Image Processing System for Early Detection of Cocoa Fruit Pest Attack

Basri*, Harli, Indrabayu, Intan Sari Areni
Universitas Al Asyariah Mandar, Polewali Mandar, West Sulawesi, Indonesia

Rosmawati Tamin
Universitas Hasanuddin, Makassar, South Sulawesi, Indonesia

*Corresponding author: basri@unasman.ac.id

Abstract. Cocoa (*Theobroma cacao* L.) is one of the leading commodities that developed in quality and quantity. Efforts through socialization and counseling in the field of cocoa cultivation for the farmers are always reminding know the importance of identifying symptoms from pest attacks could earlier makeup so that preventive actions that do not damage the environment or agricultural land. The study aims to create an early detection system based on image processing on symptoms of pest attacks on cocoa fruits. The early detection system involves training the data and test the data as a result of capture pictures of ordinary cocoa fruits and reviews those attacked by pests in real time at the location of cocoa plantations. Image processing techniques are integrated into the application software to be able to identify the pixel characteristics of capture images of cocoa fruits. The results of the study showed that the ability of the system to detect the symptoms of pests in training the data was 100% and the test of data was 70%. This result showed that the applications could be recommended to be developed on a larger scale so that it will be helpful for cocoa farmers.

Keywords: image processing system, cocoa fruit, pest attack

1. Introduction
Cocoa (*Theobroma cacao* L.) is one of the plantation crops that need attention for development [1], and it is one of the leading commodity in Indonesia. For example, cocoa plantation subsector in West Sulawesi province has a strategic role in the economy for employment into the main population by 49.69% with a contribution to the GDP [2] amounted to 21.98 [3]. Also, cocoa is now the third-largest non-oil source of foreign exchange after rubber and oil palm. According to data from the International Cocoa Organization, global cocoa demand growth of about 2-4% per annum [4]. According to statistics, Indonesian cocoa plantation area in 2016 was 1,722,315 ha with a total production reached 760,429 tonnes. Indonesia is the third largest cocoa producer after Republic of *Côte d'Ivoire* and Ghana. However, the productivity of the cocoa plant approximately 865 kg/ha/year. That is still very low compared to the potential of the cocoa plant that can reach 1500-2000 kg/ha/year. Low productivity of cocoa at the farm level, as an impact by pests and diseases [5]. Pests and diseases can reduce the productivity of cocoa up to 5-80% [6].
The effort to handle the problem, many things have been done by the government, but with the application of an expert system based on image processing, it will be beneficially for farmers to early detection of pests and diseases so that preventive and curative actions before severe damage can be done. Also, involving information technology in the field of agriculture will undoubtedly have an impact on the development of agricultural business [7]. Scanning expertise into programming is expected to facilitate cocoa farmers to access information about the symptoms of the cocoa plant by pests and diseases in their respective regions [8]. The expert system that initially in this study, integrated of image processing techniques to help identify the initial symptoms of cocoa fruit pest attacks, so that, an expert system integration based on image processing technology will facilitate the cocoa farmers to know pests on cocoa earlier before symptoms get worse.

2. Literature Review

2.1. Pest impact of cocoa plantation

Pests are a nuisance from the cultivation of crops. Essential pests that often attack cocoa plants are cocoa fruit borer (Conopomorpha cramerella Snellen), ladybug cocoa fruit (Helopeltis), stem/branch borer (Zeuzera coffee), mice and squirrels [9]. Helopeltis sp. as a cocoa pest can damage cocoa fruit quickly. This pest attack can reduce the production of crops up to 50-60% of the crops production. Characteristics of Conopomorpha cramerella Snellen is including moths that have the smallest size among members of the order Lepidoptera, the mature insects often lay eggs on the surface of the fruit, and the most preferred fruit for laying eggs is the most fruit grooves on the surface with size more than 5 cm long [3]. One of the most disturber pests is the ladybug pest. These pest symptoms are difficult to distinguish from the early symptoms of fruit rot because they both display black patches. Conopomorpha cramerella Snellen attacks cause the death of seed placental tissue so that the seeds cannot develop completely and then become sticky. Attacks on young fruit result in higher yield loss because the fruit will experience early damage so that the fruit cannot be harvested. Figure 1 shows one form of pest attack that shows the characteristics of black spots on the skin of cocoa pods.

![Figure 1. Symptoms of Helopeltis sp. pests display black patches](image)

The highest attack on cocoa pests is to overcome and prevent the growing imago from laying eggs on the cocoa skin tissue and the leaf buds. Not only that, the adult imago will mate and continue by placing the egg of the sapling which can damage the cocoa fruit and the cocoa tree network. In the process of marriage, this pest can produce 200 eggs during its lifetime. This condition is classified as very fast inbreeding so that early eradication must be conducted immediately. Consequent of this condition crack cocoa fruit and cause abnormal growth or malformation. The worst impact without handling, the fruit will dry and then die, the cocoa beans will dry out. It is evident that the production of cocoa fruit will decrease dramatically with many fruits attacked by pests.

2.2. Implementation of image processing techniques in agriculture

Image Processing is a process of improving image quality so that humans or computers easily interpret it. Image processing techniques are conducted by transforming the image into another image, for example, image compression. Operations performed in image processing can be classified into several
types, namely image restoration, image quality improvement, image segmentation, image analysis, and image reconstruction [10].

Image processing implementation in the study of early identification of pest attack in cocoa fruit has never been done before. In the field of Agriculture, image processing techniques have been done to support cocoa cultivation like the program determining the maturity of cocoa fruit using the Euclidean Distance algorithm method to determine the suitability value of the test image and sample image [11]. This program can provide cocoa fruit farmers who are still in the process of starting a cocoa plantation business and who still have little knowledge in determining the maturity of the cocoa fruit. Meanwhile, in identifying types of defects in cocoa fruit using Backpropagation Artificial Neural Network method with parameters alpha = 0.6, fault tolerance = 0.0001, and target = 0.9, obtained an accuracy rate of (76%) in determining the quality of cocoa beans [12]. Research on cocoa fruit was then carried out by Ardiansyah, Y. in 2016 who developed an application to identify the variety of maturity of cocoa fruit using image processing and fuzzy logic methods, where the identification process used rules that have been created to find the smallest distance between input data and database [13]. In that study, the color image of the cocoa obtained was classified according to the cocoa maturity class and processed into a dataset to process training data with an accuracy rate of 60%.

Research that implements image processing techniques in other fields of agriculture which classified mango based on internal quality during cold storage (low temperature) [14]. In that study used processing techniques using unsupervised PCA and pre-processing spectra as well as the Mahalanobis Distance technique to validate the classification results. Image processing-based identification system for cocoa fruits image condition conducted by Yogiswara, et al., which develop an application for identification of the disease of cocoa fruits based image capture and compared with database [15]. Same of a system that identified the quality of rice physically using digital image processing techniques [16].

Digital image processing techniques in the field of agriculture are widely implemented. In this study, the application of image processing techniques will adjust the type of data and parameters. Image of cocoa fruit as shown in Figure 1 as the object of research. It has the same characteristics in objects that will be identified in previous studies, so the potential method for applying image processing techniques will be simplified to do. Image processing techniques using features of object-based [17] classify with Artificial Neural Networks [18], feature extraction techniques [19], underlie this study. While in the preprocessing stage of the binary cocoa image data can be segmented with pixel analysis techniques and can be analyzed the localization of object parts [20]. Furthermore, it can also be done with Region of Interest Segmentation Techniques [21], [22], while the classification of objects can be done using Singular Value Decomposition Technique [23], and the Super Vector Machine (SVM) method [24].

2.3. Gabor filter method

Gabor filter in this study was used as a filter capable of simulating the characteristics of the human visual system [25] in isolating certain frequencies and orientations from the image of cocoa fruit. This characteristic was used for texture recognition applications in computer vision. The implementation of Gabor Filter is currently widely used in recognition applications [26]–[30] while in this study Gabor Filter was implemented in the feature recognition of pest attack features on the image of cocoa fruits. Spatially, a Gabor function is Sinusoidal on certain frequencies and orientations that are modulated by the Gaussenvelope function. The rate and orientation define the location of the filter center. All Gabor Filters with varying wavelengths and orientations are applied to one particular point (x, y), then there are many filter responses for that point. Orientation and frequency will have different values in different cases. To filter fruit images, 5 frequencies are used (u = 0, 1, 2, 3, 4) and 8 orientations (Θ = 0, 1, 2, 3, 4, 5, 6, 7) as the calculation formula for generating the Gabor kernel as follows:

\[
G(x, y, \Theta, \mu, \sigma) = \frac{1}{2 \pi \sigma^2} \exp \left( -\frac{x^2 + y^2}{2 \sigma^2} \right) \exp \left\{ 2 \pi i (u \cdot x \cos \Theta + u \cdot y \sin \Theta) \right\}
\]  

(1)
Where \( i = \sqrt{-1} \) is the frequency of the sinusoidal wave, \( \omega \) is the control of the orientation of the Gabor function and \( \sigma \) is the standard deviation of the Gaussian Envelope, and \( X, Y \) is the coordinates of Gabor filter. The formation of the Gabor filter is divided from the Gaussian Envelope component as well as equation (2) and Sinusoidal waveform as in equation (3).

\[
g(x, y) = \frac{1}{2 \pi \sigma^2} \exp \left\{ -\frac{x^2 + y^2}{2 \sigma^2} \right\}
\]

\[
s(x, y) = \exp\{2 \pi i (u \cdot x \cdot \cos \theta + u \cdot y \cdot \sin \theta)\}
\]

3. Methodology

3.1. Data source

Primary data that used in this study were images of cocoa fruit. It was obtained at the center of cocoa plantations. Figure 2 shows the variation of data based on the database training and validation dataset. The image presented in Figure 2 is a sample of data that will be processed by the designed application. The amount of data used in this study as shown in Table 1.

![Figure 2. Samples of image coca fruits, normal (a) and pest attack (b)](image)

Table 1. Data source

| Data type       | Training | Testing |
|-----------------|----------|---------|
| Normal          | 20       | 10      |
| Pest Impact     | 40       | 10      |

The data processed as Table 1 consists of two types of data classification, namely data with the classification of usual cocoa fruit and cocoa fruit identified as being attacked by pests. Figure 3 shows the data retrieval performed using a camera with a resolution above 5 MP, f / 2.0 with a camera distance of 25 centimeters from cocoa as an object.

![Figure 3. Object capture design](image)
3.2. System framework

System development as shown in Figure 4 was made based on a system framework that was designed.

![System Framework](image)

**Figure 4.** System Framework

The system framework consists of data input processes, object recognition processes and system outputs. The stages started with the data retrieval process as the data retrieval technique in Figure 3. The cocoa fruit that processed becomes an image (a), using a camera (b), then collected and prepared using preprocessing techniques to change the raw data based Red Green Blue (RGB) becomes Grayscale level data with a size of 92x112 pixels at 72 DPI color depth. Furthermore, data based on Grayscale color was processed using the Convolution with Gabor Kernel and Feature Extraction process techniques.

In processing the training data, all data that has been extracted, stored in the database. Meanwhile, for the recognition system, it will compare the new input data that has been processed with the existing database. To ensure the development of the system runs well, a waterfall approach model was used so that the system can be designed according to system development standards [31]. This approach was conducted according to the stages of system development starting with (a) Requirements analysis and definition, (b) System and software design, (c) Implementation and unit testing, (d) Integration and system testing, and (e) Operation and maintenance [32].

3.3. Testing Method

Testing models for the results of expert recognition for normal and pest attack condition, conducted separately to measure its performance in training and testing data collection. This test measures the accuracy of the two types of data. Accuracy measurement was done using Receiver Operating Character (ROC) technique as well as equation (4) [33].

\[
ACC = \frac{\sum TP + \sum TN}{\sum TP + \sum TN + \sum FP + \sum FN}
\]

Where ACC is a formula for measuring accuracy, TP is the number of the real image cocoa fruits that accurately detect objects. Conversely, False Positive (FP) is the number of unwanted objects that are calculated by the system as a real object of cocoa. False Negative (FN) is the number of the actual image cocoa fruits that fail to recognize as a necessary object, while True Negative (TN) is the number of the unwanted image cocoa fruits that identification as an unwanted object. In this study, TN values were not used.
4. Results

4.1. Early detection system

Early detection of cocoa fruit that was identified as normal and attacked by pests was done through applications that were built using the Matlab Programming Application. The system development process was made for application management, starting from inputting the data capturing from the camera, then there was a preprocessing feature to change the file at the RGB color level to Grayscale level. All data received then processed, and the folder management was arranged to facilitate the classification process.

![Early detection system for cocoa pest impact](image)

**Figure 5.** Early detection application

![Data flow diagram](image)

**Figure 6.** Data flow diagram
Figure 5 is a screenshot of the development of the designed application interface. An interface is a prototype application that is made based on system requirements analysis as the flowchart of the system development in Figure 6. The system development is useful for each user to enter data for preprocessing stage. The output of the data was Grayscale image data with PGM format according to the amount of data processed as Table 1. The process for identifying cocoa fruit images was done by comparing the vector image distance values generated by the convolution process by Gabor Filter. This benchmarking aims to measure the extent of similarity between new objects that will be detected with images that have been stored in the database. System testing results using the white box testing as presented in the following flowgraph as a sample.

![Flowgraph System](image)

**Figure 7. Flowgraph system**

The summary of the results for Independent Path (IP), Region (R) and Cyclomatic Complexity (C), as shown in Table 2.

| Flowgraph Type | Independent Path | Region | Cyclomatic Complexity |
|----------------|------------------|--------|-----------------------|
| Preprocessing  | 4                | 4      | 4                     |
| Recognition    | 6                | 6      | 6                     |
| Total          | 10               | 10     | 10                    |

Based on the white box test results as presented in Table 2, the calculation of the number of (R), (CC), and (IP) is the same value, so that it can be concluded that the application was free from logic errors.

4.2. Algorithm testing result

Algorithm testing was conducted to ensure that the algorithm implemented in the development of the system can run well. This test used a black box testing approach [34]. This test provides that every code that was implemented can run properly according to the expected system functionality [35]. The results of application testing based on equation (4) obtained accuracy as shown in Table 3.
Table 3. Testing results

| Data type       | Training | Testing |
|-----------------|----------|---------|
| Normal          | 100%     | 100%    |
| Pest Impact     | 100%     | 70%     |

The table shows data on the condition recognition of cocoa fruit in data testing. The testing process was carried out on two types of data, namely training data and test data. Training data is classified as processed data and stored in a database as data that has been positively known according to the analysis of cocoa experts. Meanwhile, data testing is data that has been analyzed and verified by experts but not stored in the database.

Table 3 shows that all data were tested after data training process. The system was able to recognize both the classification of normal cocoa fruits and those attacked by pests. In the training data for both conditions the cocoa fruit obtained 100% accuracy. System accuracy testing on training data achieves maximum results so that it can be said that the system was valid in the testing stages. Although the classification of cocoa fruit on the test data which was classified as normal cocoa fruit achieves a maximum accuracy, the test data of cocoa fruit attacked by Pest can only be identified by 70%.

Development systems are designed using the waterfall model have been tested and successfully demonstrated a complete process. Stages of system analysis spawned the application interface, as shown in Figure 5. To ensure that the system pass from the logic error, white box testing indicate a positive result as shown in Table 2. Likewise, on the algorithm performance with Black Box testing, reach positive result too. The accuracy of the system as presented in Table 3 indicate that the system has worked. Although the accuracy was still showing the value that has not been perfect, the image processing application for the introduction of the cocoa fruit condition can run as expected.

5. Conclusions
The result of the system development performed in this study showed that the system could be used as a recommended premier application for the early detection system of cocoa fruits condition. This result concludes based on the process of applying white box testing and algorithm testing. Tests using a white-box approach shows that the system was free from errors of logic, while the black box approach proved the functionality of the system. Also, from the results of the test algorithm development system, showed that the ability of the system to detect the attack of pests condition of cocoa fruits in training the data reached 100% and the test of data by 70%. This result will undoubtedly continue to be improved, especially in increasing accuracy through the hybrid algorithm.

Acknowledgments
This work was supported by the research team at the Research Laboratory of the University of Al Asyariah Mandar, and Ministry of Research, Technology and Higher Education of Republic Indonesia under F. Number 113/SP2H/LT/DRPM/2018.

References
[1] I. M. Fahmid, H. Harun, M. M. Fahmid, and N. Busthanul, “Competitiveness, production, and productivity of cocoa in Indonesia,” in IOP Conference Series: Earth and Environmental Science, 2018, vol. 157, no. 1, p. 12067.
[2] T. Callen, “What Is Gross Domestic Product?,” Finance Dev., pp. 48–49, 2008.
[3] K. Pertanian, “Buku Panduan Teknis Budidaya Tanaman Kakao (Theobroma Cacao L),” Gerak. Nas. Peningkatan Produksi dan Mutu Kakao. Direktorat Jenderal Perkeb., 2011.
[4] ICCO, International Cocoa Organization, 2008/2009. London: International Cocoa Organization (ICCO), 2009.
[5] D. Molden, T. Oweis, P. Steduto, P. Bindraban, M. A. Hanjra, and J. Kijne, “Improving agricultural water productivity: Between optimism and caution,” Agric. Water Manag., vol. 97, no. 4, pp. 528–535, 2010.

[6] D. P. Sudjatmiko, “Menjadikan Kakao Sebagai Pondasi Ekonomi Daerah Sulawesi Barat. Seminar Peningkatan Daya Saing dan Nilai Tambah Kakao Indonesia,” in Proceeding Seminar Peningkatan Daya Saing dan Nilai Tambah Kakao Indonesia, 2015, pp. 9–19.

[7] S. Wolfert, L. Ge, C. Verdouw, and M.-J. Bogaardt, “Big data in smart farming—a review,” Agric. Syst., vol. 153, pp. 69–80, 2017.

[8] A. Armansyah and D. Y. Prasetyo, “SISTEM PAKAR IDENTIFIKASI HAMA DAN PENYAKIT TANAMAN JAGUNG BERBASIS WEB (STUDI KASUS: DINAS TANAMAN PANGAN DAN HORTIKULTURA KAB INHIL),” SISTEMASI, vol. 5, no. 3, pp. 1–7, 2018.

[9] R. Adu-Acheampong et al., “Strategy for insect pest control in cocoa,” Am. J. Exp. Agric., vol. 6, no. 6, p. 416, 2015.

[10] J. C. Russ, The image processing handbook. CRC Press, 2016.

[11] F. Amaluddin, “Fitroh Pemanfaatan Pengolahan Citra Pencocokan buah Kakao menggunakan Euclidean Distance,” Gener. J., vol. 1, no. 1, pp. 43–48, 2017.

[12] S. Nurmuslimah, “Implementasi Metode Backpropagation Untuk Mengidentifikasi Jenis Biji Kakao Yang Cacat Berdasarkan Bentuk Biji,” Netw. Eng. Res. Oper. [NERO], vol. 2, no. 2, 2016.

[13] Y. Ardiansyah, “Aplikasi Identifikasi Mutu Kematangan Buah Kakao menggunakan Image Processing dan Metode Fuzzy Logic,” 2015.

[14] Y. T. Suci, I. W. Budiastra, and Y. A. Purwanto, “Penggolongan Mangga cv Arumanis Berdasarkan Mutu Internal,” J. Keteknikan Pertan., vol. 3, no. 2, 2016.

[15] G. H. Yogiswara, R. Magdalena, and H. F. T. S. Putra, “Identifikasi Jenis Penyakit Pada Kakao Dengan Pengolahan Citra Digital Dan K-nearest Neighbor,” proceedings Eng., vol. 3, no. 1, 2016.

[16] A. A. Nurcahyani and R. Saptono, “Identifikasi Kualitas Beras dengan Citra Digital,” Sci. J. Informatics, vol. 2, no. 1, pp. 63–72, 2016.

[17] S. Anraeni, I. Nurtanio, and I. Indrabayu, “Detection of Kidney Organ Condition Using Hidden Markov Models,” Indones. J. Electr. Eng. Comput. Sci., vol. 15, no. 2, pp. 294–300, 2015.

[18] A. A. Ilham, R. Hasanuddin, and D. M. Putri, “Wavelet analysis for identification of lung abnormalities using the artificial neural network,” in Electrical Engineering and Informatics (MICEEI), 2014 Makassar International Conference on, 2014, pp. 156–160.

[19] S. A. A. Bowo, A. Hidayatno, and R. R. Isnanto, “Analisis deteksi tepi untuk mengidentifikasi Pola daun.” Diponegoro University, 2011.

[20] V. K. Bakti, D. Dairoh, and M. Huda, “Segmentasi Dan Perbaikan Citra Untuk Proses Pengukuran Dimensi Beras,” J. Infotel, vol. 8, no. 1, pp. 88–93, 2016.

[21] Basri, Indrabayu, and A. Achmad, “Gaussian Mixture Models optimization for counting the numbers of the vehicle by adjusting the Region of Interest under heavy traffic condition,” 2015 Int. Semin. Intell. Technol. Its Appl., pp. 245–250, 2015.

[22] R. F. Falah, O. D. Nurhayati, and K. T. Martono, “Aplikasi Pendeteksii Kualitas Daging Menggunakan Segmentasi Region of Interest Berbasis Mobile,” J. Teknol. dan Sist. Komput., vol. 4, no. 2, pp. 333–343, 2016.

[23] N. K. El Abbadi and N. E. Kadhim, “Brain Tumor Classification Based on Singular Value Decomposition,” Brain, vol. 5, no. 8, 2016.

[24] A. Santoso, I. Arif, and M. Hatta, “Pembelajaran Supervised SVM Untuk Identifikasi Obyek Pisau Pada Mesin X-Ray Bandara Juanda,” Nusant. J. Comput. its Appl., vol. 1, no. 1, 2017.
[25] S. Inaba and S. Arai, “Synthesis of Gabor filtered signal by convolving low-Q filtered signal with a Gaussian function for efficient computation,” in 2013 IEEE Global Conference on Signal and Information Processing, 2013, pp. 646–649.

[26] W. Lu, Z. Fan, L. Wei, X. Xiao-ming, and H. Wei, “A method of SAR target recognition based on Gabor filter and local texture feature extraction,” J. Radars, vol. 4, no. 6, pp. 658–665, 2015.

[27] K. Radhakrishnan, D. Sri, R. Dhanalakshmi, M. Elakkiya, and G. Gayathiri, “Design and implementation of high speed Gabor filter with variable thresholding process for disease detection,” in Sensing, Signal Processing and Security (ICSSS), 2017 Third International Conference on, 2017, pp. 430–435.

[28] S. Avinash, K. Manjunath, and S. S. Kumar, “An improved image processing analysis for the detection of lung cancer using Gabor filters and watershed segmentation technique,” in Inventive Computation Technologies (ICICT), International Conference on, 2016, vol. 3, pp. 1–6.

[29] Y. C. See, N. M. Noor, J. L. Low, and E. Liew, “Investigation of face recognition using Gabor filter with the random forest as learning framework,” in Region 10 Conference, TENCON 2017-2017 IEEE, 2017, pp. 1153–1158.

[30] D. Zhuang and D. Zhou, “Face recognition based on Log-Gabor filter binary transformation,” in Proceedings of the 29th Chinese Control Conference, 2010, pp. 2792–2795.

[31] S. Madgunda, U. Suman, G. S. Praneeth, and R. Kasera, “Steps in requirement stage of the waterfall model,” Int. J. Comput. Math. Sci., pp. 86–87, 2015.

[32] M. Hammoudi, G. Rothermel, and A. Stocco, “Waterfall: An incremental approach for repairing record-replay tests of web applications,” in Proceedings of the 2016 24th ACM SIGSOFT International Symposium on Foundations of Software Engineering, 2016, pp. 751–762.

[33] T. Fawcett, “An introduction to ROC analysis,” Pattern Recognit. Lett., vol. 27, no. 8, pp. 861–874, 2006.

[34] M. L. Larrea, “Black-Box Testing Technique for Information Visualization. Sequencing Constraints with Low-Level Interactions,” J. Comput. Sci. Technol., vol. 17, 2017.

[35] C. Henard, M. Papadakis, M. Harman, Y. Jia, and Y. Le Traon, “Comparing white-box and black-box test prioritization,” in Software Engineering (ICSE), 2016 IEEE/ACM 38th International Conference on, 2016, pp. 523–534.