95%, as demonstrated by Chen, et al. (2021).

2.3 Hybrid Approach

While more advanced classification methods have been developed in recent years, the vast majority of ripeness classification efforts remain focused on a single feature of fruit, a single species of fruit, or both, as found in a survey by Rizzo, et al. (2023). These previous classifiers, with their high specificity, can only find wide use in industrial or commercial contexts [2][4]. When it comes to the consumer, convenience and simplicity should be the key. A combined approach, where the two principal features of sight and sound are combined, should offer users a comfortably wide range of application.

3. PROPOSED DESIGN

The classifier is designed to use both image data and audio data, typically captured through a mobile phone's camera and microphone, respectively. This is because visual classifiers are effective when used on fruits that visibly ripen, while auditory classifiers are effective when used on fruits that do not. It will then process this data to make a determination on both fruit and ripeness. However, there are multiple valid approaches to building such a classifier, each of them with their benefits and downsides.

3.1. Classifier Objectives

The classifier would be designed with the following overarching objectives:

1. Take in a picture of a fruit and an audio recording of the fruit being knocked upon.
2. Determine what type of fruit the picture is of, or take such a type as an input from the user.
3. Return a specific "ripeness" value for the specific type of fruit identified.
4. Return confidence levels for each determination.
5. Be able to gracefully handle failures to achieve objective 2, objective 3, or both.

In this way, the user would be able to understand the function of this classifier, as well as the utility provided. The user would also be able to consider the confidence the classifier has in its determination and therefore be able to make informed decisions when it comes to evaluating a fruit's ripeness before purchasing it.

3.2. Single Neural Network

We believe the most ambitious approach would be to use a single neural network, likely a DNN, to classify both a fruit and the fruit's specific ripeness. We would design our classifier to take in an image of known size and an audio recording of known length, which would be a picture of a given fruit and the sound produced by knocking upon that fruit, respectively. The two inputs would then be combined into a single input. This singular input would then be input into the classifier, which would then output two values: the fruit type, if necessary, and the "ripeness" value.

There are upsides and downsides to this approach. The benefit to using a single neural network would be simplicity. Unlike the approach discussed in 3.3, a single neural network would not have to depend on the functioning of other neural networks in order to return an informative result. It would simply present its results, the confidence values associated with them, and the user would likely take all of those into account to make a decision. The biggest challenge in using a singular neural network, however, is its likely size as well as the computing power needed to train it. The single neural network, with the dual responsibilities to both determine a fruit and its level of ripeness would both likely be very large and must be trained on a very large dataset in order to accommodate a wide variety of fruits and their sounds and various levels of ripeness.

3.3. Multiple Neural Networks

Another approach is to depend on multiple, more specialized neural networks. In this approach, we would implement separate neural networks responsible for making determinations for specific fruit and ripeness. We would design our classifier to first determine what fruit the picture was of, and then use that result to send both the picture and audio to one or two fruit-specific neural networks. The results would then be combined into the same two values discussed above.

Once again, we have upsides and downsides. The most apparent benefit is that we lower the complexity of each neural network, which only have to deal with one feature (picture, audio) at a time instead of having to process multiple features at the same time. Multiple neural networks can be trained at once, which allows for us to develop functionality in a more efficient manner. However, a notable downside is that the neural networks responsible for classifying ripeness depend upon the neural network that identifies fruit for an accurate determination. This opens the possibility for an entirely erroneous result with erroneous confidence values, possibly misinforming the user.

4. ANTICIPATED RESULTS

Overall, each implementation of the classifier may have its own benefits and drawbacks, but a sufficiently trained classifier should be able to determine a particular fruit and its ripeness somewhat reliably in most cases. It should be able to perform at a higher level of accuracy than a classifier relying only on visual or auditory data, and it should be able to perform at a significantly higher level of accuracy than a human's own wisdom and judgement. Last, it should be able to help users and customers maximize their satisfaction by preventing them from wasting money on fruits that are not at a desirable ripeness.

5. CONCLUSION

We proposed a method for both identifying and classifying the ripeness of a particular fruit by combining two features often used individually to determine ripeness. As most consumers rely on potentially unreliable folk wisdom when it comes to checking fruits are ripe or not, this method could allow them to be much better informed and therefore waste less money. With recent developments in classifier algorithms, this previously complex task could likely be handled by today's neural networks such as CNNs and DNNs. While combining two features such as sight and sound can increase complexity, we believe recent advancements can accommodate such an increase and open the way forward for more comprehensive classifiers that work with multiple features at once.

6. FUTURE WORK

While this proposal assumes such a classifier can be trained, feasible training methods would need to be developed since using actual fruit in bulk quantity can become impractical. In addition, Rizzo, et al. (2023) maintain that most of the work when determining fruit ripeness still focus on a singular feature. However, the number of features continue to increase, including aspects of fruits such as fluorescence and aroma. Including those features into classifiers may increase accuracy and confidence further. Last, the concepts of using a single general neural network or multiple specialized neural networks present an interesting fork in the road when designing classifiers, and more exploration may be required in order to determine the benefits and drawbacks of each.

REFERENCES

[1]Muthulakshmi. A and P N. Renjith. 2021. Comprehensive Systematic Review on Fruit Maturity Detection Technique. *IEEE Xplore*, 1234–1240. DOI:https://doi.org/10.1109/ICESC51422.2021.9532772

[2]Delan Zoe H. Arenga, Jennifer C. Dela Cruz, and Delan Zoe H. Arenga. 2017. Ripeness classification of cocoa through acoustic sensing and machine learning. *2017IEEE 9th International Conference on Humanoid, Nanotechnology, Information Technology, Communication and Control, Environment and Management (HNICEM)* (December 2017). DOI:https://doi.org/10.1109/hnicem.2017.8269438

[3]Yenming J. Chen, Yeong-Cheng Liou, Wen-Hsien Ho, Jinn-Tsong Tsai, Chia-Chuan Liu, and Kao-Shing Hwang. 2021. Non-destructive acoustic screening of pineapple ripeness by unsupervised machine learning and Wavelet Kernel methods. *Science Progress* 104, 3_suppl (July 2021), 003685042211108. DOI:https://doi.org/10.1177/00368504221110856

[4]Kenji Contreras, Anel Henry, Danilo Cáceres-Hernández, and Javier E. Sanchez-Galan. 2022. Comparing Convolutional Neural Networks and Deep Metric Learning Methods for Classification of Export Watermelon (Citrullus lanatus) Varieties. *IEEE Xplore*, 1141–1146. DOI:https://doi.org/10.1109/ISIE51582.2022.9831572

[5]Jonathan Q De Leon, Flordeliza R Fernandez, and Charvin Kelsey Lacsina. 2022. Microcontroller Based Portable Coconut Tapping Mechanism With Audio Signal Filtering. *ECS Transactions* 107, 1 (2022). DOI:https://doi.org/10.1149/10701.20147ecst

[6]Nemilyn A. Fadchar and Jennifer C. Dela Cruz. 2020. A Non-Destructive Approach of Young Coconut Maturity Detection using Acoustic Vibration and Neural Network. *2020 16th IEEE International Colloquium on Signal Processing & Its Applications (CSPA)* (February 2020).

DOI:https://doi.org/10.1109/cspa48992.2020.9068723
[7]Weangchai Kharamat, Manop Wongsaisuwan, and Norrarat Wattanamongkhol. 2020. Durian Ripeness Classification from the Knocking Sounds Using Convolutional Neural Network. *IEEE Xplore*, 1–4. DOI:https://doi.org/10.1109/iEECON48109.2020.229571
[8]Marcus Guozong Lim and Joon Huang Chuah. 2018. Durian Types Recognition Using Deep Learning Techniques. *IEEE Xplore*, 183–187. DOI:https://doi.org/10.1109/ICSGRC.2018.8657535
[9]Matteo Rizzo, Matteo Marcuzzo, Alessandro Zangari, Andrea Gasparetto, and Andrea Albarelli. 2023. Fruit ripeness classification: A survey. *Artificial Intelligence in Agriculture* 7, (March 2023), 44–57. DOI:https://doi.org/10.1016/j.aiia.2023.02.004
[10]Wei Zeng, Xianfeng Huang, Stefan Müller Arisona, and Ian Vince McLoughlin. 2013. Classifying watermelon ripeness by analysing acoustic signals using mobile devices. *Personal and Ubiquitous Computing* 18, 7 (August 2013), 1753–1762. DOI:https://doi.org/10.1007/s00779-013-0706-7