

Partial Centroid Contour Distance (PCCD) in Mango Leaf Classification

Eko Prasetyo^{a,1}, Raden Dimas Adityo^{a,2}, Nanik Suciati^{b,1}, Chastine Fatichah^{b,2}

^a Department of Informatics, Faculty of Engineering, Universitas Bhayangkara Surabaya, Surabaya, 60231, Indonesia
E-mail: ¹eko@ubhara.ac.id; ²dimas@ubhara.ac.id

^b Faculty of Intelligent Electrical and Informatics Technology, Institut Teknologi Sepuluh Nopember, Surabaya, 60111, Indonesia
E-mail: ¹nanik@if.its.ac.id; ²chastine@if.its.ac.id

Abstract—The research in the classification of mango leaf varieties requires appropriate features and classification methods to achieve high accuracy. The system used 263 features, texture and color features included Boundary Moments features that generated from Centroid Contour Distances (CCD). The CCD measures distance from center to the edge along 360 degrees, this causes enormous computational loads. On the other hand, the final part of mango leaf to recognize the mango varieties simply by observing the leaf base and leaf tip, so the mango leaf as the special case of CCD can be solved by only generating features at these parts. We propose Partial CCD (PCCD) by calculating the distance from boundary point does not to the center point of the leaf but to the midpoint-cut of the leaf base or leaf tip. PCCD has two parts, PCCD Leaf Base and PCCD Leaf Tip to capture leaf base and leaf tip features, respectively. On experiment testing with PCCD or another color, shape, and texture features only, the system can't achieve high accuracy, but the combination of all features increase accuracy up to 10%. The comparison among all various features are used in classification. It is compared the original features, individual PCCD features (Leaf base and Leaf tip), and combination of Leaf base and Leaf tip. These results show that combination of original features and PCCD features achieve the best accuracy 80.17% and average accuracy 78.41%. The highest accuracy performance obtained by SVM classification is 81.73%. The comparison with other features also proved that the combination obtains better performance.

Keywords—centroid contour distance; mango leaf; partial; shape features.

I. INTRODUCTION

Image-based classification research is now multi-disciplinary research in various fields related to computing. The system is also called computer vision. Research shows that the performance of a computer system is influenced by feature extraction and/or classification methods. The research in image-based classification was conducted [1] using the Otsu global segmentation thresholding method, morphological operations, and watershed transformation to classifies early detection of breast cancer successfully. The accuracy performance achieved up to 98.9%. In medical images, classification also can be used to detect tuberculosis diseases early by using hybrid classification between Artificial Neural Network (ANN) and Support Vector Machine (SVM)—the system achieves a sensitivity performance of 89.87% [2]. The Otsu method was also used in medical image processing to find the position of the vein for injection process. The right position vein was important to avoid repeat injection by the nurses, because wrong injection position made the patient uncomfortable or scared.

From the experiment, detection of the vein was successfully done and capable show the position of the vein [3]. The other research conducted weighting scheme for K-Nearest Neighbor (K-NN) in order to optimize the accuracy and precision fingerprint indoor localization system for multiple object tracking. The experiment result showed that system performance increased up to 25% better than the conventional system [4]. In remote sensing imagery, image processing was also used to the enhancement of Landsat8 imagery by developing algorithm for denoising and modifying homographic filter for edge preservation. The result of algorithm was worked well on the images of Landsat8 [5].

In the agriculture field, previous study [6] created a system to classify tree species by combining spectral features and LiDAR metrics. The experiment result shows the combination features are higher than individual features. Previous study [7]–[9] also used spectral information to classify fruit into several categories. The system also achieved very satisfying results. Research by [10] combined color, shape and texture features to classify fruit with neural

network, the system applied to 1653 color fruit images from the 18 categories achieved classification accuracy up to 89.1%. Another approach is to combine feature extraction and classification is Convolutional Neural Network (CNN) based to classify fruit and vegetables [11]. Computer vision research is also used in solving leaf classification; research Leaf texture analysis is used to classify olive spot diseases [12]. Texture features (Moment invariants for multi-component shapes) can be created to classify several leaf varieties [13]. Shape features of the leaf were also used in many plant classification [14], to classify plant varieties and compared between traditional classifier and CNN based classifier [15]. The experiment result shows that CNN based classifier is better than a traditional classifier. Peak detection algorithm is created to support other leaf features in leaf classification, and successfully complement each other and achieve better performance [16].

The related research in leaf shape recognition uses Centroid Contour Distance as shape features is conducted in [17]. This research classifies four classes of leaf shapes on 200 leaf images of a tropical plant. Each class consists of 50 images. By using Probabilistic Neural Network classification, it achieves accuracy 96.67% [17]. The improved version of CCD features, which are called Width CCD features, are also combined with Band Limited Phase Only Correlation (BLPOC) to calculate the similarity of finger vein image. This research uses a score-level fusion method based on the weighted SUM rule. Using a database collected from 123 volunteers, the combination features achieve an efficient recognition performance with the equal error rate (EER) 1.78% [18]. The other research in leaf shape classification proposes a shape features for mobile retrieval of leaf images. This feature is called multi-scale arch height (MARCH). Some hierarchical arch height features at different chord spans are extracted from each contour point to provide a compact, multiscale shape descriptor. Both the global and detailed features of the leaf shape can be effectively captured by the proposed algorithm. The MARCH features can achieve higher classification rate and retrieval accuracy than the other features benchmarks with a more than 500 times faster retrieval speed [19]. The other research in the field of leaf shape classification is proposed a leaf shape descriptor based on sinuosity coefficients and leaf geometrical features [20]. The sinuosity coefficients are defined using the sinuosity measure, which is a measure expressing the degree of meandering of a curve. By using the Radial Basis Function Neural Network (RBF) and Multilayer Perceptron (MLP) classifiers achieve accurate classification up to 93%. The other research conducts explaining the geometric differences between manual tracings of paint-injected and un-manipulated Placental chorionic surface vascular networks (PCSVNs) under the framework of a shape-context model. The results are these can be matched with nearly 100% accuracy [21]. Related to the leaf classification, the latest study of mango leaf varieties classification uses of K-Support Vector Nearest Neighbor (K-SVNN) to solve multi-class classification [22]. That research uses 300 data generated from mango leaf images, each consist of 256 texture features, two-color features, and 2 shape features. The main features are Weighted Rotation- and Scale-invariant Local Binary Pattern features with average weights

(WRSI-LBP-avg) [23] and achieve the highest accuracy for data with and without reduction is 71.33% and 71.00%, respectively. The other study calculates Boundary Moments as the aggregation of Centroid Contour Distances shape features to help improving performance. By adding these features, the accuracy increases up to 3.8% [24].

CCD provides good shape textures but is inefficient for mango leaf problems because mango leaves have an oval and pointed shapes at the base and tip. The difference among mango leaves can be observed in the edge pattern shown by the base and tip only, while CCD [25] obtained the shape features by calculating the distance from the edge to the center of the object along with 360-degree angles. A high number of features cause high computation and many useless features, especially CCD, that are not from the base and edges. We propose Partial Centroid Contour Distance (PCCD) to generate CCD only from useful edges (leaf base and leaf tip) and increase the accuracy performance of mango leaf classification. PCCD is a modification of CCD, where the calculation of distance is not from the center of the leaf but from the midpoint-cut of the leaf base or leaf tip to the boundary point. The purposes of these features are to capture the leaf base features and leaf tip features of the mango leaf, where specifically, each mango varieties have a slightly different leaf base and leaf tip. This characteristic would be captured by PCCD to improve performance accuracy. The PCCD features are influenced by the width of the leaf base and the leaf tip (Δ), and the angle (α). For features generated, the lower α , the more features are generated. Because the leaves have a base side and tip side, the PCCD is applied to the leaf base and leaf tip as well. The features generated from the leaf base is called PCCD Leaf Base (PCCD-LB), while the features generated from the leaf tip is called PCCD Leaf Tip (PCCD-LT). For feature extraction of the system, if we only generate leaf base and leaf tip features, the large computation of CCD can be reduced. We also explain how PCCD deals with invariant problems such as rotation, translation, or scaling. We prove that PCCD is invariant to these three problems by explaining how PCCD is generated by involving several related leaf components. To justify the quality of PCCD features, we conduct a comparison between PCCD features and the original features of previous studies. Comparisons are conducted between individual features and combinations of features. We also compare performance both without and with data reduction. In this study, we compare the original features of mango leaf classification [24], Centroid Contour Distance (CCD) [25], and a combination of all their features. We also prove the effectiveness of the proposal by comparing our proposed method with other methods, including Moment Invariants [26], Moment Statistics [26], Compactness and Circularity [26], Moment Color, and CCD. We hope that PCCD could contribute to the performance improvement of the mango leaf varieties classification system.

II. MATERIALS AND METHOD

A. Mango Leaf Dataset

The authors use a data set generated from 300 images of mango leaves [24]. Each data is represented by texture, color,

and shape features. The texture features indicated 256 Weighted Rotation-and Scale-invariant Local Binary Pattern features with average weights (WRSI-LBP-avg) [23]. The color features are subject to mean and standard deviation. The shape features are subject to compactness and circularity, and Centroid Contour Distance (CCD). We also generate 181 PCCD-LB features and 181 PCCD-LT features, as we proposed in this study.

B. Centroid Contour Distance (CCD)

Contour shape usually features only exploit shape boundary information. Generally, there are two types of approaches for contour shape modeling: continuous approach (global) and discrete approach (structural). Continuous approaches do not divide the shape into segments, usually a feature vector derived from the integral boundary that is used to describe the shape. The measure of shape similarity is usually a metric distance between the acquired feature vectors. Discrete approaches break the shape boundary into segments, called primitives using a particular criterion. The final representation is usually a string or a graph (or tree); the similarity measure is calculated using string matching or graph matching [25].

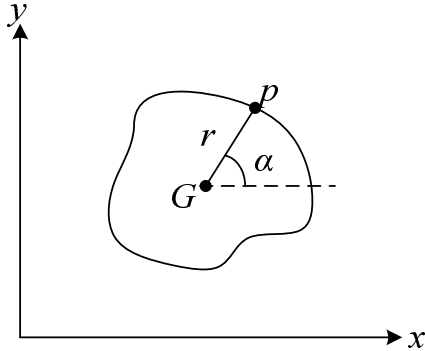


Fig. 1 Illustration of Centroid Contour Distance

Contour shape features can be considered as the distance from the center to the boundary point in a circle with the same angular distance, as described in Figure 1. Point $p(x,y)$ on the boundary is selected from angle direction α from the center point $G(g_x, g_y)$. The distance between center point G to the point p would be Centroid-distance function r [27], using equation (1).

$$r = \sqrt{(G_x - p_x)^2 + (G_y - p_y)^2} \quad (1)$$

The CCD generates distance features from the center to the leaf edge, where each feature is generated at angle α among features. For example, by using α 10 degrees, there are 36 CCD features, begin 0, 10, 20, until 350. On objects with informative shapes along the edges, using CCD would be effective. But in the mango leaf case, only the leaf base and leaf tip are informative in distinguishing the mango leaf varieties; the CCD would be ineffective because most features are useless and deep computation. With the CCD modification, which only takes the leaf base and leaf tip, the informative features only would be generated. These leaf base and leaf tips are the features generated by PCCD. The distance will be calculated from the midpoint-cut of the leaf to the leaf edge. The PCCD still uses α among features. For

leaf base and leaf, the tip would generate along 180 degrees, respectively.

C. Mango Leaf Detection Framework

The system framework for classifying mango leaf varieties is presented in Figure 2. This framework has nine stages, i.e., image acquisition, pre-processing 1, image segmentation, pre-processing 2, features extraction, data splitting, reducing the training data, classifier Training, and data prediction. Image acquisition is conducted by capturing mango leaf using a phone cell camera with resolution 2592x1944 and no effect. The first pre-processing is conducted to remove high-intensity light in the image. Then, we segment the image to obtain the leaf object as the foreground using Otsu thresholding on C , color component [28]. The second pre-processing is stage resizing, cropping, morphological operations, and texture sampling. Then, feature extraction is conducted to obtain 263 features as in previous research [3] and PCCD features as proposed. We use two-fold cross-validation in experiment testing by using 50:50 splitting for training and testing data, respectively. To simplify the training data and speed up the computation, we conduct data reduction. Then, we train the classifier using training data and predict the testing data.

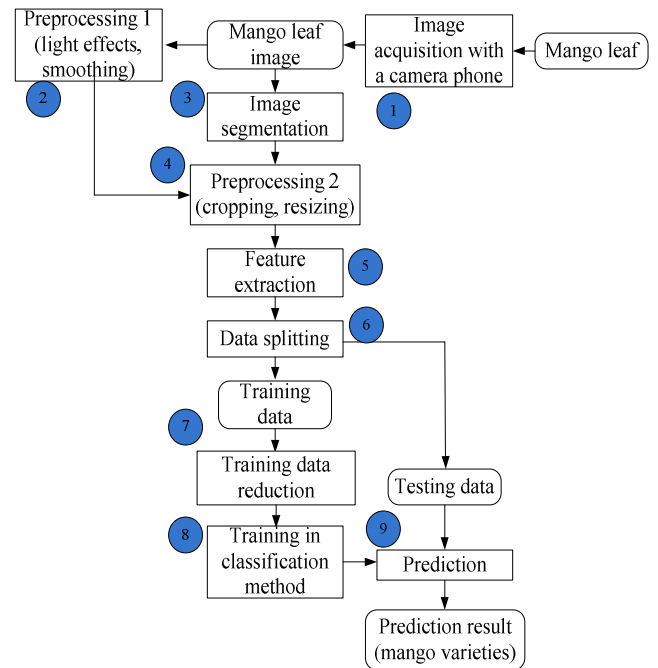


Fig. 2 Mango leaf detection framework

D. Partial Centroid Contour Distance (PCCD) Features

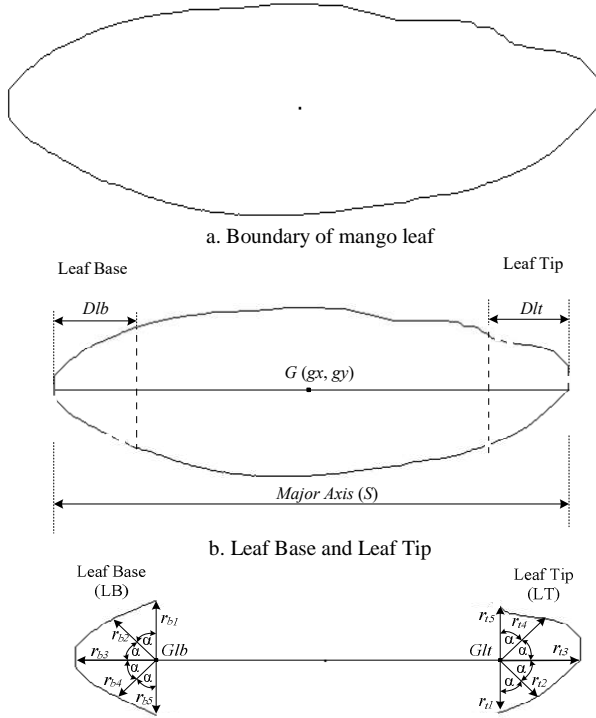
Partial Centroid Contour Distance (PCCD) is the modification of Centroid Contour Distance (CCD) where the CCD measures the distance from the object center $G(g_x, g_y)$ to each selected boundary point, while PCCD measures distance from the midpoint-cut of the base (Glb) or tip of the leaf (Glt) to the selected boundary point of the leaf base (rb_i) and the leaf tip (rt_i).

In the PCCD, when the boundary of the leaf is resulted, as presented in Figure 3(a), we determine the Distance of Leaf Base (Dlb) and Distance of Leaf Tip (Dlt). Dlb is the length of the leaf base which is used to generate PCCD Leaf Base

(PCCD-LB), while Dlt is the length of the leaf tip that is used to generate PCCD Leaf Tip (PCCD-LT). The equations to get the length of Dlb and Dlt as equation (2).

$$Dlb = \frac{\Delta}{100} S, Dlt = \frac{\Delta}{100} S \quad (2)$$

To get Dlb and Dlt , we determine Δ . The Δ is the percentage of leaf base or leaf tip width to the length of the major axis of the leaf, as illustrated in Figure 3(b). In this research, we use to fix the value of the Δ , which is 20.



c. Distance of Glb and Glt to the selected boundary point
Fig. 3 Partial Centroid Contour Distance

Next, based on the Dlb value, we cut the leaf base along with the Dlb value and cut the leaf tip along with the Dlt value, as in Figure 3(b). In the Leaf Base (LB) area, we specify the Glb centre point from the midpoint between the edges, while in the Leaf Tip (LT) area, we specify the Glt centre point from the midpoint between the edges, as in Figure 3(c).

Next, we calculate rb_i and rt_i , rb_i is the distance from Glb to the selected boundary point of LB area. The number of these feature distances (PCCD-LB) are determined by α . The α is the corner width between distance calculation, as in Figure 3(c). The bigger value of α the lesser number of features generated, so the lower α , the more features are generated. For example, in Figure 3(c) using $\alpha = 45$, in the LB area, we would get 5 PCCD-LB samples, which are rb_{b1} , rb_{b2} , rb_{b3} , rb_{b4} , and rb_{b5} . The rb_i distance is determined based on the following equation (3).

$$r_{bi} = \sqrt{(Glb_x - m_{ix})^2 + (Glb_y - m_{iy})^2} \quad (3)$$

Where for m_i is the selected boundary point of the object based on the angle direction α . That calculation method is also conducted to the Leaf Tip (LT) area. The equation for

getting a PCCD-LT sample in LT area, r_{ti} , uses the following equation (4).

$$r_{ti} = \sqrt{(Glt_x - m_{ix})^2 + (Glt_y - m_{iy})^2} \quad (4)$$

The r_{bi} and r_{ti} calculations use the Euclidean distance between the centre point of Glb and the Glt to the selected boundary point so that when the leaf object is rotated, this distance calculation is not affected. So, PCCD is invariant to rotation. The subtraction used at Euclidean distance is also based on both Glb and Glt location and the boundary points. The object's location shifting in the image also does not give distance difference obtained. Hence, PCCD is also invariant to translation.

To solve the scale-invariant problem, we use the major axis length (S) as the control of r_{bi} and r_{ti} values. On the same leaf, when taken with different image size or different leaf object size, the r_{bi} and r_{ti} length are also different. To avoid this, we do normalization by dividing the r_{bi} and r_{ti} values with the major axis (S). The equation is as follows.

$$r_{bi}^{norm} = \frac{r_{bi}}{S} \text{ and } r_{ti}^{norm} = \frac{r_{ti}}{S} \quad (5)$$

So, based on the example in Figure 3(c), using $\alpha = 45$, then 5 r_b leaf base features and 5 r_t leaf tip features are generated. The PCCD features generated from the leaf base is called PCCD Leaf Base (PCCD-LB), while PCCD features generated from the leaf tip area is called PCCD Leaf Tip (PCCD-LT).

III. RESULTS AND DISCUSSION

A. Testing intra-PCCD

In our experiment of PCCD, we use $\Delta = 20$, so the Dlb and Dlt width is 20% of the leaf width. To prove the quality of PCCD features, we conduct empirically testing using $\alpha = 90$ until 5. The lower alpha used, the more PCCD features are generated. In PCCD-LB, the $\alpha = 90$ indicates 3 features generation by calculating the distance at 270, 0, and 90 angles. While in PCCD-LT, the $\alpha = 90$ indicates 3 features generation by calculating distances at angles 90, 180, and 270, as illustrated in Figure 2(c). Some examples of variation α are presented in Table 1.

TABLE I
NUMBER OF GENERATED FEATURE

α	PCCD-LB	PCCD-LT
90	3	3
45	5	5
30	7	7
15	15	15
10	19	19
5	37	37

The empirically testing conducted is encompassed with and without K-SVNN data reduction. The results are presented in Figure 3. We compared the classification accuracy between PCCD-LB vs. PCCD-LT vs. the combination of both. Both comparisons were conducted on classification without data reduction and data reduction. The

results presented in Figure 4 (a), graphs with dashed lines, dotted lines and solid lines respectively are leaf base features, leaf tip features and combination of both. From the graph, it can be observed that the Leaf tip feature provides the lowest accuracy, around 35% to 47%. Leaf base features provide better accuracy by 44% to 60%. The combination of the two features provides the best accuracy, ranging from 48% to 64%; the highest accuracy achieved at 64.67%. The accuracy chart pattern on all features always fluctuates up or down. In general, using a combination of features can achieve higher accuracy, while the highest accuracy is achieved by the leaf range of 45 to 65.

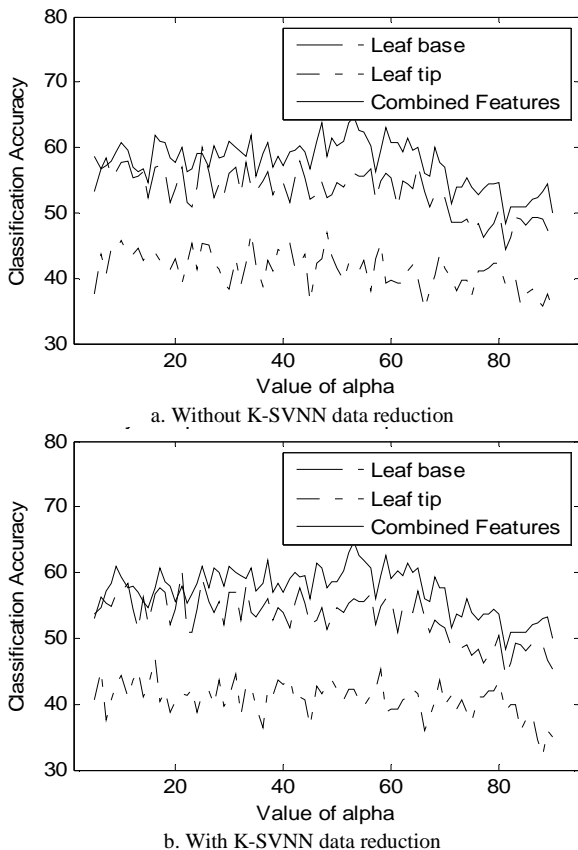


Fig. 4 The result of empirically testing

In the same testing but accompanied by K-SVNN data reduction, the highest accuracy was achieved using the combination of both features, followed by Leaf base and Leaf tip features, respectively. The highest accuracy with the combination of both features is achieved at alpha 53. The pattern of results on all combinations of features is almost the same between without and with data reduction, where the highest accuracy is achieved by the combination of both leaf base and leaf tip features. This proves that data reduction does not affect the performance of all these features.

B. Comparison Result with Other Features

The authors also conduct classification testing and compare PCCD with some other features as follows:

- 7 Moment Invariants
- 6 Moment Statistics
- Local Binary Pattern
- WRSI-LBP, Compactness, and Circularity

- 2 Moment Colors
- 3 Boundary Moments
- CCD. In this test, we use the assumption $\alpha = 45$.

Each PCCD-LB and PCCD-LT uses 5 features, while CCD uses 8 features. From the results presented in Table 2, PCCD relatively uses fewer features than other features, such as LBP and WRSI-LBP, where each PCCD-LB and PCCD-LT uses 5 features while LBP and WRSI-LBP 256 features, respectively. Comparison with other methods such as Moment Invariants [26], Moment Statistics [26], Compactness and Circularity [26], Moment Color, and CCD, it is seen that PCCD-LB and PCCD-LT use almost the same number of features, this indicates that PCCD needs same computation load compared to the other features. In this testing, the combination of PCCD-LB and PCCD-LT provides the best accuracy among all other methods but with a slight difference with WRSI-LBP, where the accuracy is 53.13% and 52.93% respectively for the combination of PCCD and WRSI-LBP. From these results, WRSI-LBP has similar performance to PCCD, but PCCD has fewer features than WRSI-LBP.

From the results of this comparison, it can be concluded that the combination of PCCD-LB and PCCD-LT achieves better performance than others, but the system performance is not optimal because the accuracy achieved is only 53.13%. We also conduct combination testing between PCCD and the original features of mango leaf classification, as presented in the next section.

TABLE II
COMPARISON OF ACCURACY (%) WITH VARIOUS OTHER FEATURES

Features	Count of Features	Accuracy
Moment Invariants	7	33.77
Moment Statistics	6	47.47
LBP	256	51.80
WRSI-LBP	256	52.93
Compactness and Circularity	2	42.40
Moment Color	2	39.30
Boundary Moments	3	38.83
CCD	8	50.3
PCCD Leaf Base	5	48.8
PCCD Leaf Tip	5	36.2
Combined PCCD-LB and PCCD-LT	10	53.13

C. Comparison Result with Original Features

The original features of mango leaf classification use 263 features, consisting of 256 WRSI-LBP features, an average of grey images, the standard deviation of the grey image, compactness, circularity, and 3 Boundary Moments of CCD [24]. PCCD-LB and PCCD-LT use feature under Table 1, CCD uses generated features from [27], while the combination features use a combination of all features mentioned before. We conducted a comparison of original features, PCCD-LB, PCCD-LT, CCD, and all combination features without and with data reduction. The results are presented in Table 3 and Table 4.

From the data presented in Table 3, at all α , the highest accuracy is obtained when using a combination of all features. The highest accuracy is obtained for $\alpha = 15$ with an accuracy of 81.73%. The result presented in Table 3 shows

our comparison of all various features used in classification. We compare the original features, individual PCCD features (Leaf base and Leaf tip), and a combination of Leaf base and Leaf tip. These results show that the combination of original features and PCCD features achieve the best accuracy of 80.17% and average accuracy of 78.41%. This accuracy is increased by almost 10% from the accuracy of the original features, 0.17%, and 69.85% for the best and average accuracy, respectively. We also use Centroid Contour Distance (CCD) features as comparison features with PCCD. When the system uses CCD features only, the accuracy achieved is 56.67% and 55.48% for the best and average accuracy, respectively. Still, when the system uses a combination of original and PCCD features, the performance increases up to 79.97% and 78.73% for the best and average accuracy, respectively. We also conduct a comparison of the combination of CCD and PCCD as a feature in classification. There is an interesting one where the accuracy achieved is just 59.77% and 58.62% for the best and average, respectively, but when the system uses a combination of the original features, CCD, and PCCD, the accuracy achieve 81.73% and 79.87% for the best and average, respectively. These results show the combination of textures, color, and shape features can increase the classification accuracy.

TABLE III
ACCURACY (%) OF SVM CLASSIFICATION WITH VARIOUS SELECTED α AND WITHOUT K-SVNN DATA REDUCTION

Features	α					Avg
	90	45	30	15	10	
Orig. Features (263 features)	70.17	69.67	69.53	69.93	69.93	69.85
Leaf base	47.83	54.07	53.57	54.73	56.00	53.24
Leaf tip	35.23	41.87	42.93	41.73	43.00	40.95
PCCD (Leaf base + Leaf tip)	52.57	59.53	59.07	58.33	58.47	57.59
Orig. Features + PCCD	75.83	77.60	79.53	80.17	78.93	78.41
CCD	56.43	56.67	53.73	56.53	54.03	55.48
Orig. Features + CCD	78.03	78.40	78.47	79.97	78.80	78.73
CCD + PCCD	55.70	59.93	59.77	59.50	58.20	58.62
Orig. Features + CCD + PCCD	78.50	79.17	80.00	81.73	79.93	79.87

TABLE IV
ACCURACY OF SVM CLASSIFICATION WITH VARIOUS SELECTED α AND WITH K-SVNN DATA REDUCTION

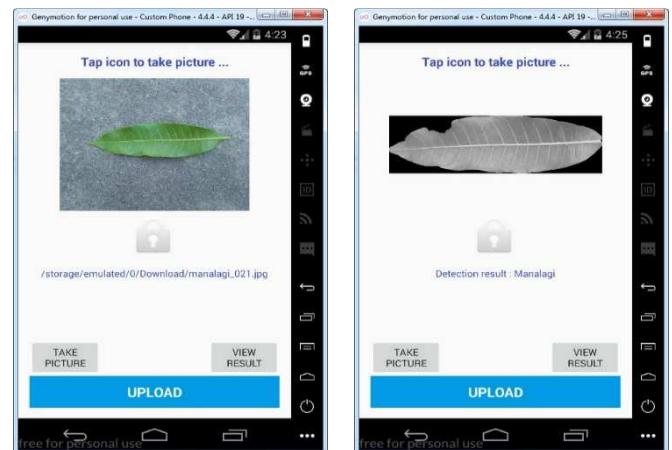
Features	α					Avg.
	90	45	30	15	10	
Orig. Features (263 features)	66.80	66.13	66.37	66.57	67.07	66.59
Leaf base	47.43	53.60	53.67	54.53	56.23	53.09
Leaf tip	35.50	41.30	42.47	41.97	42.13	40.67
PCCD (Leaf base + Leaf tip)	52.50	59.40	58.03	56.87	57.57	56.87
Orig. Features + PCCD	72.40	74.63	75.37	76.33	76.47	75.04
CCD	57.23	57.30	54.90	57.20	53.67	56.06
Orig. Features + CCD	73.70	74.87	74.03	75.67	75.93	74.84
CCD + PCCD	56.77	60.27	58.63	57.73	57.23	58.13
Orig. Features + CCD + PCCD	74.77	76.53	75.73	77.87	77.73	76.53

The authors also conduct comparison by adding data reduction with K-SVNN before the training session. Data reduction aims to simplify the data processed in the training

session. The number of data used during training is reduced, so the training process runs faster, but the accuracy performance is decreased in all variation of α and all feature combination options. As presented in Table 3 and Table 4, the accuracy of original features decreases from 69.85% to 66.59% from without to with data reduction, respectively. The leaf base, leaf tip, combination of leaf base and leaf tip, a combination of original features and PCCD, CCD, a combination of CCD and PCCD, and a combination of all features obtain decreasing performance because of data reduction. The combination of original features and PCCD features achieves an accuracy of 76.47% and 75.04% for the best and average, respectively. The results also show that when the system uses a combination of original features and CCD, the accuracy achieves 75.93% and 74.84% for the best and average. The combination of all features also achieves the best accuracy among all; the system achieves an accuracy of 77.87% and 76.53 for the best and average, respectively. Moreover, compared to all feature combinations, it is concluded that the combination of PCCD increases the accuracy performance of the mango leaf classification. From all testing results, it shows that the system does not achieve good performance when using PCCD as features individually, but when combined with the other features, i.e., texture features or CCD features, the system achieves good accuracy performance.

D. Testing in Android Application

The authors implement the classification of mango leaf varieties in software that works on the Android operating system. The application is developed using Android Studio 2.3.3. The testing is conducted using the Genymotion emulator 2.9.0 and a phone cell with Android versions 5 and 6. The environment testing used by the authors as follow: image acquisition on one leaf only, the time testing is in the morning where the sun exposes mango leaves directly, the camera effect used is normal, and the distance between the leaves and the camera about 10-20 cm.



a. Capture the mango leaf b. The result of the classification

Fig. 5 Application of mango leaf detection

As shown in Figures 5 (a), it appears that the image acquisition can be made using the camera or using files stored on the phone cell, while Figure 5(b) presents the results of the detection. As soon as the detection process is

complete, the 'see results' button can be used to display the detected image and the name of mango species obtained.

IV. CONCLUSION

This study concluded that PCCD is an informative shape feature that supports improved classification performance. Using only the leaf base and leaf tip as generated features, PCCD can increase accuracy by up to 81.74%. The PCCD features are divided into two parts, namely PCCD-LB and PCCD-LT. On individual testing, the system cannot achieve high performance but increase accuracy up to 10% when combined with original features. The comparison testing with other features also proves that the PCCD combination features are more effective than others, although the difference is slight compared to WRSI-LBP PCCD uses fewer features, so the computing load system is also lighter. The important next study is the number of raw features generated by PCCD is still a lot, so it requires high computation. We need further modification so that PCCD can be summarized into fewer features with the same classification strength but more straightforward in the calculation.

ACKNOWLEDGMENT

This research was funded by the Directorate of Research and Community Service (DRPM) DIKTI year 2018 with contract number 009/SP2H/LT/K7/KM/2018 on 26 February 2018 to Universitas Bhayangkara Surabaya collaborated with Institut Teknologi Sepuluh Nopember in Inter-Universities Research Cooperation (PKPT) scheme.

REFERENCES

- [1] S. Thawkar and Ingolikar, R., "Automatic Detection and Classification of Masses in Digital Mammograms," *Int. J. Intell. Eng. Syst.*, vol. 10, no. 1, pp. 65–74, 2017.
- [2] K. G. Satheesh and A. N. J. Raj, "Medical image segmentation and classification using MKFCM and hybrid classifiers," *Int. J. Intell. Eng. Syst.*, 2017, doi: 10.22266/ijies2017.1231.02.
- [3] G. J. Jong, Hendrick, Z. H. Wang, D. Kurniadi, Aripriharta, and G. J. Horng, "Implementation of Otsu's method in vein locator devices," *Int. J. Adv. Sci. Eng. Inf. Technol.*, 2018, doi: 10.18517/ijaseit.8.3.4414.
- [4] A. P. Rahmadini, P. Kristalina, and A. Sudarsono, "Optimization of fingerprint indoor localization system for multiple object tracking based on iterated weighting constant - KNN method," *Int. J. Adv. Sci. Eng. Inf. Technol.*, 2018, doi: 10.18517/ijaseit.8.3.6086.
- [5] K. S. Kumar and G. Sreenivasulu, "Image enhancement through denoising and retrieval of vegetation parameters from Landsat8," *Int. J. Adv. Sci. Eng. Inf. Technol.*, 2018, doi: 10.18517/ijaseit.8.1.4059.
- [6] Y. Shi *et al.*, "Tree species classification using plant functional traits from LiDAR and hyperspectral data," *Int. J. Appl. Earth Obs. Geoinf.*, vol. 73, pp. 207–219, Dec. 2018, doi: 10.1016/j.jag.2018.06.018.
- [7] Y. Zhang, W. S. Lee, M. Li, L. Zheng, and M. A. Ritenour, "Non-destructive recognition and classification of citrus fruit blemishes based on ant colony optimized spectral information," *Postharvest Biol. Technol.*, vol. 143, pp. 119–128, Sep. 2018, doi: 10.1016/j.postharvbio.2018.05.004.
- [8] M. A. Peña, R. Liao, and A. Brenning, "Using spectrottemporal indices to improve the fruit-tree crop classification accuracy," *ISPRS J. Photogramm. Remote Sens.*, vol. 128, pp. 158–169, Jun. 2017, doi: 10.1016/j.isprsjprs.2017.03.019.

- [9] O. M. Ben Saeed *et al.*, "Classification of oil palm fresh fruit bunches based on their maturity using portable four-band sensor system," *Comput. Electron. Agric.*, vol. 82, pp. 55–60, Mar. 2012, doi: 10.1016/j.compag.2011.12.010.
- [10] Y. Zhang, S. Wang, G. Ji, and P. Phillips, "Fruit classification using computer vision and feedforward neural network," *J. Food Eng.*, vol. 143, pp. 167–177, Dec. 2014, doi: 10.1016/j.jfoodeng.2014.07.001.
- [11] J. Steinbrener, K. Posch, and R. Leitner, "Hyperspectral fruit and vegetable classification using convolutional neural networks," *Comput. Electron. Agric.*, vol. 162, pp. 364–372, Jul. 2019, doi: 10.1016/j.compag.2019.04.019.
- [12] A. Sinha and R. S. Shekhawat, "Olive Spot Disease Detection and Classification using Analysis of Leaf Image Textures," in *Procedia Computer Science*, Jan. 2020, vol. 167, pp. 2328–2336, doi: 10.1016/j.procs.2020.03.285.
- [13] M. Ben Haj Rhouma, J. Žunić, and M. C. Younis, "Moment invariants for multi-component shapes with applications to leaf classification," *Comput. Electron. Agric.*, vol. 142, pp. 326–337, Nov. 2017, doi: 10.1016/j.compag.2017.08.029.
- [14] S. Zhang, W. Huang, Y. Huang, and C. Zhang, "Plant Species Recognition Methods using Leaf Image: Overview," *Neurocomputing*, May 2020, doi: 10.1016/j.neucom.2019.09.113.
- [15] G. Saleem, M. Akhtar, N. Ahmed, and W. S. Qureshi, "Automated analysis of visual leaf shape features for plant classification," *Comput. Electron. Agric.*, vol. 157, pp. 270–280, Feb. 2019, doi: 10.1016/j.compag.2018.12.038.
- [16] C. Kalyoncu and Ö. Toygar, "Geometric leaf classification," *Comput. Vis. Image Underst.*, vol. 133, pp. 102–109, Apr. 2015, doi: 10.1016/j.cviu.2014.11.001.
- [17] A. Hasim, Y. Herdiyeni, and S. Douady, "Leaf Shape Recognition using Centroid Contour Distance," in *IOP Conference Series: Earth and Environmental Science*, Feb. 2016, vol. 31, no. 1, doi: 10.1088/1755-1315/31/1/012002.
- [18] M. S. Mohd Asaari, S. A. Suandi, and B. A. Rosdi, "Fusion of Band Limited Phase only Correlation and Width Centroid Contour Distance for finger based biometrics," *Expert Syst. Appl.*, 2014, doi: 10.1016/j.eswa.2013.11.033.
- [19] B. Wang, D. Brown, Y. Gao, and J. La Salle, "MARCH: Multiscale-arch-height description for mobile retrieval of leaf images," *Inf. Sci. (Nijl.)*, 2015, doi: 10.1016/j.ins.2014.07.028.
- [20] J. R. Kala and S. Viriri, "Plant species classification using sinuosity coefficients of leaves," *Image Anal. Stereol.*, 2018, doi: 10.5566/ias.1821.
- [21] E. Farnell *et al.*, "A shape-context model for matching placental chorionic surface vascular networks," *Image Anal. Stereol.*, 2018, doi: 10.5566/ias.1708.
- [22] E. Prasetyo, R. D. Adityo, N. Suciati, and C. Fatchah, "Multi-class K-support vector nearest neighbor for mango leaf classification," *Telkonnika (Telecommunication Comput. Electron. Control.)*, vol. 16, no. 4, 2018, doi: 10.12928/TELKOMNIKA.v16i4.8482.
- [23] E. Prasetyo, R. D. Adityo, N. Suciati, and C. Fatchah, "Average and maximum weights in weighted rotation- and scale-invariant LBP for classification of mango leaves," *J. Theor. Appl. Inf. Technol.*, vol. 95, no. 21, 2017.
- [24] E. Prasetyo, R. D. Adityo, N. Suciati, and C. Fatchah, "Mango Leaf Classification with Boundary Moments of Centroid Contour Distances as Shape Features," 2018, doi: 10.1109/ISITIA.2018.8711115.
- [25] D. Zhang and G. Lu, "Review of shape representation and description techniques," *Pattern Recognit.*, 2004, doi: 10.1016/j.patcog.2003.07.008.
- [26] R. C. Gonzalez and R. E. Woods, *Digital Image Processing (3rd Edition)*. 2007.
- [27] Y. Mingqiang, K. Kidiyo, and R. Joseph, "A Survey of Shape Feature Extraction Techniques," in *Pattern Recognition Techniques, Technology and Applications*, 2008.
- [28] E. Prasetyo, R. D. Adityo, N. Suciati, and C. Fatchah, "Mango leaf image segmentation on HSV and YCbCr color spaces using Otsu thresholding," 2017, doi: 10.1109/ICSTC.2017.8011860.