Lampiran_II-26.pdf

by

Color and Texture Feature based for Maize Leaf Disease Identification

Indah Agustien Siradjuddin, Arif Riandika, Wahyudi Agustiono
Informatics Department
University of Trunojoyo Madura
Kamal, Madura, Indonesia
indah.siradjuddin@trunojoyo.ac.id, arief.rian87@gmail.com, Wahyudi.agustiono@trunojoyo.ac.id

***Abstract*— This paper presents a combination of color and texture features to identify the maize plant disease based on maize leaf image. Global Color Histogram and Color Coherence Vector extract color features, and Local Binary pattern is used to extract the texture features. There are four classes in the proposed identification model: Cercospora Leaf Spot, Common Rust, healthy, and Northern Leaf Blight disease. For the identification process, we trained the Voting classifier with five CARTs and a plant-village dataset. The trained identification model achieved an average accuracy is 75,935%. In addition, we added a segmentation image for the preprocessing stage to improve the accuracy. As a result, this preprocessing stage increased the accuracy to 82.645%.**

Keywords—Maize Plant Disease, Global Color Histogram, Color Coherence Vector, Local Binary Pattern, voting classifier.

I. Introduction

Agriculture is the world's oldest and largest industry that plays a pivotal role in safeguarding food security. In addition to soil degradation, climate change, and pollution, the spread of pest and disease is a serious threat causing crop damage, and severe harvest loss especially for maize, a staple food throughout most of the world. For instance, according to Rwomushana as cited by Supartha, et al[1]. , farmers in Honduras, Argentina and Africa lose up to 40%, 72% and 53% respectively of their maize crop to pests and diseases attacks. In the recent years, the threat of pest and disease attack has become serious, due the rapid development of pest resistance as the results of the excessive use of pesticides and climate change [2].

Accordingly, reducing off-farm chemical inputs and using only selective agrochemicals in favor of more environment friendly methods of pest or disease control are important practices in Building a Resilient Future in Food and Farming. However, this is in not easy for farmers because they often lack of access to sufficient information about the recent good agriculture practices in protecting crops against pest or disease attacks. Lots of them still depend on the local information they get from other farmers or field officers from Agriculture Department. The problems become more complicated, because there are limited number of field officers available and at the same time, they also need to do tough agricultural field-works, such as, manual crop tracking, log keeping, using pen and paper which are time consuming, and tedious. All these challenges, if not taken seriously, can cause huge loses, to crop production, and in the long run could threaten food security.

For the last few years, smart farming has been introduced as a one way to help farmers face serious challenges especially in monitoring and protecting their crops against pest and disease attacks more efficiently. For instance, image processing and computer vision for crop monitoring and early detection systems are considered as the major application of smart farming to solve difficult task in pest and disease control. Unlike traditional pest and disease manual detection that is inefficient, expensive, and time-consuming, image-based recognition using artificial intelligence techniques provides a smart solution for automatic identification. Moreover, the implementation of image processing and machine learning methods that are able to remove noises and illumination of image sources proved robustness and accuracy in detection and classification [3].

Two general identification approaches are based on obtaining features of the image data. They are feature learning and feature engineering based. First, feature learning, features are learned based on image datasets using the advantage of Convolutional Neural Networks[4], [5]. However, the characteristic of the obtained features is difficult to identify. The second is feature engineering. In this approach, we have to define the representative features based on the problem and the dataset, and also, we have to decide on the feature extraction method [6], [7]. With this approach, we can see the characteristic of the obtained features.

For this matter, we proposed maize disease identification using a feature engineering approach. There are four kinds of maize disease based on leave images, including the healthy leave. As seen in the image, the differences between each type of disease and the healthy one are the color and texture of the leaves. Drawing from the characteristic of the leaf images, this article, therefore, aims to identify the maize leaf disease with color and texture features. The finding of the proposed approach is that we applied the preprocessing stage, i.e. leaf images segmentation. In this preprocessing stage, we extract the images' red pixels, blue pixels, and green ones before the feature extraction process.

This paper is organized as follows, section two described the approach for color and texture feature, section three is the proposed model for the identification, section four explained the conducted experiments, and final section is the conclusion.

II. Feature Extraction

A. Global Colour Histogram

The proposed model's first feature is color features extracted using Global Colour Histogram (GCH). GCH creates a histogram of pixels in each channel of the colour image [8]. First, the image is transformed into quantization to reduce the dimension of features. We use eight-level quantization in each channel. Therefore, the number of extracted features is 24. The GCH process is depicted in Fig. 1.

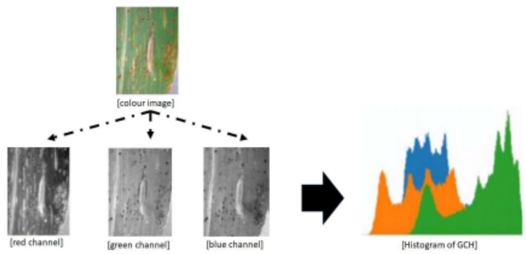

Fig. 1. Global colour Histogram

B. Color Cohorent Vector

Color Coherence Vector (CCV) extracts the color feature of an image. It obtained whether the pixels in the same area (neighborhood) have similar color (coherent) or not (incoherent pixels) [9]. There two primary processes are involved in this feature extraction. The first is the quantization image, and the second is labeling pixels using n-Connected Components Labelling (CCL). The quantization image is to convert pixels into several ranges or levels to reduce the number of features to be extracted. Second, every level from the quantization is grouped or labeled using the Connected Components approach. The *n*-CCL clustered pixels based on their *n*-neighborhood. The *n* will be four or eight neighborhoods [10].

These two processes are illustrated in Fig 2. In this figure, the pixels range from 0-29 and are converted into three levels of pixels {0,1,2}. Each level of pixels is then grouped or labeled using 8-CCL. For example, the pixel with level 0 is grouped into two areas, i.e., areas A and B (two labels), and level 1 is grouped into one area only (C). Meanwhile, level 2 is grouped into two areas (D and E)

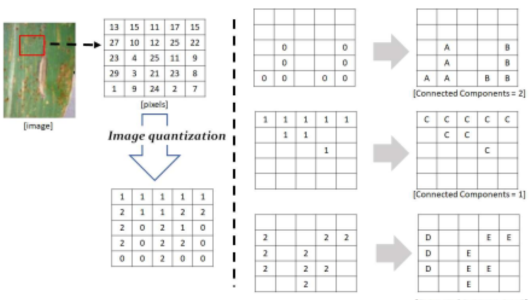

Fig. 2. Quantization Image and Connected Components on each level of quantization

The grouped (labeled) pixels are then calculated. If the number of pixels in each group is more than a determined threshold, it is included in coherent pixels, otherwise incoherent pixels. The calculation of CCV is illustrated in Fig. 3. The algorithm of CCV is written in Algorithm 1. Our proposed model used 8-level of quantization. Therefore, there are 16 features from the CCV feature extraction.

Labeled	A	B	C	D	E
Quantized (Level) Pixels	0	0	1	2	2
Number of Pixels	4	4	8	3	6

CCV, threshold =4
Coherent label : A, B, C, E ; Incoherent : D

Quantized (Level) Pixels	0	1	2
Coherent Pixels	8	8	6
Incoherent Pixels	0	0	3

Fig. 3. Calculation of Color Coherent Vector of Grouped or Labeled components

```
Algorithm 1 : Feature Extraction - Color Coherent Vector
INPUT : Image I, Binary Image B
OUTPUT : Color Coherent Vector CCV //size : 2x levelPixels, //[0]:coherent, [1]:incoherent

LevelPixels ← ImageQuantization (I)
for each LevelPixels of quantizedImage :
    label ← ConnectedComponents (B)
    for each label:
        if the_number_of_pixels ≥ threshold :
            increment(CCV[levelPixels,0])
        else:
            increment(CCV[levelPixels,1])
```

C. Local Binary Pattern

The proposed model's last feature is texture features obtained using Local Binary Pattern (LBP). The LBP captures the texture feature of a pixel within neighboring pixels by encoding the pixel. The pixel is encoded by comparing it (g_c) with its neighborhood pixels $\{g_0, g_1, \dots, g_7\}$ using Formula in (1) and (2).

$$LBP = \sum_{i=0}^{7} s(g_i - g_c)2^i \tag{1}$$

$$s(x) = \begin{cases} 1, & if\ x \geq 0 \\ 0, & otherwise \end{cases} \tag{2}$$

The LBP is illustrated in Fig. 4. The histogram of LBP represents the features. However, to reduce the number of features, first, the LBP is quantized into several levels. We used eight-level quantization in the LBP feature extraction in the prosed model. Second, we built a histogram of the quantization result to represent texture features. Therefore, from LBP feature extraction, we have eight features.

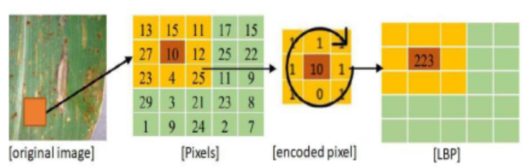

Fig. 4. Quantization Image and Connected Components on each level of quantization

D. Colour Processing Detection Algorithm

The proposed model identifies maize disease based on leaf images. Table 1 shows that the disease or healthy images are mainly shown by the green or red pixels. Therefore, in the experiment, we add pre-processing stage, i.e., the segmentation of the input image. Before the feature extraction, we segment the images into green and red pixels. However, we also did the blue pixels segmentation to compare the result. We use Colour Processing Detection Algorithm (CPDA) [11] in the segmentation pre-processing stage. The CPDA algorithm is written in Algorithm 2.

III. Maize Disease Identification Model

The proposed maize disease identification model is depicted in Fig. 5. First is the segmentation process to extract the green or red pixels area of the maize leaves. The second is feature extraction (color and texture) using GCH, CCV, and LBP from the using image from the segmentation preprocessing.

Finally, these features are concatenated and used as input for the classifier. We use a voting classifier that consists of five CART (Classification And Regression Trees) with depths 2, 4, 6, 8, and 10.

We measured the performance of the conducted experiments using two kinds of measurement, i.e. F1-score and accuracy as written in Formula (3) and Formula (4). F1-score measurements only used True Positive (TP), False Positive (FP), and False Negative (FN) values, as seen in Formula (5) and Formula (6). Meanwhile, the accuracy score is the ratio of all correctly predicted and the total number of data.

$$F1\,Score = 2 * \frac{Precision * Recall}{Precision + Recall} \tag{3}$$

$$Accuracy = \frac{TN+TP}{TN+FP+TP+FN} \tag{4}$$

$$Precision = \frac{TP}{TP+FP} \tag{5}$$

$$Recall = \frac{TP}{TP+FN} \tag{6}$$

```
Algorithm 2 : Image Segmentation - CPDA
INPUT : ColorImage I
OUTPUT :
RedPixelOutputImage (outR),
GreenPixelOutputImage(outG),
BluePixelOutputImage (outB)
Divide_ColourImg_into_ThreeChannels (R, G, B)
GrayImg ← GrayScale(I)
RCP ← R - GrayImg //Red Component Pixels
GCP ← G - GrayImage // Green component Pixels
BCP ← G - GrayImage // Blue component Pixels
RP ← RCP - GCP/2 - BCP/2
GP ← GCP - RCP/2 - BCP/2
BP ← BCP - RCP/2 - GCP/2
//Red Pixel Segmentation
for each (x,y) in image coordinat :
    if RP(x,y) ≥ threshold:
        outR(x,y) ← I(x,y)
    else:
        outR(x,y) ← 0
//Green Pixel Segmentation
for each (x,y) in image coordinat :
    if GP(x,y) ≥ threshold:
        outG(x,y) ← I(x,y)
    else:
        outG(x,y) ← 0
//Blue Pixel Segmentation
for each (x,y) in image coordinat :
    if BP(x,y) ≥ threshold:
        outB(x,y) ← I(x,y)
    else:
        outB(x,y) ← 0
```

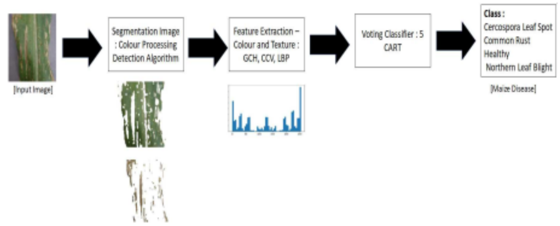

Fig. 5. Maize Disease Identification Model

IV. Result and Analysis

We tested the proposed maize plant disease identification on maize leaf images from the Plant-Village Dataset [12]. The dataset consists of three maize disease classes (i.e., Cercospora Leaf Spot/Gray Leaf Spot, Common Rust, Northern Leaf Blight) and one class for healthy leaf. The distribution of total images in each class is written in Table 1. In the experiment, we augmented the input image since we have an imbalanced dataset, as seen in Table 1. Therefore, for the augmented dataset, we have 3000 images in each class.

The first experiment compares the number of features used for the identification. The first model used 24 features from GCH, the second used 16 features from CCV, the third used 8 features from LBP, and the last combined all features (48 feature). Table 2 shows the performance using F1-score measurement of the first experiment.

TABLE I. Maize Leave Classes in Plant-Village Dataset

Class / Disease	Number of Images	Images
Cercospora Leaf Spot/Gray Leaf Spot	513	
Common Rust	1192	
Healthy	1162	
Northern Leaf Blight	985	

TABLE II. Performance Accuracy of Features Comparison on the Maize Plant Disease Identification.

Class	F1 – Score (%)			
	GCH	*CCV*	*LBP*	*GCH+CCV+LBP*
Cercospora Leaf Spot/Gray Leaf Spot	45,94	43,44	51,66	50,9
Common Rust	98,42	91,68	94,14	98,36
Healthy	88,42	82,4	71,08	86,48
Northern Leaf Blight	67,2	67,14	45,88	68

Class Cercospora Leaf Spot has the least performance, as seen in Table 2. Table 1 shows that this class has the fewest images in the plant village dataset; even to train the model, we have augmented the dataset to balance the number of images in each class. However, to increase the performance, more data are required. Table 2 also shows that, on average, the performance of maize disease identification using GCH, CCV, LBP, and the combination of all features are 74.995, 71.165, 65.69, and 75.935, respectively. The combined features achieved the highest F1-score. The combined features achieved the highest F1-score since this approach concatenates texture and color features (48 features in total) to identify the disease.

TABLE III. PERFORMANCE WITH F1-SCORE USING PREPROCESSING SEGMENTATION STAGE ON THE COMBINED FEATURES.

Class	F1 – Score (%)			
	Without segmentation	*Green Pixel*	*Red Pixel*	*Blue Pixel*
Cercospora Leaf Spot/Gray Leaf Spot	50,9	58,14	67,92	41,7
Common Rust	98.36	87,7	94,22	99,08
Healthy	86,48	98,1	99,66	88,7
Northern Leaf Blight	68	65,2	68,78	68,6

TABLE IV. PERFORMANCE WITH ACCURACY SCORE USING PREPROCESSING RED PIXEL SEGMENTATION STAGE ON THE COMBINED FEATURES.

Class	Accuracy Score (%)
Cercospora Leaf Spot/Gray Leaf Spot	83,19
Common Rust	97,15
Healthy	99,73
Northern Leaf Blight	85,01

For The second experiment, we applied the preprocessing stage on the input image, i.e., the segmentation using Color Processing Detection Algorithm (CPDA), and combined all features from GCH, CCV, and LBP. Table 3 shows the performance F1-score for the second experiment. In the segmentation preprocessing stage, we did three kinds of segmentation. They are Green Pixel, Red Pixel, and Blue Pixel segmentation. As a result, the average F1-score using the original and segmented input images (Green, Red, and Blue pixels) are 75.935, 77.285, 82.645, and 74.52, respectively.

The result in Table 3 shows that red and green segmentation of an input image during preprocessing stage increases the performance accuracy of maize disease identification. The most significant improvement is obtained in the class Cercospora Leaf Spot, which has the lowest F1-Score using the original image (see Table 2). We also calculate the performance using the accuracy score for the red pixel segmentation. The result is shown in Table 4. As in F1-Score, the Cercospora Leaf Spot has the lowest accuracy score.

The increasing F1-score is achieved since the red, and green pixels' segmentation emphasizes input images required for identification, as seen in Fig. 6.

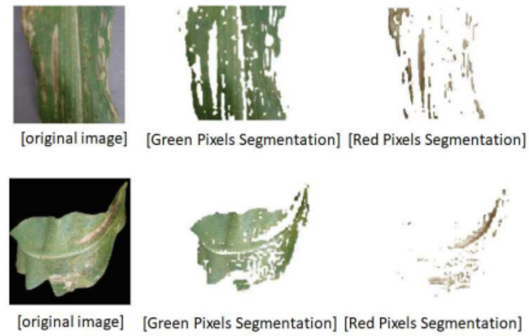

Fig. 6. Red and Green Pixels Segmentation using Color Processing Detection Algorithm

Fig. 6 shows that the segmentation extracts the required valuable pixels (Red and Green) to identify the maize disease. Furthermore, the segmentation process keeps the greenish pixels of the healthy leaf and the blank image for the reddish pixel segmentation, as seen in Fig. 7. Therefore, this segmentation increases the accuracy.

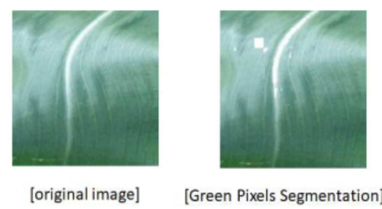

Fig. 7. Green Pixels Segmentation of the healthy leaf using Color Processing Detection Algorithm

However, the result of blue pixel segmentation is depicted in Fig. 8. As seen in the figure, the left pixels using blue segmentation do not give enough features for identification since the leaf images from Table 1 show that the images consist of reddish and greenish pixels. Therefore, the blue pixels segmentation achieved the lowest accuracy compared to the original image.

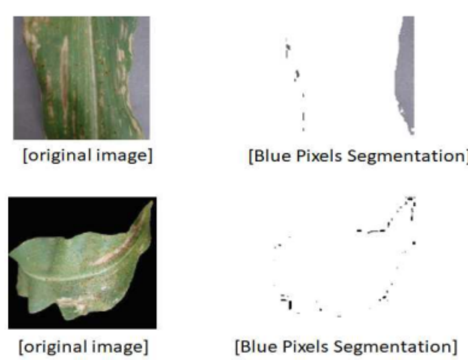

Fig. 8. Blue Pixels Segmentation using Color Processing Detection Algorithm

V. Conclusion

The paper proposes identifying maize plant disease based on a combination of texture and color features using Global Color Histogram, Color Coherence Vector, and Local Binary Pattern. The model is trained and tested on a Plant-Village Dataset. The experiments showed that the combined features achieved higher performance than the identification model with only color or texture features. We also include the segmentation image for the preprocessing stage to obtain specific input image pixels necessary for disease identification. The results showed that the preprocessing stage increased the performance, especially for the red pixels segmentation, since the disease on maize leaf is characterized mainly by the area of the red pixel in the image. For further research, the proposed identification model using the combination of features can be compared to the Convolutional Neural Networks architecture, which has feature learning layers.

References

[1] I. W. Supartha, I. W. Susila, A. A. Agung Sri Sunari, I. G. F. Mahaputra, I. K. W. Yudha, and P. A. Wirdana, "Damage characteristics and distribution patterns of invasive pest, Spodoptera frugiperda (J.E Smith) (Lepidoptera: Noctuidae) on maize crop in Bali, Indonesia," *BIODIVERSITAS*, vol. 22, no. 6, pp. 3378–3389, Jun. 2021.

[2] T. Miedaner and P. Juroszek, "Global warming and increasing maize cultivation demand comprehensive efforts in disease and insect resistance breeding in north-western Europe," *Plant Pathology*, vol. 70, no. 5, pp. 1032–1046, Jun. 2021.

[3] M. E. Karar, F. Alsunaydi, S. Albusaymi, and S. Alotaibi, "A new mobile application of agricultural pests recognition using deep learning in cloud computing system," *Alexandria Engineering Journal*, vol. 60, no. 5, pp. 4423–4432, 2021, doi: https://doi.org/10.1016/j.aej.2021.03.009.

[4] R. Ahila Priyadharshini, S. Arivazhagan, M. Arun, and A. Mirnalini, "Maize leaf disease classification using deep convolutional neural networks," *Neural Computing and Applications*, vol. 31, no. 12, pp. 8887–8895, Dec. 2019, doi: 10.1007/s00521-019-04228-3.

[5] E. Rocha, L. Rodrigues Moreira, and J. Mari, *Maize leaf disease classification using convolutional neural networks and hyperparameter optimization*. 2020. doi: 10.5753/wvc.2020.13489.

[6] K. R. Aravind, P. Raja, K. V. Mukesh, R. Aniirudh, R. Ashiwin, and C. Szczepanski, "Disease classification in maize crop using bag of features and multiclass support vector machine," in *2018 2nd International Conference on Inventive Systems and Control (ICISC)*, 2018, pp. 1191–1196. doi: 10.1109/ICISC.2018.8398993.

[7] D. Chauhan, D. R. Walia, D. C. Singh, D. M. Deivakani, and M. Makhan, "Detection of Maize Disease Using Random Forest Classification Algorithm," p. 6, 2021.

[8] S. Khan and M. Narvekar, "Survey of Techniques for Disorder Detection in Tomato(Solanum Lycopersicum)," in *2018 Second International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2018, pp. 1853–1858. doi: 10.1109/ICCONS.2018.8662933.

[9] S. R. Dubey and A. S. Jalal, "Apple disease classification using color, texture and shape features from images," *SIViP*, vol. 10, no. 5, pp. 819–826, Jul. 2016, doi: 10.1007/s11760-015-0821-1.

[10] I. A. Siradjuddin, A. Sakinah, and M. K. Sophan, "Combination of Feature Engineering and Feature Learning Approaches for Classification on Visual Complexity Images." *International Journal of Innovative Computing Information and Control*, vol. 17, no. 3, pp. 991-1005, Jun 2021.

[11] E. Kiani and T. Mamedov, "Identification of plant disease infection using soft-computing: Application to modern botany," *Procedia Computer Science*, vol. 120, pp. 893–900, 2017, doi: https://doi.org/10.1016/j.procs.2017.11.323.

[12] "Plant Village Dataset." https://github.com/spMohanty/PlantVillage-Dataset

Lampiran_II-26.pdf

ORIGINALITY REPORT

10%	5%	10%	1%
SIMILARITY INDEX	INTERNET SOURCES	PUBLICATIONS	STUDENT PAPERS

MATCH ALL SOURCES (ONLY SELECTED SOURCE PRINTED)

3%

Selected Wahyudi Agustiono. "An ensemble-of-classifiers approach for corn leaf based diseases detection", 2021 IEEE 7th Information Technology International Seminar (ITIS), 2021

Publication

Exclude quotes	Off	Exclude matches	Off
Exclude bibliography	On		