The 1st International Seminar August 2022
The Changing Role of Knowledge and Living Sustainability in ASEAN Community
Website: https://proceeding.unpkediri.ac.id/index.php/ip

Disease Detection System in Tobacco Leaves based on Edge Detection with Decision Tree Classification Method

[1*]Firman Matiinu Sigit, [2]Rahmad Syaifudin, [3]Dwi Ari Suryaningrum
[1,2,3] Teknik Elektro, Universitas Tulungagung
E-mail: [1]firmanmatiinusigit@unita.ac.id, [2]syaifudinrahmad@unita.ac.id, [3]dwiarisuryaningrum@unita.ac.id
*Corresponding Author

Abstract—Indonesia has long been known as an agricultural country, one of Indonesia's best agricultural products is tobacco. Tobacco with good quality can be seen from the leaves, in fact the diseases that attack tobacco are of various types which can be seen from the changes in the tobacco leaves starting from the seeding and planting period. In the tobacco growing period, it is divided into two major parts, namely seeding period and the planting period, so that diseases that attack tobacco are also divided into two, namely diseases that attack during seeding and planting. This research is limited to diseases that attack tobacco at the time of planting, because at the time of seedling the tobacco has not yet produced leaves. When tobacco enters the planting period, at this time tobacco leaves begin to form. Good care is needed at this time such as fertilization, nutrition, vitamins, and pest control in order to obtain healthy tobacco so that tobacco is not susceptible to disease. Tobacco that lacks nutritional intake will be susceptible to diseases including fungi, bacteria, and viruses. The disease attack on tobacco has its own characteristics that appear on tobacco leaves. Early detection of the disease is very important so that disease control can be precise and the spread of the disease can be prevented so as not to cause endemic. In this research, an early detection system of tobacco leaf disease based on image processing will be designed. Normalization image, grayscale technique, folllowed by edge detection will be applied in these image so that from here the entropy, energy, and inertia values of the image can be obtained using statistical measures, and the last one using the decision tree classification technique can be classified as uninfected leaf or infected leaf. In this study, feature extraction from images of tobacco leaves that are not infected with the virus using grayscale techniques followed by edge detection produces an average statistical measure with entropy (h) values between 2,341 to 2,676, energy (e) values between 6,112 to 6,665, and inertia values. (i) between 3,322 to 3,576, while for leaves infected with the virus the average value of entropy (h) is between 4,543 to 5,576, the average value of energy (e) is between 12,212 to 13,455, and the average value of inertia (i) between 5,343 to 6,597.
Keywords— Tobacco leaves; Tobacco Mozaic Virus (TMV); Edge Detection; Decision Tree.

Corresponding Author:

Firman Matiinu Sigit,
Study Program Teknik Elektro,
Universitas Tulungagung,
Email: firmanmatiinusigit@unita.ac.id

I. INTRODUCTION

Indonesia has long been known as an agricultural country, because it is blessed with fertile soil. The agricultural sector is one of the keys to the economy in Indonesia, because it contributed about 14.43 percent of the national GDP (Gross Domestic Product) in 2013, although it experienced a slight decline compared to the previous decade (2003) which reached 15.19 percent. Currently Indonesia is the world's largest producer of palm oil, cloves and cinnamon, the second largest producer of nutmeg, natural rubber, cassava, vanilla and coconut oil, the third largest producer of rice and cocoa, the fourth largest producer of coffee, and the fifth largest producer of tobacco (Wikipedia, 2022).

Because Indonesia is the fifth largest tobacco producing country in the world, the tobacco commodity is one of the government's concerns. Tobacco is a product that is very sensitive to cultivation method, planting location, season/weather, and processing method. Therefore, a tobacco cultivar will not produce the same quality if it is planted in a different agroecosystem. Tobacco products are very specific to a certain area and certain cultivars. As a result, various tobacco products are usually named after the location where they are grown (Wikipedia, 2022). Due to the very sensitive character of tobacco plants, tobacco plants need attention in their care.

Tobacco is an agricultural product that is processed from the leaves of the plant which is also named the same, so that the main concern is the tobacco leaves. Healthy tobacco will have resistance to disease, therefore adequate nutrition and pest control are needed in its treatment. If not, tobacco will be susceptible to diseases including pests, fungal and viral infections. Pest attacks are usually in the form of caterpillar attacks that cause leaf holes or leaf quality to decrease, infections due to fungi and viruses can cause damage to leaves, leaves do not grow perfectly, leaf color changes, leaves shrivel, twist and fold, leaf size shrinks, stunted plants, and etc. Infections due to fungi and viruses that are known to attack tobacco include lanas disease, Fusarium wilt disease, tobacco mosaic disease, cucumber mosaic disease, cracker disease, betok disease, root-knot Nematodes, and leaf rust.

These diseases if not treated properly at early detection will spread and cause even greater damage. At the beginning of the disease attack, the quality of the leaves decreases and causes a decrease in yields qualitatively and quantitatively slightly, but if proper control/prevention is not carried out it will spread and increase the damage, if viruses and fungi have spread, an epidemic will occur and control will be difficult and ineffective. For example in the Bojonegoro area, lanas disease can cause tobacco deaths of more than 50% and also cracker disease which attacks about 30% of the land area of the epidemic which caused a loss of 5 billion rupiah in 1984 (Dalmadiyo et al., n.d.).]Therefore, early detection to find out about the disease is very important so that the character of the disease can be known, it is hoped that treatment / control can be carried out appropriately and effectively.

Related research that examines tobacco disease are:

Research conducted by Cece Suhara and Titiek Yulianti in 2017 in the journal entitled "Effectiveness of the Carna-5 (Cucumber Mosaic Virus Associated RNA-5) Vaccine against Cucumber Mosaic Virus (CMV) Infection in Cigar Tobacco Plants". Mosaic disease caused by Cucumber Mosaic Virus (CMV) causes losses in both production and quality of tobacco leaves. Until now, chemical control of the virus has not been implemented, except for the control of vector insects. One alternative is the use of the CMV (Carna-5) satellite as a biological controller. The aim of this study was to evaluate the effectiveness of Carna-5 in inhibiting the development of CMV in cigar tobacco. The Carna-5 vaccine did not affect plant growth, but was able to reduce the incidence and severity of disease (Suhara & Yulianti, 2018).

Research conducted by Gembong Dalmadiyo, Supriono, and Bagus Hari-Adi, which is contained in the sub-chapter of the book entitled "Virgina Tobacco Diseases and Its Control" which briefly explains that diseases caused by fungal infections can reduce the quality and quantity of production, if initially not handled properly it will become an epidemic/endemic and will cause even greater damage and handling will be difficult and less effective. There are several diseases that often arise in tobacco that can harm farmers, including lanas disease, crackers, mosaics, and sprouting. In the Bojonegoro area, this lanas disease can cause tobacco deaths of more than 50%. Cracker disease can also be found in the Bojonegoro area and became endemic in 1984 with an attack area of about 30% of the land and caused a loss of around 5 billion rupiah (Dalmadiyo et al., n.d.).

Research conducted by researchers from Balittas Malang, namely Bagus Hari-Adi and Gembong Dalmadiyo about diseases that attack Madura Tobacco, which is stated in the Journal entitled "Madura Tobacco Diseases and Their Control" briefly explains that most tobacco diseases caused by fungi and viruses are divided into two according to the life span of tobacco, namely diseases during the seedling period and diseases during planting. Diseases during the nursery include germination and seedling disease. Diseases during planting for example are Lanas Disease, Fusarium wilt disease, Tobacco virus mosaic disease, Cucumber virus mosaic disease, Cracker disease, and Betok disease. Control efforts are prevention efforts because if it is widespread prevention will be difficult and ineffective (Adi & Dalmadiyo, n.d.).

Mr. Hrishikesh P. Kanjalkar and Prof. S.S. Lokhande in the International Journal of Advanced Research in Computer Engineering & Technology (IJARCET) Volume 3, published January 1, 2014 wrote a journal entitled "Feature Extraction of Leaf Diseases" mentioning feature extraction in detecting diseases through plant leaves can use the HSI method (Hue, Saturation, Intensity) (Kanjalka & Lokhande, 2014).

In this study, the disease was limited to a disease caused by a virus that attacks tobacco leaves. Diseases that attack during seeding period are not described in detail in this study because they attack before the plant develops leaves. Tobacco diseases caused by viruses include Tobacco Mosaic Virus (TMV), Cucumber Mosaic Disease, Cracker Disease, and Betok Disease.

II. RESEARCH METHOD

In general, the research methodology is divided into two major parts, namely Pre-engineering and Engineering. Pre-engineering involves many activities in the agricultural sector, while Engineering includes image processing and decision tree classification.

A. Pre-engineering (Agricultural Sector)

Pre-engineering activities include the process of searching for diseases in the field, conducting RNA tests in the laboratory and identifying diseases or viruses, after obtaining the name and type of virus, the virus is stored. The next step land is made, where the amount of land is determined by the number of types of viruses found (Here, we only identify one type of virus, namely the tobacco mosaic virus), with a land size of 5 x 5 meters with a number of plants of 20 plants on each land and a greenhouse is formed which is expected to be free from outside diseases or pests. One land to grow tobacco free from virus as control parameter, so we made two artificial lands in this experiment.

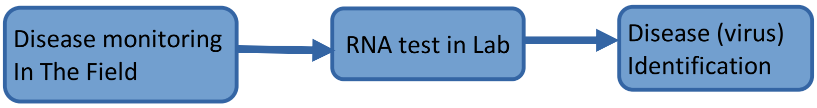

Figure 1. Diagram Block of Pre-engineering Agrucultural Sector

In this experiment, we use virginia tobacco type, which is grown on podzolic soil with an acidity of pH 5.5 - 6.0 and air temperature ranges from 21-32.3 degrees Celsius at an altitude of 0 - 900 meters above sea level (Ali & Hariyadi, 2018).

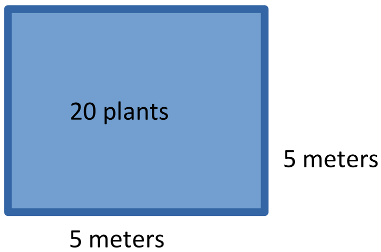

Figure 2. Land Size of Tobacco Virus Experiment

Disease treatment begins when planting on artificial land in the first week. Taking pictures when tobacco starts to show symptoms of disease on the leaves (3rd week). Every day 10 images are taken on one land so that in one week there are 5 working days, so in one field in one week 50 images are taken, for the 3rd week to the 12th week (tobacco growing period) the total leaf image obtained is 50 x 10 = 500 sick leaf images dataset. The total image of healthy leaves obtained is 50 X 10 = 500 pieces, the total image of leaves obtained is 1000 pieces image dataset.

B. Engineering Process

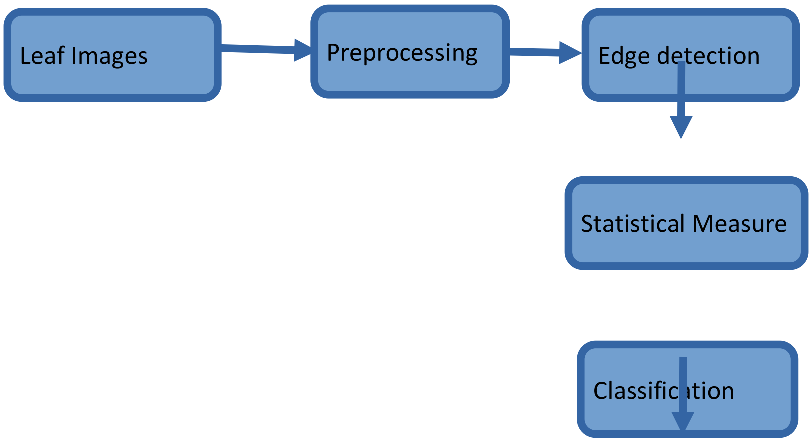

Figure 3. Block Diagram of Engineering Process

The engineering process consists of two major processes including image processing and classification. Image processing consists of preprocessing, feature extractions and statistical measuring of these features. The classification here uses a decision tree to determine the type of virus.

Preprocessing

Preprocessing here, the image is converted into a grayscale image and the size is normalized to 32x32. After the image becomes grayscale and the size is 32x32, then the image undergoes filtering with convolution sobel for edge detection.

Edge Detection

Edge Detection feature extraction can use Sobel Edge-Detection Operator or Canny Edge-Detection Operator, the working principle is to convolute each image pixel value after converting it to grayscale.

Statistical Measure

The result of the edge detection process is then the value of the image pixel is processed according to statistical equations to find the entropy (h), energy (e), and inertia (i) values. If each value in the matrix resulting from normalization has been obtained, assume that the value is FP, then the Normalized Fourier Coefficients or NFP can be obtained with the equation,

$$\mathbf{NFP}_{u,v} = \frac{\left|\mathbf{FP}_{u,v}\right|}{\sqrt{\sum_{(u\neq 0)\wedge(v\neq 0)} \left|\mathbf{FP}_{u,v}\right|^2}} \qquad (1)$$

spectrum data can be obtained using the *entropy* equation, (h) as below,

$$h = \sum_{u=1}^{N}\sum_{v=1}^{N} \mathbf{NFP}_{u,v} \log\left(\mathbf{NFP}_{u,v}\right) \qquad (2)$$

energy, e can be obtained by the equation,

$$e = \sum_{u=1}^{N}\sum_{v=1}^{N} \left(\mathbf{NFP}_{u,v}\right)^2 \qquad (3)$$

and the magnitude of *inertia*, (i) can be explained by the equation,

$$i = \sum_{u=1}^{N}\sum_{v=1}^{N} (u-v)^2 \mathbf{NFP}_{u,v} \qquad (4)$$

III. RESULT AND DISCUSSION

The following are examples obtained from experiments on healthy tobacco leaves (Uninfected Leaf) and tobacco leaves exposed to the Tobacco Mozaic Virus (TMV-Infected Leaf).

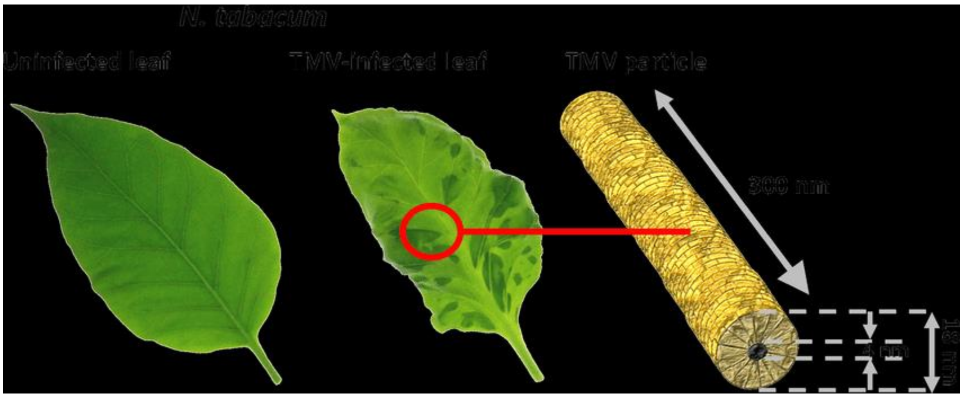

Figure 4. Uninfected Leaf and TMV-Infected Leaf (Koch et al., 2016)

After obtaining data regarding feature extraction from tobacco leaf images in the form of statistical data (h, e, i) and color data (HSI extraction), the data obtained can be presented in the following table.

From the table 1, the decision tree algorithm that can be arranged as an example is if the image of virginia tobacco leaves has an entropy value (h) between 2,341 to 2,676 and an energy value (e) between 6,112 to 6,665 and an inertia value (i) between 3,322 to 3,576, then the tobacco leaf virginia is healthy/uninfected leaf, and if the picture of virginia tobacco leaf has an entropy value

(h) between 4,543 to 5,576 and an energy value (e) between 12,212 to 13,455 and an inertia value (i) between 5,343 to 6,597 then the virginia tobacco leaf is TMV - infected leaf

Table 1. RESULT OF THE FEATURE EXTRACTIONS FROM TOBACCO LEAF EXPERIMENT

Data Leaf	Feature image extraction of Tobacco Leaf			
	Entropy (h) (mean)	**Energy (e) (mean)**	**Inertia (I) (mean)**	**Tobacco Leaf Disease**
Uninfected Leaf				
50 leaves from week 3	2.341	6.112	3.322	Uninfected Leaf
50 leaves from week 4	2.432	6.132	3.345	Uninfected Leaf
50 leaves from week 5	2.445	6.253	3.354	Uninfected Leaf
50 leaves from week 6	2.513	6.265	3.367	Uninfected Leaf
50 leaves from week 7	2.527	6.289	3.453	Uninfected Leaf
50 leaves from week 8	2.551	6.342	3.466	Uninfected Leaf
50 leaves from week 9	2.610	6.376	3.476	Uninfected Leaf
50 leaves from week 10	2.613	6.387	3.555	Uninfected Leaf
50 leaves from week 11	2.645	6.454	3.556	Uninfected Leaf
50 leaves from week 12	2.676	6.665	3.576	Uninfected Leaf
range	2.341 to 2.676	6.112 to 6.665	3.322 to 3.576	Uninfected Leaf
Infected Leaf				
50 leaves from week 3	4.543	12.212	5.343	TMV-Infected Leaf
50 leaves from week 4	4.570	12.222	5.444	TMV-Infected Leaf
50 leaves from week 5	4.665	12.323	5.465	TMV-Infected Leaf
50 leaves from week 6	4.789	12.656	5.476	TMV-Infected Leaf
50 leaves from week 7	4.897	12.754	5.667	TMV-Infected Leaf
50 leaves from week 8	5.110	13.276	6.132	TMV-Infected Leaf
50 leaves from week 9	5.254	13.289	6.342	TMV-Infected Leaf

Data Leaf	**Feature image extraction of Tobacco Leaf**			
	Entropy (h) (mean)	**Energy (e) (mean)**	**Inertia (I) (mean)**	**Tobacco Leaf Disease**
50 leaves from week 10	5.355	13.365	6.344	TMV-Infected Leaf
50 leaves from week 11	5.434	13.387	6.476	TMV-Infected Leaf
50 leaves from week 12	5.576	13.455	6.597	TMV-Infected Leaf
range	4.543 to 5.576	12.212 to 13.455	5.343 to 6.597	TMV-Infected Leaf

.

IV. CONCLUSION

In this study, of course, the conditions were made as ideal as possible starting from the type of soil, soil pH, irrigation, nutrition and fertilization, temperature, humidity, sunlight intensity, and the plants were cultivated free from pests except for diseases that were given TMV virus. Of course, if this system is used to detect a disease that attacks tobacco leaves, there are still many things that need to be improved, for example, tobacco leaves with holes due to pests will affect the statistical measure values (h, e, i) in image processing, so we still need lots of data to enrich our algorithm.

In addition to the dataset of the number of images that need to be reproduced, other things that can be improved from this system are the parameters of the image processing that can be obtained such as grayscale color feature extraction or Hue Saturation Image (HSI). In addition, another thing that can be improved from this system is the identification of disease not only one type of disease but also for the identification of several other diseases caused by virus such as Cucumber Mosaic Disease, Cracker Disease, and Betok Disease. With the increase in the number of images in the dataset, it is suggested that the classification algorithm that can be used is the Convolutional Neural Network algorithm from this type of deep learning algorithm in this classification.

ACKNOWLEDGMENTS

The author would like to thank Balittas (Balai Penelitian Tanaman Pemanis dan Serat) Malang for referrence and discussion, Laboratorium Biokimia Universitas Brawijaya and Universitas Tulungagung that have provided financial support for this research.

REFERENCES

Adi, B. H., & Dalmadiyo, G. (n.d.). Penyakit Tanaman Tembakau Madura dan Pengendaliannya. In *Hama Tanaman Tembakau* (pp. 51–59). Pustaka Balitbangtan.

Ali, M., & Hariyadi, B. W. (2018). *Teknik Budidaya Tembakau*. https://doi.org/10.31219/osf.io/zy3eb

Dalmadiyo, G., Supriyono, & Adi, B. H. (n.d.). Penyakit Tanaman Tembakau Virginia dan Pengendaliannya. In *Hama Tanaman Tembakau* (pp. 64–76). Pustaka Balitbangtan.

Kanjalka, M. H. P., & Lokhande, P. S. S. (2014). Feature Extraction of Leaf Diseases. *Journal of Advanced Research in Computer Engineering & Technology*, *3*(1).

Koch, C., Ebe, F. J., Azucena, C., Förste, A., Walheim, S., Schimmel, T., Bittner, A. M., Jeske, H., Gliemann, H., Eiben, S., Geiger, F. C., & Wege, C. (2016). Novel roles for well-known players: from tobacco mosaic virus pests to enzymatically active assemblies. *Beilstein: Journal of Nanotechnology*, *7*, 613–629. https://doi.org/10.3762/bjnano.7.54

Suhara, C., & Yulianti, T. (2018). Efektivitas Vaksin Carna-5 (Cucumber Mosaic Virus Associated RNA-5) terhadap Infeksi Cucumber Mosaic Virus (CMV) pada Tanaman Tembakau Cerutu (Nicotiana tabacum L.). *Buletin Tanaman Tembakau Serat & Minyak Industri*, *9*(1), 24. https://doi.org/10.21082/btsm.v9n1.2017.24-34

Wikipedia. (2022). Tembakau. In *Wikipedia*.