

SYSTEMATIC LITERATURE REVIEW (SLR): DISEASE DETECTION IN MELONS USING DIGITAL IMAGE PROCESSING

Frieyadie

Information Sistem
Sekolah Tinggi Manajemen Informatika dan Komputer Nusa Mandiri
www.nusamandiri.ac.id
frieyadie@nusamandiri.ac.id

Abstract—Systematic Literature Review (SLR) is a technique used in this study that is used to study techniques for identifying leaf diseases using digital images as a basis for obtaining an understanding of disease identification techniques in melon leaves with digital images. Based on data from the Central Statistics Agency for the last 3 years from 2017-2019, melon production has increased considerably. Melon production data in 2017 was 92.43 tons, in 2018 was 118,708 and in 2019, overall melon production was 122,105 tons collected from 34 provinces in Indonesia. The problem that is often encountered in melon cultivation is the presence of plant pests that can harm and not maximize the yields of farmers. Several viruses cause mosaic disease that infects Cucurbitaceae plants, namely Cucumber aphid borne yellows virus (CABYV), Cucumber green mottle mosaic virus (CGMMV), Cucumber mosaic virus (CMV), Papaya ringspot virus (PRSV), Squash mosaic virus (SqMV), Squash leaf curl virus (SLCV), Watermelon mosaic virus (WMV). Information technology has now developed to be able to manage digital image data to identify problems faced by farmers. Several classification methods that can be used to answer problems include SVM, Artificial Neural Network, Decision Tree, Convolutional Neural Network.

Keywords: digital image, leaves, melon disease, systematic literature review

Abstrak—*Systematic Literatur Review (SLR)* merupakan teknik yang dilakukan dalam penelitian ini yang digunakan untuk mempelajari teknik-teknik mengidentifikasi penyakit pada daun menggunakan citra digital sebagai dasar memperoleh pemahaman mengenai teknik identifikasi penyakit pada daun melon dengan citra digital. Berdasarkan data dari Badan Pusat Statistik selama 3 tahun terakhir dari tahun 2017-2019, produksi buah melon cukup meningkat. Data produksi melon pada tahun 2017 sebanyak 92.43 ton, tahun 2018 sebanyak 118.708 dan pada tahun 2019, secara keseluruhan produksi buah melon sebanyak 122.105 Ton yang dihimpun dari 34 provinsi di Indonesia. Permasalahan yang sering dijumpai dalam budidaya melon adanya hama penyakit tanaman yang dapat merugikan dan tidak maksimalnya hasil panen para petani. Terdapat beberapa virus penyebab penyakit mosaik yang menginfeksi tanaman Cucurbitaceae yaitu Cucumber aphid borne yellows virus (CABYV), Cucumber green mottle mosaic virus (CGMMV), Cucumber mosaic virus (CMV), Papaya ringspot virus (PRSV), Squash mosaic virus (SqMV), Squash leaf curl virus (SLCV), Watermelon mosaic virus (WMV). Teknologi informasi saat ini telah berkembang untuk dapat mengelola data citra digital untuk dapat mengidentifikasi permasalahan yang dihadapi oleh petani. Beberapa Metode Klasifikasi yang dapat digunakan untuk menjawab permasalahan diantaranya SVM, Artificial Neural Network, Decision Tree, Convolutional Neural Network.

Kata kunci: citra digital, daun, penyakit melon, tinjauan literatur sistematis

INTRODUCTION

This systematic literature review was conducted to study various leaf disease recognition technologies with digital images, as a step to understand the technology for recognizing melon leaf disease with digital images. Melon is one of the gardening commodities that deserves attention,

because it has high economic value, delicious taste, and distinctive aroma, and is loved by the community. Based on data from the Central Statistics Agency for the last 3 years from 2017-2019, melon production has increased considerably. Melon production data in 2017 was 92.43 tons, in 2018 was 118,708 and in 2019, overall melon production was 122,105 tons



collected from 34 provinces in Indonesia. The problem that is often encountered in melon cultivation is the presence of plant pests that can harm and not maximize the yields of farmers.

Several viruses cause mosaic disease that infects Cucurbitaceae plants, namely Cucumber aphid borne yellows virus (CABYV), Cucumber green mottle mosaic virus (CGMMV), Cucumber mosaic virus (CMV), Papaya ringspot virus (PRSV), Squash mosaic virus (SqMV), Squash leaf curl virus (SLCV), Watermelon mosaic virus (WMV) (Khuluq, Phabiola, & Wijaya, 2020), Begomovirus which

causes leaf curl (Wilisiani, Somowiyarjo, & Hartono, 2014). Some of the fungi that attack melon plants include powdery mildew (the fungus *Oidium* sp or *Erysiphe cichoracearum*), fusarium wilt (*Fusarium oxysporum*), Fruit rot (*Phytophthora Nicotianae* fungus, *P. Capsici*, *Pythium* sp). Bacteria that also attack melon plants, namely bacterial wilt (*Bacteria Erwinia tracheiphila*), these bacteria attack the roots of melon plants. Pests that also attack melon plants include Trips (*Thrips parvipinus*), aphids (*Aphis gossypii*), and fruit flies (*Dacus* sp).

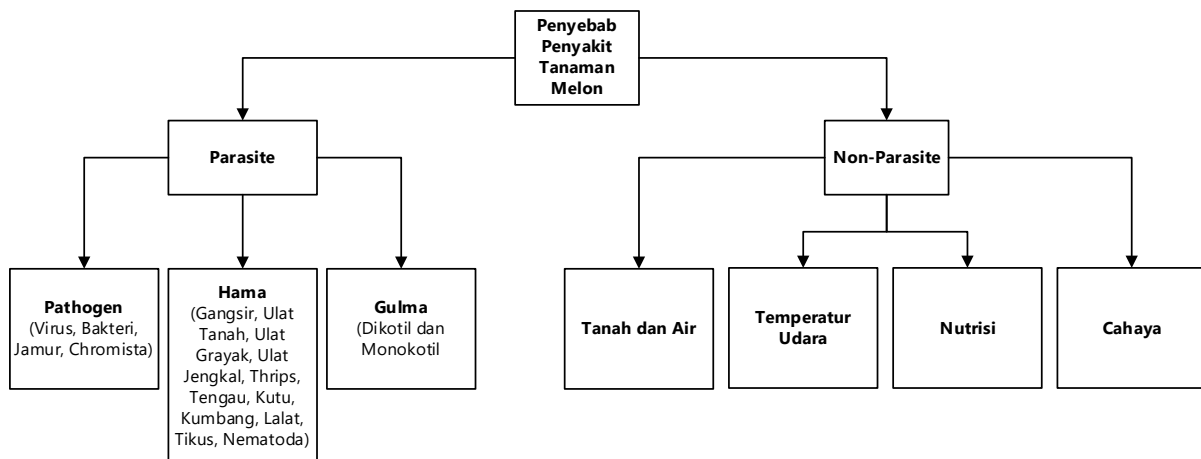


Figure 1. General Picture of the Causes of Melon Plant Diseases

The number of disease symptoms found is also a problem that is often faced by farmers. The current solution is still not optimal because there are still limited numbers of experts to produce solutions to diseases faced by farmers. The unavailability of information regarding the types and characteristics of the viral pathogens that cause the disease is an important obstacle in determining the virus management strategy (Wilisiani et al., 2014).

Information technology has now developed to be able to manage digital image data to be able to identify problems faced by farmers in identifying diseases in melon leaves. Based on the description above, this study will conduct a systematic literature review to identify, assess, and interpret research related to diseases of mango leaves and leaves in particular by using digital images. Based on the information above, this study will conduct a Systematic Literature Review to identify, evaluate, and interpret diseases in melon leaves using digital imaging.

MATERIALS AND METHODS

Systematic Literature Review can be described as a research method and process to identify and critically assess relevant research, as well as to collect and analyze data from the research. The Systematic Literature Review aims to identify all empirical evidence that fits pre-defined inclusion criteria to answer a specific research question or hypothesis. The steps taken in this research are by looking for similarities (Compare), looking for inequalities (Contrast), Providing views (Criticizing), Comparing (Synthesizing), and Summarizing (Summarize) (Solikin, 2020).

Compare is a review technique by looking for similarities between several kinds of literature and concluding. Contrast is a technique of reviewing by finding differences between several pieces of literature and concluding. Criticize is a technique of reviewing by making your own opinion on the source being read. Synthesize the technique of doing a review by combining several sources into a new idea. Summarize is a technique of reviewing by

rewriting the source in its sentence. The technique uses a Systematic Literature Review to find out things in the research stage, techniques, methods, or algorithms that can be used in further research.

RESULTS AND DISCUSSION

A. Digital Image

An image is simply a form of all the views that are around us in the form of shapes. Image is formed from 2 (two) types of images, namely two-dimensional analog images and digital images. Analog image in remote sensing, where an image of continuous variation in the view that is being perceived is represented by continuous variations in the tone of the image, such as images produced from photosensitive chemicals in photographic film (Encyclopedia.com, 2019). A digital image is a representation of a two-dimensional image as a finite set of digital values, known as an image element or a pixel (Bhat, 2014). In general, images can be divided into 3 (three) types, namely color images, greyscale images or black and white images, and binary images. A color image is an image that is built by combining 3 (three) primary colors, namely red, green, and blue. RGB color images are multi-spectral with one band for each red, green, and blue, resulting in a weighted combination of the three primary colors for each pixel (Fisher, Perkins, Walker, & Wolfart, 2000). Image processing serves as a technique for enhancing the raw image received from cameras or sensors placed on satellites, spacecraft, and aircraft. (Bhat, 2014). Image processing is a method commonly used to improve raw images received from various sources (Neetu Rani, 2017)

B. Image Format

There are many types of image files, some image types like TIFF are best for printing, while other image types like JPG or PNG are best for web graphics. (University of Michigan Library, 2021). JPEG is the name given to the algorithm developed by the Joint Photographic Experts Group whose goal is to minimize the file size of photo image files. (Alfio, Costantino, & Pepe, 2020)

C. Melon Disease Detection System Using Image Processing

To perform the disease detection process in melons by the system, it is necessary to input images using an image capture medium in the form of a digital camera. The results of the image capture collection are inputted into the system for further processing and analysis. The system that will

process the image must use an algorithm that will later get accurate and quality information. In brief, Figure 2 shows the stages of the process of a melon disease recognition system with digital image processing capabilities.

1. Input Data.

The data entered is in the form of a leaf image. Data in the form of leaf images, taking images by taking pictures with a digital camera. The image collection will continue to be inputted into the system for further processing.

2. Image Pre-Processing

This second step is a way to eliminate the interference (noise) contained in the image that is perceived by the camera so that it has smaller dimensions, and is more structured so that it can be further processed (Jumeilah, 2017). The preprocessing stage has several processes, namely case folding, removing stop words, tokenizing, and stems which can be seen in Figure 2 (Langgeni, Abdurahman Baizal, & Firdaus, 2010). Image filtering is performed to improve image quality. To smoothen the image to reduce noise, image blur can be improved. Broadly speaking, there are two types of algorithms linear and non-linear (Desai, Kushwaha, & Jha, 2020).

The smoothing algorithm itself can be divided into two types, namely the smoothing algorithm which uses the concept of isotropic diffusion and anisotropic diffusion. The main difference between these two algorithms is the distribution of the smoothing process. The algorithm that uses the isotropic diffusion method will not be able to identify the edge of the object in the image, but in the anisotropic algorithm, the side will be identified and will not be diffused, so that it becomes the same side as before. Finish smoothing. In this algorithm, only the inside and outside will be affected by the smooth process. Gaussian filters are very suitable for removing noise in the normal distribution, which usually appears in the distribution of images produced by digital camera processing because it is a natural phenomenon caused by the nature of light reflection and the light sensitivity of the camera sensor itself. The advantage of the median filter is its ability to reduce noise caused by random noise (such as salt noise and copper foil), or it can be called impulse noise. Compared to other types of nonlinear spatial (spatial) filters, the median filter is the most suitable filter for this situation. Therefore, this filter has been named the most powerful filter in processing images with similar noise (Wedianto, Sari, & H, 2016).

3. Image Segmentation

There are several techniques used for image segmentation. All of these techniques have their

respective uses and functions. All of these techniques can be approached from two basic approaches to segmentation, namely a region-based or edge-based approach. Each technique can be applied to a different image to perform the segmentation as needed. All of these techniques can also be classified into three categories of Structural Segmentation Techniques.

The image segmentation method is characterized based on two properties, namely Discontinuity and Similarity. The method based on Discontinuity is called the Edge-based method and the method based on Similarity is called the Region-based method (Karthick, Sathiyasekar, & Puraneeswari, 2014). Image segmentation systems can be classified as follows: 1) Region-based method; 2) Edge Based Method; 3) Hybrid Techniques. Region-based segmentation describes the division of the image into standard regions of connected pixels by applying the uniformity criterion between groups of candidate pixels. One pixel in a region is similar to a certain calculated feature or object (such as paint, intensity, and surface). The main disadvantage of histogram-based area detection is that it does not provide spatial information (only the grayscale range). Edge-based segmentation works at the location of pixels in the image that is parallel to the boundaries

of the objects seen in the image. The hint then covers the edges of the area or object, and the number of objects drawn is equal to the number of borders on the image. For accurate segmentation, the identified boundary edges must be approximately the same as the object in the input image. Segmentation algorithms have been developed to segment images and can be classified into the following: 1) Segmentation by Clustering; 2) Segmentation by Edge Detection; 3) Segmentation by Fuzzy Logic; 4) Segmentation by Neural Network; 5) Segmentation by Region-Based; 6) Segmentation by Thresholding. (Jayapriya & Hemalatha, 2019) (Mutlag, Ali, Aydam, & Taher, 2020)

4. Feature Extraction

The feature extraction method is used to obtain as much information as possible from the image (Dey, 2018). The selection and effectiveness of selected features and extraction is a major challenge today. Many methods are used to extract features, which may depend on the Geometric features, Statistics features, Texture features, and Color features. Each main feature type is divided into many types which are divided such as the Color feature which is divided into three types (color moment, color histogram, and average RGB) (Mutlag et al., 2020).

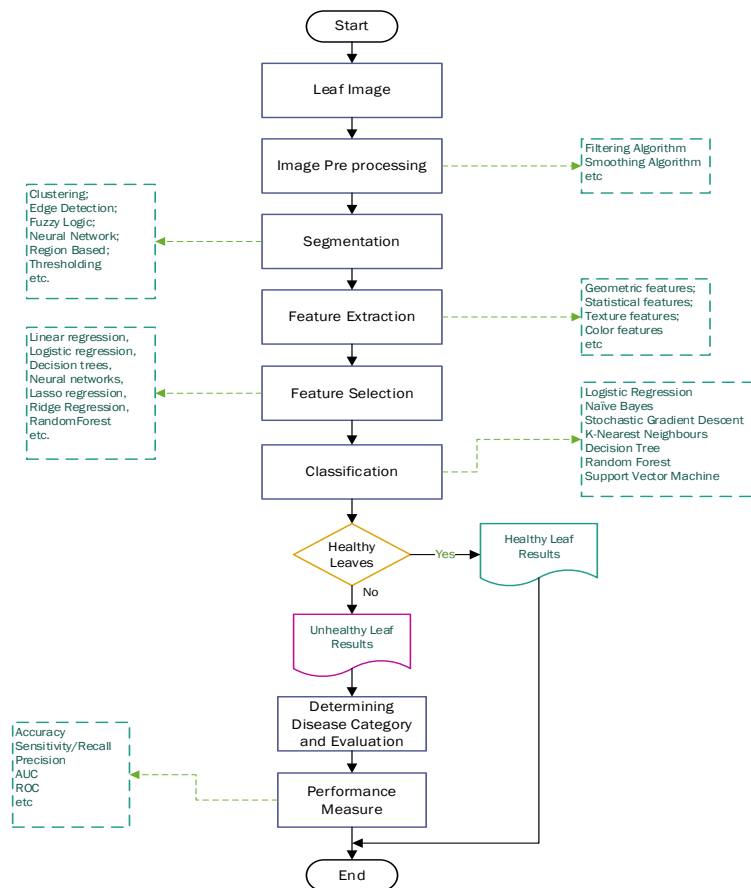


Figure 2. Identification and Classification of Diseases on Melon Leaves

5. Feature Selection

The Feature Selection method is used to reduce computation time, improve prediction performance, and better understand data in machine learning or pattern recognition applications (Chandrashekar & Sahin, 2014). Feature Selection method, related to supervised and unsupervised methods. Supervised learning algorithms include linear regression, Logistic regression, Decision trees, Neural networks, etc. The Unsupervised feature learning method of the target variable will not be considered a method that uses correlation to eliminate redundant variables. Unsupervised feature learning methods are classified into three categories, namely: Wrapper method, Filter method, Embedded or Intrinsic method.

6. Classification

Classification is a controlled approach to machine learning, in which algorithms learn from the data given to them, and then use that learning to classify new observations. Classification algorithms can generally be classified as follows: Logistic regression, Naive Bayes Classifier, Support vector machines, Decision trees, Random Forest, k-Nearest

Neighbor, etc. Logistic regression is a classification and not a regression algorithm. It estimates discrete values (Binary values like 0/1, yes/no, true/false) based on a specific set of independent variables. Naïve Bayes is a classification technique based on the assumption of independence between predictors or known as the Bayes theorem. In simple terms, the Naive Bayes classifier assumes that the presence of certain features in a class is not related to the presence of other features. The drawback of Naïve Bayes is that what is known as the 'zero probability problem', when the conditional probability is zero for a particular attribute, fails to provide a valid prediction. Decision trees, generally simple to understand and visualize, require little data preparation. This method can also handle numeric and categorical data. On the other hand, if it is complex it does not generalize well ("overfitting"), and decision trees may become somewhat unstable because small variations in the data can result in completely different trees.

7. Performance Measure

Performance measurement is very important, when it comes to classification problems, in this case relying on the AUC - ROC curve. If it is necessary

to examine or visualize the performance of a multi-class classification problem, we use the AUC (Area under the Curve) ROC (Receiver Operating Characteristics) curve. This is one of the most important evaluation metrics for checking the performance of any classification model. Also written as AUROC (Area under the Receiver Operating Characteristics). Accuracy is the ratio of correct predictions to total predictions. It is important if, to have a symmetrical data set (the number of FN and FP is close). Used if: false negatives & false positives have the same cost. Sensitivity/Recall is the ratio of true positives to total (actual) positives in the data. Precision is the ratio of true positives to the total predicted positives.

Result

From the results of the systematic literature review that has been carried out, there are research questions (Research Questions / RQ) that are made to identify research problems as follows: 1) RQ1: What Digital Image Method can be used optimally to extract leaf disease features 2) RQ2: Classification method what can be used efficiently to classify the type of disease on leaves. Meanwhile, the appropriate research objectives related to this research are: 1) RO1: Determining the digital image method that can be used optimally to extract melon leaf disease features 2) RO2: Classification method that can be used efficiently to classify species Melon leaf disease

CONCLUSION

Digital Image Methods that can be used to identify diseases on melon leaves are the stages of Image Acquisition, Preprocessing, Segmentation, Feature Extraction, and Feature Selection. Classification methods that can be used are SVM, Artificial Neural Network, Decision Tree, and Convolutional Neural Network.

REFERENCE

- Alfio, V. S., Costantino, D., & Pepe, M. (2020). Influence of image tiff format and jpeg compression level in the accuracy of the 3d model and quality of the orthophoto in UAV photogrammetry. *Journal of Imaging*, 6(5). <https://doi.org/10.3390/jimaging6050030>
- Bhat, M. (2014). Digital Image Processing. *International Journal of Scientific & Technology Research*, 3(1), 1–5. Retrieved from <http://www.ijstr.org/paper-references.php?ref=IJSTR-0114-8118>
- Chandrashekar, G., & Sahin, F. (2014). A survey on feature selection methods. *Computers and Electrical Engineering*, 40(1), 16–28. <https://doi.org/10.1016/j.compeleceng.2013.11.024>
- Desai, B., Kushwaha, U., & Jha, S. (2020). Image Filtering-Techniques, Algorithm, and Applications. *GIS Science Journal*, 7(11), 970–975.
- Dey, S. (2018). No CNN application on structured data-Automated Feature Extraction. Retrieved from Towards Data Science website: <https://towardsdatascience.com/cnn-application-on-structured-data-automated-feature-extraction-8f2cd28d9a7e>
- Encyclopedia.com. (2019, April 16). Analogue Image. Retrieved April 20, 2021, from A Dictionary of Earth Sciences website: <https://www.encyclopedia.com/science/dictionaries-thesauruses-pictures-and-press-releases/analogue-image>
- Fisher, R., Perkins, S., Walker, A., & Wolfart, E. (2000). Glossary - Color Images. Retrieved April 20, 2021, from homepages.inf.ed.ac.uk website: https://homepages.inf.ed.ac.uk/rbf/HIPR2/c_olimage.htm
- Jayapriya, P., & Hemalatha, S. (2019). Comparative Analysis of Image Segmentation Techniques and Its Algorithm. *International Journal of Scientific and Technology Research*, 8(10), 2209–2212.
- Jumeilah, F. S. (2017). Penerapan Support Vector Machine (SVM) untuk Pengkategorian Penelitian. *Jurnal RESTI (Rekayasa Sistem Dan Teknologi Informasi)*, 1(1), 19. <https://doi.org/10.29207/resti.v1i1.11>
- Karthick, S., Sathiyasekar, D. K., & Puraneeswari, A. (2014). A Survey Based on Region-Based Segmentation. *International Journal of Engineering Trends and Technology*, 7(3), 143–147. <https://doi.org/10.14445/22315381/ijett-v7p262>
- Khuluq, M., Phabiola, T. A., & Wijaya, I. N. (2020). Penularan Virus Bergejala Mosaik Pada Tanaman Melon (Cucumis melo L.) Secara Mekanis dan Melalui Vektor Kutu Daun. *Jurnal Agroekoteknologi Tropika*, 9(1), 76–86.
- Langgeni, D. P., Abdurahman Baizal, Z. K., & Firdaus, Y. (2010). Clustering Artikel Berita Berbahasa Indonesia Menggunakan Unsupervised Feature Selection. *Seminar Nasional Informatika 2010 (SemnasIF 2010)*, 1(4), 1–

10. Yogyakarta: UPN "Veteran" Yogyakarta. Retrieved from <http://jurnal.upnyk.ac.id/index.php/semnasi/article/view/1175>
- Mutlag, W. K., Ali, S. K., Aydam, Z. M., & Taher, B. H. (2020). Feature Extraction Methods: A Review. *Journal of Physics: Conference Series*, 1591(1). <https://doi.org/10.1088/1742-6596/1591/1/012028>
- Neetu Rani. (2017). Image Processing Techniques: A Review. *Journal on Today's Ideas - Tomorrow's Technologies*, 5(1), 40-49. <https://doi.org/10.15415/jotitt.2017.51003>
- Solikin. (2020). Deteksi Penyakit Pada Tanaman Mangga Dengan Citra Digital: Tinjauan Literatur Sistematis (SLR). *Bina Insani Ict Journal*, 7(1), 63-72.
The University of Michigan Library. (2021, February
- 11). Image File Formats - All About Images. Retrieved April 20, 2021, from Research Guides at University of Michigan Library website: <https://guides.lib.umich.edu/c.php?g=282942&p=1885348>
- Wedianto, A., Sari, H. L., & H, Y. S. (2016). Analisa Perbandingan Metode Filter Gaussian, Mean Dan Median Terhadap Reduksi Noise. *Jurnal Media Infotama*, 12(1), 21-30. <https://doi.org/10.37676/jmi.v12i1.269>
- Wilisiani, F., Somowiyarjo, S., & Hartono, S. (2014). Identifikasi Molekuler Virus Penyebab Penyakit Daun Keriting Isolat Bantul pada Melon. *Identifikasi Molekuler Virus Penyebab Penyakit Daun Keriting Isolat Bantul Pada Melon*, 18(1), 47-54. <https://doi.org/10.22146/jpti.15602>