Analysis of Canny edge detection algorithm and K-means on the leaf image

Sarah Syifara Baadilla
Dr. Budi Utami Fahnun, Skom., MM.

Email: sarahsyifara23@gmail.com
Email: bufahnun@gmail.com
Jurusan Teknik Informatika, Fakultas Teknologi Industri, Universitas Gunadarma

Article Info

Article History

Created September 11, 2019

Keyword:

K-means
Clustering
Canny
Segmentasi

ABSTRACT

Types of diseases in leaves such as Anthracnose and Blackspot can be detected using digital image processing applications . Steps were made in the processing of the image of one of them is the segmentation of textures that become things important in processing the image digitally to identify an image . In texture segmentation using the method K-means clustering is the segmentation of the method of clustering that classifies the data by disease by way of determining the iteration in the process and the number of clusters that is desired to be given as an insert . Then the results of the method of K-means generate some clusters that eventually there is an input to insert the image of the detected disease and the results of the insert that form the image of the binary which then perform the process of detection of edge canny to clarify the view image of that detected the disease in order to look more vivid and detailed. Results of the end of the application can be seen clearly that the appearance of the image that detected the disease can be seen clearly by using the detection edge canny.

.

PRELIMINARY

Segmentation is the process of separating objects or regions from one another based on certain characteristics, then the results of this segmentation process will be used for further processes that can be carried out on an image, for example the image classification process and the object identification process. Therefore, image segmentation can be needed to detect diseases in plants, especially in agriculture.

in agriculture there is research from (Khirade and Patil, 2015) aims to identify disease Anthranose and Blackspot on multiple datasets leaves by using the method of classification and segmentation . Methods of classification are used in research this is a Support Vector Machine (SVM) with the results of accuracy which is obtained 93.5%. While the method of segmentation that is used in research this is the K-means clustering. K-means clustering is a method of analyzing the data or methods of Data Mining that perform process modeling without supervision (unsupervised) and is one of the methods

that perform grouping of data to the system partition .

Furthermore, in agriculture there is research from (Singh and Misra, 2017) explaining that agricultural productivity is something that is very dependent on the economy so that detecting diseases in plants plays an important role in agriculture, because it has fairly natural disease in plants. In this study using several images of leaf disease, namely banana leaves, rose leaves, lemon leaves and bean leaves. The classification was first carried out using the minimum distance criteria with K-Means Clustering and showed its efficiency with an overall accuracy of 86.54%. Detection accuracy was improved to 93.63% with the proposed algorithm. In this study presents an algorithm for image segmentation techniques used to detect and classify leaf diseases in plants.

The research (Chitra and Ponmuthuramalingam, 2015) explains that the Canny Operators can be applied to different situations. In this study using images as an experiment. To calculate the accuracy of the image using parameters and time. The Canny operator can detect edges clearly.

To improve its performance, it is proposed to calculate the magnitude of the gradient and based on the gradient direction also make an adaptive calculation from the threshold of the Canny operator method. Then the time obtained from the adaptive threshold in the first image is 41ms and the second image is 51ms. The experimental results show that the edge detected by the improved Canny operator has more continuity, and a greater noise ratio. The Canny operator provides a low error rate, localizes edge points (the distance of edge pixels found for detection and actual edges are very short), and only gives one response for one edge (Kadir and Susanto, 2013).

RESEARCH METHODS

Method of K-means clustering is a method of analyzing the data or methods of Data Mining that perform process modeling without supervision (unsupervised) and is one of the methods that perform grouping of data to the system partition . In data mining there are two types of clustering methods used in data grouping, namely hierarchical clustering and non-hierarchical clustering. Hierarchical clustering is a method of grouping data that starts by grouping two or more objects that have the closest similarity. Then the process is forwarded to other objects that have a second closeness. And so on so that the cluster will form a kind of tree where there is a clear hierarchy of levels between objects, from the most similar to the least similar. Logically all objects will only eventually form a cluster.

Different the hierarchical clustering method, the non-hierarchical clustering method actually starts by first determining the desired number of clusters (two clusters, three clusters, etc.). After the number of clusters is known, then the cluster process is carried out without following the hierarchy process (Bastian, Ade., Sujadi, Harun., And Febrianto, Gigin., 2009). The K-Means Clustering method seeks to group existing data into several groups, where data in one group has the same characteristics with each other and has different characteristics from the data in other groups.

The proximity of two objects is determined based on the distance of the two objects. Likewise, the proximity of a data to a certain cluster is determined by the distance between the data and the center of the cluster. The closest distance between one data with a certain cluster

will determine which data belongs to which cluster. (Fahmi and Suprapto, 2015). Calculation of the distance of all data to each cluster center point using the Euclidean distance theory is formulated as follows:

$$d(i,j) = \sqrt{(x_1j - x_1j)^2 + x_2j - x_2j)^2 + .. + (xki - xkj)^2}$$

Where :

d (i,j) = Distance of data to i to cluster center j

xki = Data to i in the attribute data to k

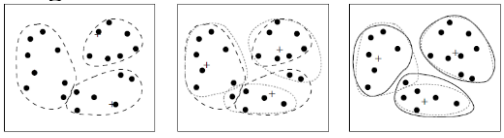
xkj = Center point to j in attribute to k

In research from (Prasetyo, 2009) explains the steps to do the K-means Clustering Algorithm, to see the flowchart algorithm display of k-means can be seen in Figure 1 .

1. Determine the value of K as the number of clusters.
2. Select K from dataset X as centroid.
3. Allocate all data to centroid with distance metric using the formula equation
4. Recalculate centroid C based on data that follows each cluster.

Figure 1. K-means Algorithm Flowchart

In the k-means segmentation clustering an illustration and can be seen in Figure 2.



Initial Cluster Iteration Process Final Cluster

Figure 2. Illustration of K-means Clustering
Source: Jiawei, Micheline and Jian (2012)

One of the modern edge detection algorithms is edge detection using the Canny method. Canny's edge detection was discovered by Marr and Hildreth who examined the modeling of human visual perception. There are some of the most optimum edge detection criteria that the Canny algorithm can meet:

a. Detect well (detection criteria) Ability to place and mark all edges according to the choice of convolution parameters. While also providing a very high flexibility in terms of

determining the level of detection of the desired edge thickness.

b. Localize well (localization criteria) With Canny it is possible to produce a minimum distance between the edge detected and the original edge.
c. Clear response (response criteria) There is only one response for each edge. So it is easily detected and does not cause confusion in subsequent image processing.

The difference between canny edge detection and other edge detection is that at canny edge detection, the image edge can be seen more clearly and in detail, so the canny operator is the most frequently used edge detection process. To see the results of the canny operator, see Figure 3.

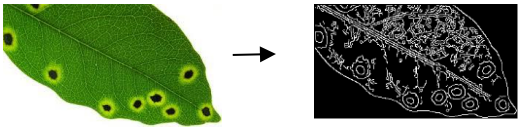


Figure 3. Canny Operators

For Canny Operators , this algorithm provides a low error rate, localizes edge points (the distance of edge pixels found for detection and the actual edge is very short), and only gives one response for one edge (Kadir and Susanto, 2013). Canny can detect actual edges with minimum error rates in other words, Canny operators are designed to produce optimal edge images. Following are the steps in implementing the Canny edge detection algorithm :

Step 1
First , do the thinning of the image with the goal to eliminate the noise . It is can be done by using a Gaussian filter with a veil simple . The veil which used size is much more smaller than the size of the image .

Step 2
After smoothing the image of the noise carried , carried the process to obtain the power of the edges (edge strength). It is done by using the operator Gaussian. Furthermore , the image gradient can be calculated via the formula :

$$|G| = |G_x| + |G_y|$$

Step 3
The third step is to calculate the edge direction . The formula that is used for the purposes of this :

$$theta = tan^{-1}(G_y, G_x)$$

Step 4
Determined the direction of the edge by using the inverse tangent of the gradient magnitude Y (Gy) divided by the gradient magnitude X (Gx). Directions are obtained from the calculation is then mapped to 0, 45, 90, or 135 degrees based on its proximity to the fourth degree of direction earlier .

Step 5
After the edge direction is obtained , non- maximum removal is carried out . The elimination of non- maximum carried in along the edge of the direction of the edge and eliminating pixels (pixel set becomes 0) are not considered as an edge. By way as it was , was obtained edge thin.

Step 6
The sixth step is a process called hysteresis . This process removes lines that are like discontinuous edges on objects . The trick is to use two thresholds T 1 and T2. Then , all of the pixel image that is worth more substantial than T1 regarded as pixel edge . Furthermore , all the pixels are connected to the pixels are and have a

value of more substantial than T2 is also regarded as the pixel edges . Flowchart operator Canny can be seen in Figure 4.

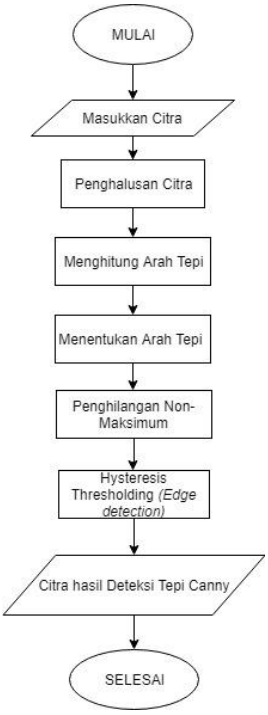


Figure 4. Canny Operator Flowchart

DISCUSSION

To design the application of canny edge-based segmentation and k-means clustering in this research method, there is a flow of research on leaf disease segmentation in Figure 5.

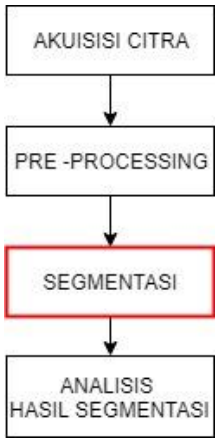


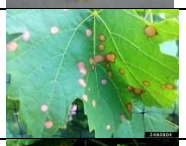
Figure 5. The disease process segmentation leaf

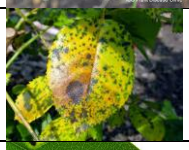
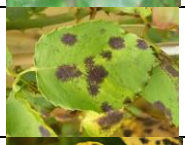
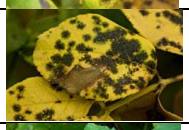
Below is an explanation of each flow segmentation process in leaf disease, as follows:

Image Acquisition

Image Acquisition is the process of capturing or scanning an analog image so that a digital image is obtained. The image used in this study is the leaf image obtained from the study (Alsraf, 2017) with JPG format. The original image of the leaf used has blotches on the leaf surface, which indicates that the leaf is diseased. In Anthracnose leaf disease there are brownish yellow spots, whereas for Blackspot leaf disease there are blackish brown spots. Dataset of leaf diseases obtained from research (Alsraf, 2017) .

Table 1. Image of leaf disease

Image Name	Original Image	Disease Name
1.jpg		Anthracnose
2.jpg		Anthracnose
3.jpg		Anthracnose
4.jpg		Anthracnose
5.jpg		Blackspot
6.jpg		Blackspot
7.jpg		Blackspot

8.jpg		Blackspot
9.jpg		Blackspot
10.jpg		Blackspot
11.jpg		Blackspot
12.jpg		Blackspot
13.jpg		Blackspot
14.jpg		Blackspot
15.jpg		Blackspot
16.jpg		Blackspot

Source: Alsraf (2017)

Pre Processing

Pre-processing that is used is cropping the image or cutting the image used to reinforce the area of the image detected by the disease. The steps taken in pre-processing can be seen in Figure 6 below..

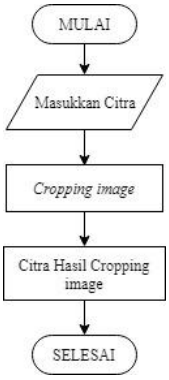


Figure 6. Pre-processing Flowchart

Examples of the results of pre-processing in the diseased leaf image can be seen in Figure 7 .

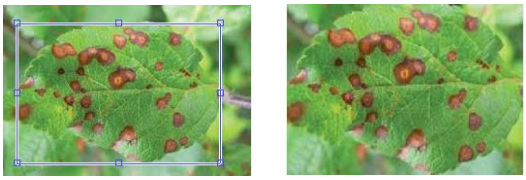


Figure 7. Results of Pre-processing Segmentation

Segmentation

The image results obtained from digital images through Pre Processing, will enter the segmentation stage. The purpose of segmentation is to separate areas of objects or images using the K-means Clustering and Canny edge detection methods. In the K-means algorithm it is a method of grouping data with a system partition. Whereas Canny edge detection is an image segmentation process to detect clear edges with a low error rate. Segmentation with K-means Clustering algorithm and edge detection with canny algorithm can be applied to segment the diseased leaf image on agricultural images. The segmentation process flowchart can be seen in Figure 8 .

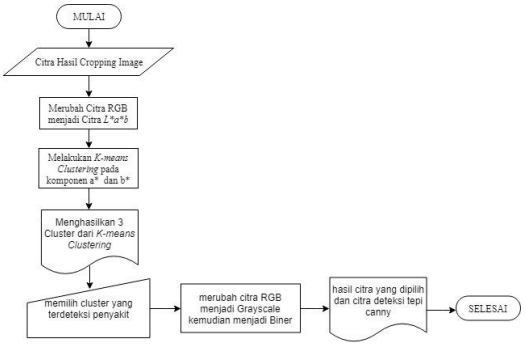


Figure 8. Diagram of the segmentation process flow

An example of K-means clustering on leaves infected with Blackspot leaf disease can be seen in Figure 9 .

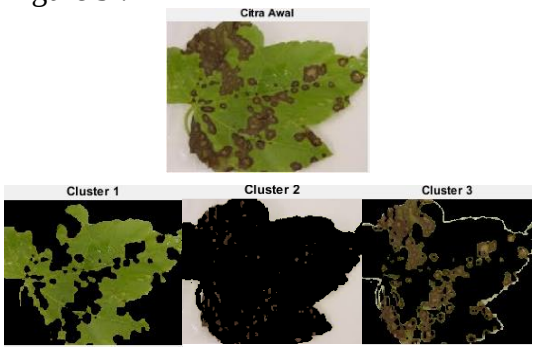


Figure 9. Example of Results from K-means Clustering

At the K-means Clustering segmentation stage there are 3 steps to be taken, namely:

1. Change RGB image to L * a * b image .
2. Conduct a K-means cluster on components a * and * b .
3. Produces 3 clusters, because the image has 3 colors.
4. An example of Canny edge detection in leaves infected with Blackspot leaf disease can be seen in Figure 10 .

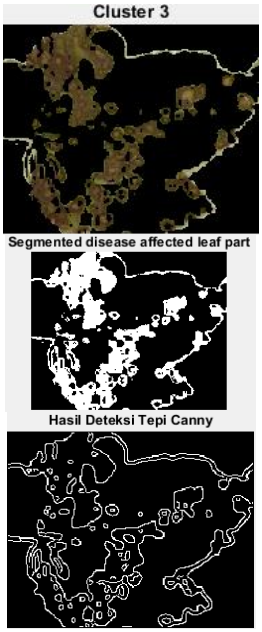


Figure 10. Example of Canny edge detection results

In the segmentation stage of Canny edge detection, there are 3 steps to be performed, namely:

1. Change RGB image to Grayscale image .
2. Change the Grayscale image to a binary image .
3. Perform Canny edge detection to clarify the edges of detected disease images.

RESULTS

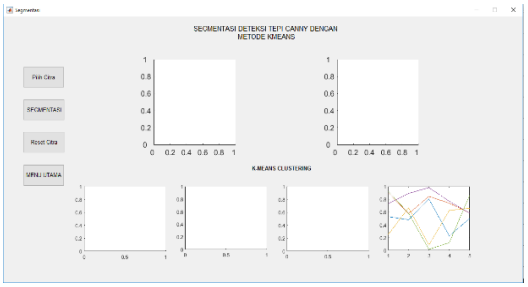


Figure 11. Main Page Display

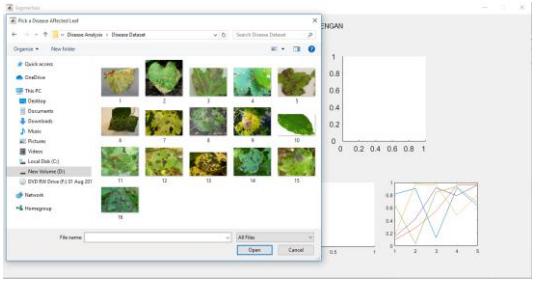


Figure 12. Display Select Image

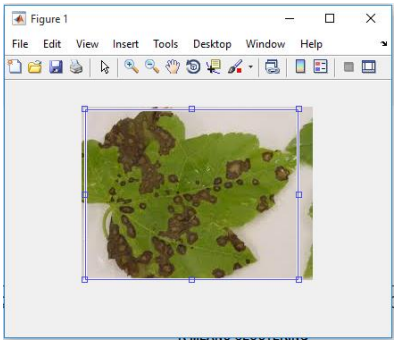


Figure 13. Selected Leaf Parts

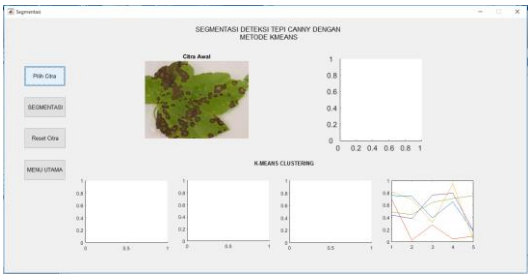


Figure 14. Display of Initial Leaf Image

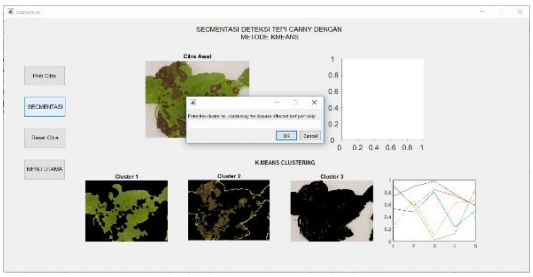


Figure 15. Display of K-means Clustering Segmentation

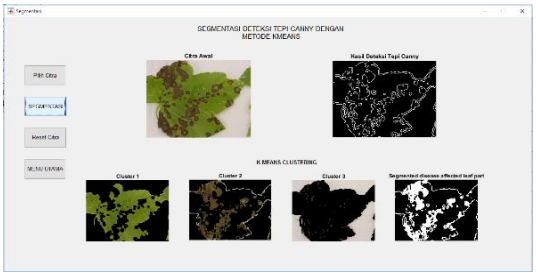


Figure 16. Display of results from K-means and Canny edge detection

CONCLUSION AND SUGGESTION

The conclusion of this research is the image of disease on leaves detected by the edge with the Canny method has a minimum error, the Canny operator is made to produce an optimal edge image and show the edges in the image to be seen more clearly. In the K-means Clustering method the number of clusters can be determined in advance so as to make this method less fast.

And for each diseased leaf image that has been tested it is produced that the diseased leaf image after carrying out the K-means Clustering process produces several clusters which are then inserted to select which clusters are detected by disease, the results of these entries produce a binary image that can be used to detect the edge of the canny so that the leaf image detected by the disease can be seen more clearly and in detail.

The suggestion from this research is that the application made still has many shortcomings so that the application can be developed and improved. To do edge detection can be done with other algorithms so that the image can be seen more clearly. In addition to edge detection, segmentation can be used with other methods.

BIBLIOGRAPHY

Alsraf , Muhammad. 2017. ***'Shape Feature extraction for Plant Disease'*** . *https://www.mathworks .com/matlabcentral/fileexchange/ 62825-shape-feature-extraction- for-plant- diseases?s_tid=prof_contriblnk*

Bastian, Ade., Sujadi , Harun., And Febrianto , Gigin . 2009. ***'AP PLICATION OF K-MEANS CLUSTERING***

ANALYSIS ALGORITHM IN HUMAN INFECTED DISEASE (CASE STUDY OF MAJALENGKA DISTRICT)' , (1), pp. 26–32.

Chitra, AD and Ponmuthuramalingam , P. 2015. ' ***An Approach for Canny Edge Detection Algorithms on Face Recognition*** ', International Journal of Science and Research (IJSR), 4 (11), pp. 24-27. doi : 10.21275 / v4i11.nov151012.

Kadir, Abdul. and Susanto, Adhi. 2013 ' ***Image Processing*** ', Image Proce ssing , p. 700. Available at: http://amearl.staff.gunadarma.a c.id/Downloads/folder/0.58 .

Khirade , Sachin D. and Patil, AB 2015. ' ***Plant disease detection using image processing'*** , *Proceedings - 1st International Conference on Computing, Communication, Control and Khirade , SD and Patil, AB (2015) 'Plant disease detection using image processing' , Proceedings - 1st International Conference on Computing, Communication, Control and* , pp. 768–771. doi : 10.1109 / ICCUBEA.2015.153.

Prasetyo , Eko . 2012.' ***DataMining con cepts and applications using MA TLAB*** ' https://scholar.google.co.i d/citations?user=9N8JRC4AAAA J&hl=en

Singh, Vijai . and Misra ,AK.2017. *' **Det ection of plant leaf diseases using image segmentation and soft computing techniques'*** ,*Information Processing in Agriculture* . China Agricultural University, 4 (1), pp. 41–49. doi : 10.1016 / j.inpa.2016.10.005.