Universal Journal of Agricultural Research 11(6): 1071-1076, 2023 http://www.hrpub.org
DOI: 10.13189/ujar.2023.110614

Automated Strawberry Ripeness Detection Using Convolution Neural Network

Wesly Septiandi Sidabutar, Jos Timanta Tarigan*, Amer Sharif

Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Indonesia

Received August 21, 2023; Revised October 17, 2023; Accepted November 20, 2023

Cite This Paper in the Following Citation Styles

(a): *[1] Wesly Septiandi Sidabutar, Jos Timanta Tarigan, Amer Sharif , "Automated Strawberry Ripeness Detection Using Convolution Neural Network," Universal Journal of Agricultural Research, Vol. 11, No. 6, pp. 1071 - 1076, 2023. DOI: 10.13189/ujar.2023.110614.*

(b): *Wesly Septiandi Sidabutar, Jos Timanta Tarigan, Amer Sharif (2023). Automated Strawberry Ripeness Detection Using Convolution Neural Network. Universal Journal of Agricultural Research, 11(6), 1071 - 1076. DOI: 10.13189/ujar.2023.110614.*

Abstract The demand of agricultural product has been steadily increased by times. This condition requires related industry to increase their productivity. Automation in smart farming is one of the currently discussed solutions to develop a productive and sustainable farming solution. In this research, we introduce a tool that might help the automation process in farming industry by speeding up the ripeness detection while increasing the accuracy. We designed and developed a system that is able to detect fruit's ripeness, specifically strawberry. It was built based on machine learning using CNN algorithms. We developed the model using Keras library. We performed the training by feeding 200 images of strawberry fruit with various ripeness levels that covers the entire growth cycle of the fruit. Additionally, these images were also taken from various level of distance to simulate harvester point of view during harvesting. Besides, we performed automated preprocessing to the image data by converting it to HSI color domain. We also built and evaluated the most optimal CNN architecture to retrieve the best training result. The trained model is fed into our own desktop-based application. The testing is performed by feeding more data to the application. The test data shows that our system was able to predict correctly 92-99% in HSI color spectrum compared to 61-67% accuracy in RGB format.

Keywords Smart Farming, Strawberry, Convolution Neural Network

1. Introduction

Optimizing the productivity of farming to ensure food security is one of the most discussed topics in recently. There are several methods that have been proposed or implemented such as optimizing yield prediction [1] or developing efficient agricultural development model [2]. Additionally, smart farming is one of the most discussed topics in recent years. The idea of smart farming is intended to revolutionize traditional agricultural practices by enhancing overall efficiency, productivity, and sustainability. It mostly involves the use of advanced technology to achieve the goal. One of the most discussed issues in smart farming is to reduce the cost of manpower that usually exists in traditional agricultural activities such as harvesting. This activity is mostly repetitive, time-consuming, and costly. Hence, many researches have aimed to produce a solution related to this issue such as robotic harvester, harvesting prediction, and fruit maturity detection.

Strawberry (Fragaria x anannassa) in particular, is one of the most popular fruits which has massive amount of demand throughout the world. Aside from the unique taste of it, strawberry is also claimed to be one of the healthiest fruits to consume [3]. However, the characteristic of strawberry requires an optimum timing to harvest the fruit to have the best quality result. The best method to judge the maturity of strawberry is to measure some of its criteria such as firmness, total soluble solids, titratable acidity, and total anthocyanins [4]. However, this method is mostly destructive as it requires the fruit to be squished to collect the data.

However, there is various research aimed at automated strawberry harvesting without extracting the fruit. Most of the methods use image classification to measure the ripeness of the fruit. Feng et al. developed a fruit detachment and classification method for strawberry harvesting robot [5]. It uses image processing to automatically detect strawberry fruit. Huang et al. presented a solution to specifically detect strawberry picking point during harvesting [6] and later used this solution to develop a robot-powered harvester in a complex farming environment [7]. They were able to develop a robot that learns how to harvest strawberry fruit from its tree; from detecting the fruit, deciding the picking point, and performing motion to harvest them.

Similar to these previous works, our research is aimed at developing an automated ripeness detection of strawberry. Our approach is to use machine learning to evaluate the colour and shape of strawberries. Although the RGB colour model can display colours well, it has several obstacles to displaying colours that are not available in the colour library. Image identification is usually easier when viewed in terms of the difference in hue values[8]. We used images from the fruit on the plant. These images will be pre-processed using the HSI (Hue-Saturation-Intensity) Colour Model due to its ability to increase accuracy in fruit-related image processing [9]. After the image has been changed to Hue Saturation Intensity, it will be classified using the Convolutional Neural Network algorithm. Additionally, we also observe how the viewing condition affects the performance of the proposed solution. We took images from harvester distance and without clearing the view between the camera and the fruit. This additional feature is important since it simulates the viewing distance of the farmers during the harvesting process. This method is usually used to solve difficult image-based object recognition tasks using an appropriate but simple architecture.

2. Methodology

2.1. Convolutional Neural Networks

Machine Learning is a subset of artificial intelligence that focuses on the ability of the system to improve its performance through data rather than built-in code. It is intended to learn and get better from additional experience and data. One of the types that are commonly used in machine learning solution is artificial neural networks (ANNs). ANN is a computational model that mimics the structure of the human brain. It consists of interconnected nodes with multiple layers with specific functions. Each connection has weight and bias values that adapt during training process. Theoretically, ANN can be trained to solve any domain based on its training process. However, its effectiveness and efficiency are varied and highly depend on the architecture.

Convolution Neural Network (CNN) is one of the most popular types of ANN. The idea of convolutional neural networks was originally coined with the design of LeNet by Lecun et al. [10]. It is defined as a collection of neurons arranged like an acyclic graph (a graph without loops in it). CNN has the characteristic that there is a hidden layer that is connected to a subset of neurons in the previous layer. Because of this, it enables the convolutional neural network to learn features effectively. CNN is able to produce filters that are taught for a specific purpose, for example, the initial layer usually focuses on understanding the edges of the image, then the next layer is more focused on understanding shapes, then the next layer is more focused on studying partial parts of objects, as seen as a little, some, or a lot, and the last layer is used to recognize objects[11].

CNN is composed of several layers with different types and functions. The most common types of CNN layers are as follows:

1 Convolutional: capture local patterns and features
2 Pooling: reduce spatial dimensions, usually called subsampling. Max Pooling is a type of pooling layer that is able to reduce the dimension while retaining the information in a local region.
3 Flatten: convert multi-dimensional values into one-dimensional vector
4 Dense: usually called fully connected layer, connects every neuron from previous to subsequent layers.
5 Dropout: prevent overfitting by ‘dropping out’ neurons during training

2.2. HIS Color Spectrum

HSI (Hue, Saturation, Intensity) is a color space developed to be more intuitive in color manipulation and designed to predict how humans perceive and interpret color. HSI was developed when colors needed to be determined manually and is rarely used nowadays because users can choose a color pantone. This color space is commonly studied for historical purposes or an interest in history. The HSI color space is the best for traditional image processing functions such as convolution, equalization, histograms, etc., which are operated using manipulation of brightness values because intensity is equally dependent on R, G, and B [12].

HSI describes colours as having hue, saturation, and intensity. Where hue serves to reveal the true colours like red, yellow, and green. Hue is used to distinguish colours such as the redness and greenness of light. Hue relates to the wavelength of light. Saturation is used to express how pure the colour is and to state how much white is distributed among the colours. Intensity is the part that indicates how much light the eye gets[13]. Figure 1 shows the geometric representation of HSI color spectrum and its relation to other color model [14].

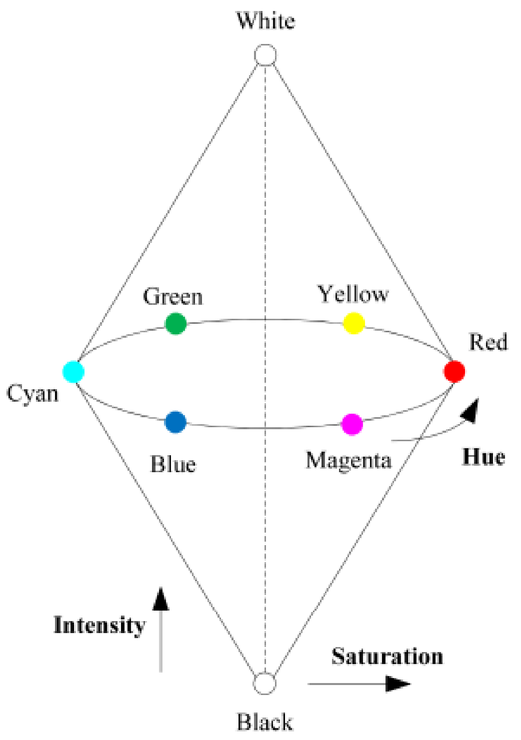

Figure 1. Geometric Representation of HIS Color Spectrum

3. Proposed Solution

Our objective is to develop a system that can automatically detect strawberry maturity based on its color. We built the detection system using machine learning as its core. The engine is built using convolutional neural network as its engine. The system uses training models received from the training process that has been done previously. The overall architecture of our system can be seen in figure 2.

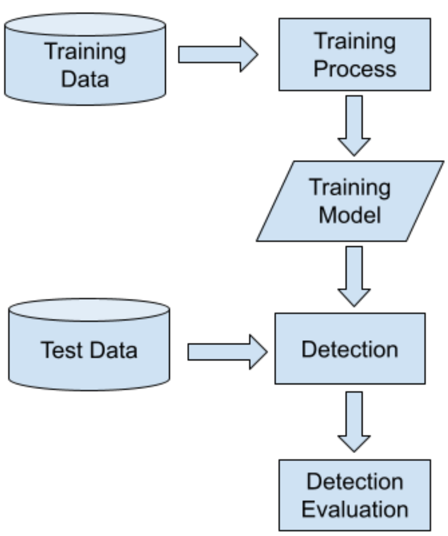

Figure 2. System Architecture

The first step to build the detection system is to train the model using the training dataset. We collected the data using nearby strawberry farm in Brastagi Grand Forest Park, Sumatera Utara, Indonesia. This area is famous for its strawberry farming tourism. In this area, most of strawberries are grown in open field covered with silver mulch plastic. Figure 3 below shows the common strawberry plantation technique in Brastagi.

Figure 3. Common strawberry plantation in Brastagi. Source: [15]

We collected 300 strawberry images which are divided into 150 images of unripe/immature strawberries and 150 images of ripe/mature strawberries. Furthermore, we also collect 2 kinds of different sets of images based on the distance of the camera while capturing the image. The first set was taken from a short distance, focusing on the fruit. The second set was taken from a longer distance to emulate the viewing position of the pickers during harvesting. We define the distance between 60-70 cm from the plant (compared to 20-30 cm in the previous samples). We found this set of data is necessary since the decision-making during harvesting is done at this position where the leaves and branches obstructed part of the fruit. Hence, in total, there are 4 categories of images in our dataset.

4. Implementation

The image datasets were captured using a smartphone camera and converted into images of 500×500. The image is then converted into HSI spectrum by our system. We split the datasets into 2 categories based on their purposes; training and validation with the ratio of 8:2 respectively. Figure 4 shows the example of images from each category and its HSI counterpart.

Figure 4. Image data in RGB (Top) and HSI (Bottom)

We generated the model by using a simple phyton-based script. We used Keras library to simplify building the CNN. The architecture of the CNN model is shown in table 1.

Table 1. The Architecture of CNN Model

Layer Type	Size
Convolutional	148, 148, 16
Max Pooling	74, 74, 16
Convolutional	72, 72, 32
Max Pooling	36, 36, 32
Convolutional	34, 34, 64
Max Pooling	17, 17, 64
Flatten	18496
Dense	200
Dropout	200
Dense	500
Dropout	500
Dense	1

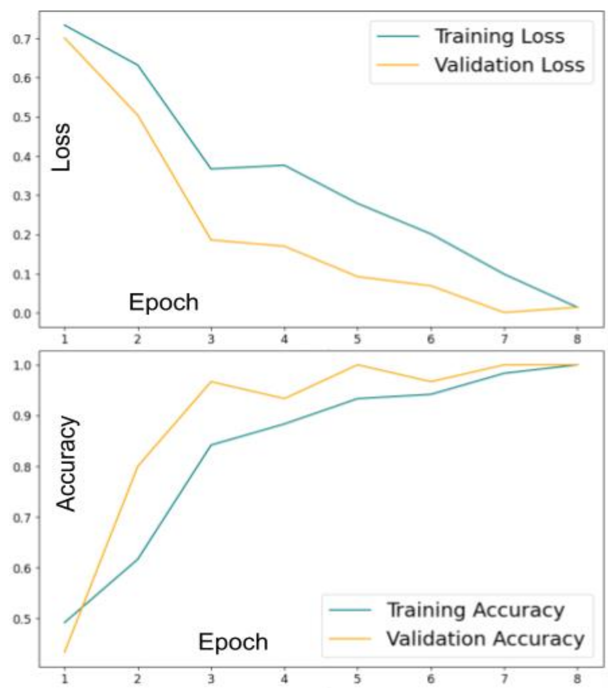

Figure 5. The Loss (loss) and Accuracy (right) during training

The training process is done by feeding the training data to generate the model. We run the training on a laptop with AMD Ryzen 5 4600H CPU and nVidia Geforce GTX 1650 TI with 3949MB VRAM GPU. During the training, we limited the epoch to 30. In machine learning, accuracy is a measure of how well a model correctly predicts the labels of the given dataset and loss represents the error in the predictions made by the model. In training a machine learning, the objective is to increase the accuracy and lower the loss. The training process can be shown in figure 5. However, after 8 iterations (or epoch 7), the training reached the desired accuracy and loss. The training was completed in less than 20 seconds. We then generate the machine learning model using Hierarchical Data Format version 5 (H5 or HDF5 [16]) to be used in the predictor application.

To test the performance of our model, we developed a simple desktop application that aimed to detect the maturity of strawberry using its image. The main interface of the application is shown in figure 6. The application is intended to help users upload an image and it will detect the ripeness using the model trained earlier. The system accepts image file with .jpg and .png format with various sizes. However, it is recommended by the application to use a 500×500 image to avoid distortion due to image rescaling.

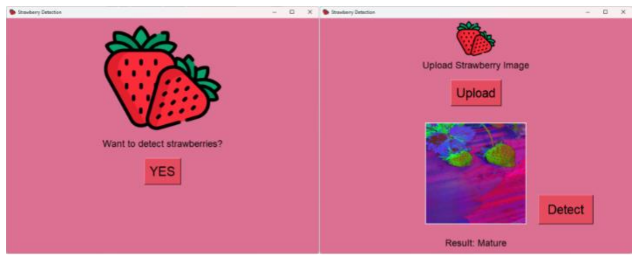

Figure 6. The interface of applications of proposed system

5. Test and Results

The testing is carried out twice; the first test was done by using an RGB image and the second one by HSI. We decided to perform the test this way to see how the HSI conversion can increase the accuracy of the system. Additionally, we also performed different tests for images that focused on the fruit and images that were taken from harvester's point of view. Each of these categories contains 50 images. To measure the performance of the system, we calculate the accuracy of the detection. This value can be measured by using this formula.

$$accuracy = \frac{correct\ \ prediction}{total\ \ prediction} \times 100\% \quad (1)$$

Table 2. The Result from RGB images focus on the fruit (not highlighted) and from distance (grey highlighted)

Actual	Prediction		Total
	Mature	Immature	
Mature	35	15	50
Immature	24	26	50
Mature	34	16	50
Immature	17	33	50

Table 2 shows the result from the test with the RGB images that focus on the fruit and taken from harvester's point of view. Overall, the system gives 61% accuracy with a better prediction result for mature fruit. The test with images taken from distance gives similar results. The system gives 67% overall accuracy with a better, but similar, precision for mature fruit compared to immature.

In the second test, we switched to HSI images. Table 3 shows the result of the test. The data shows a significant improvement over the previous test where the system was able to give 99% accuracy for images focusing on the fruit and 92% for images taken from harvester's point of view.

Table 3. The Result from HSI images focuses on the fruit (not highlighted) and from distance (grey highlighted)

Actual	Prediction		Total
	Mature	Immature	
Mature	49	1	50
Immature	0	50	50
Mature	45	5	50
Immature	3	47	50

Our test shows that the proposed system can produce good strawberry ripeness detection. Additionally, the image in HSI color spectrum also gives a very high accuracy in both focus and distance tests. Nevertheless, the system also produced both false positives and false negatives result. Figure 7 shows some of the images that give false results during the test.

Figure 7. The samples with distorted color (left) and multiple objects (right)

Based on our observation, the false predictions are mostly caused by the color distortion, either by sunlight specular shading or by shadows from the surrounding leaves. Additionally, our system is limited to analyze the whole image which means that if the image contains more than one strawberry, it will also affect the result. This condition mostly happened during the distance test for the immature prediction.

6. Conclusions

In this paper, we have introduced a CNN model that is able to detect strawberry maturity. We trained the model using 300 images and performed the training using RGB and HSI colour spectrums. We also perform the test using images taken from harvester's point of view to simulate the system's ability to detect the ripeness from distance. The results gave a satisfying result with 92-99% accuracy using HSI colour spectrum.

For future development, our intention is to build a fully automated fruit harvesting system that is able to minimize manpower in farm industry. Our next approach is to implement an object detection so our method can detect and analyze multiple strawberries. This feature will increase the usability of the system since it is able to analyze different fruits. We also would like to build a real-time detection system that is attached to a vision-based peripheral or robotic system to fully simulate an automated fruit harvesting system.

REFERENCES

[1] S. K. Jha *et al.*, "Sugarcane Yield Prediction Using Vegetation Indices in Northern Karnataka, India," *Univers. J. Agric. Res.*, vol. 10, no. 6, pp. 699–721, Dec. 2022, doi: 10.13189/ujar.2022.100611.

[2] U. Maman *et al.*, "Initiating Model of Agricultural Development Based on Local Wisdom: The Case of Risk Mitigation of Organic Rice Farming in Indonesia," *Univers. J. Agric. Res.*, vol. 10, no. 5, pp. 449–463, Oct. 2022, doi: 10.13189/ujar.2022.100501.

[3] F. Giampieri, S. Tulipani, J. M. Alvarez-Suarez, J. L. Quiles, B. Mezzetti, and M. Battino, "The strawberry: Composition, nutritional quality, and impact on human health," *Nutrition*, vol. 28, no. 1, pp. 9–19, Jan. 2012, doi: 10.1016/j.nut.2011.08.009.

[4] C. Liu *et al.*, "Application of Multispectral Imaging to Determine Quality Attributes and Ripeness Stage in Strawberry Fruit," *PLOS ONE*, vol. 9, no. 2, p. e87818, Feb. 2014, doi: 10.1371/journal.pone.0087818.

[5] G. Feng, C. Qixin, and N. Masateru, "Fruit Detachment and Classification Method for Strawberry Harvesting Robot," *Int. J. Adv. Robot. Syst.*, vol. 5, no. 1, p. 4, Mar. 2008, doi: 10.5772/5662.

[6] Z. Huang, S. Wane, and S. Parsons, "Towards Automated Strawberry Harvesting: Identifying the Picking Point," in *Towards Autonomous Robotic Systems*, Y. Gao, S. Fallah, Y. Jin, and C. Lekakou, Eds., in Lecture Notes in Computer Science. Cham: Springer International Publishing, 2017, pp. 222–236. doi: 10.1007/978-3-319-64107-2_18.

[7] Z. Huang, E. Sklar, and S. Parsons, "Design of Automatic Strawberry Harvest Robot Suitable in Complex Environments," in *Companion of the 2020 ACM/IEEE International Conference on Human-Robot Interaction*, Cambridge United Kingdom: ACM, Mar. 2020, pp. 567–569. doi: 10.1145/3371382.3377443.

[8] E. Law *et al.*, "The Science of Citizen Science: Theories, Methodologies and Platforms," in *Companion of the 2017 ACM Conference on Computer Supported Cooperative Work and Social Computing*, in CSCW '17 Companion.

New York, NY, USA: Association for Computing Machinery, Feb. 2017, pp. 395–400. doi: 10.1145/3022198.3022652.

[9] Y. Wang, Y. Cui, H. Huang, S. Chen, P. Zhang, and G. Q. Huang, "Study on HSI color model-based fruit quailty evaluation," in *2010 3rd International Congress on Image and Signal Processing*, Oct. 2010, pp. 2677–2680. doi: 10.1109/CISP.2010.5647943.

[10] Y. Lecun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, Nov. 1998, doi: 10.1109/5.726791.

[11] I. Cholissodin, "Buku Ajar AI, Machine Learning & Deep Learning," *Mach. Learn.*, 2020.

[12] K. Jack, *Digital Video and DSP: Instant Access*. in Instant Access. Elsevier Science, 2008. [Online]. Available: https://books.google.ne/books?id=9xQbHjZwfz4C

[13] E. R. Swedia and M. Cahyanti, "Algoritma Transformasi Ruang Warna," *Depok Univ Gunadarma*, 2010.

[14] D. Sun, G. Lu, H. Zhou, Y. Yan, and S. Liu, "Quantitative Assessment of Flame Stability Through Image Processing and Spectral Analysis," *IEEE Trans. Instrum. Meas.*, vol. 64, pp. 1–1, Dec. 2015, doi: 10.1109/TIM.2015.2444262.

[15] Kover Magazine, "Petik Sendiri, hingga Pemandangan Indah Kebun Stroberi Berastagi," Kover Magazine. Accessed: May 24, 2023. [Online]. Available: https://kovermagz.com/petik-sendiri-hingga-pemandangan-indah-di-kebun-stroberi-berastagi/

[16] M. Folk, G. Heber, Q. Koziol, E. Pourmal, and D. Robinson, "An overview of the HDF5 technology suite and its applications," in *Proceedings of the EDBT/ICDT 2011 Workshop on Array Databases*, in AD '11. New York, NY, USA: Association for Computing Machinery, Mar. 2011, pp. 36–47. doi: 10.1145/1966895.1966900.