TEKNIKA: JURNAL SAINS DAN TEKNOLOGI

Homepage journal: http://jurnal.untirta.ac.id/index.php/ju-tek/

Application of AutoML Vision for palm identification at UIN Syarif Hidayatullah Jakarta

Victor Amrizal [a], ***Hendra Bayu Suseno*** [a], ***Priyanti Priyanti*** [b,1]

[a] *Department of Informatics Engineering, Faculty of Science and Technology, Universitas Islam Negeri Syarif Hidayatullah Jakarta, Jl. Ir. H. Juanda No. 95, Ciputat, 15421, Indonesia*

[b] *Department of Biology, Faculty of Science and Technology, Universitas Islam Negeri Syarif Hidayatullah Jakarta, Jl. Ir. H. Juanda No. 95, Ciputat, 15421, Indonesia*

[1] Corresponding author: priyanti@uinjkt.ac.id

ARTICLEINFO

Article history:
Submitted 25 March 2022
Reviewed 31 May 2022
Revised 28 June 2022
Accepted 29 June 2022
Available online on 29 June 2022

Keywords:
AutoML Vision, expert system, feature extraction, plant classification.

Kata kunci:
AutoML Vision, sistem pakar, ekstraksi fitur, klasifikasi tanaman.

ABSTRACT

Palm is a monocot plant that belongs to the Arecaceae family. The plant has unique characteristics and reasonably high diversity in style and shape. However, some people do not know the species of palms around them. Plant identification is made by looking at the shape of the leaves because each species of palm has leaves with unique features. Digital identification of plants is important because it is effective and efficient. One of the tools that can be used to identify plants is AutoML Vision because the tool is easy to develop and can identify quite accurately. The application of AutoML Vision in identifying palms is carried out at the Syarif Hidayatullah State Islamic University Jakarta campus. The application of AutoML Vision is carried out at the UIN Syarif Hidayatullah Jakarta campus. The results showed an accuracy rate of 90%, and for a threshold value of 0.9, a precision value of 100%.

ABSTRAK

Palem adalah salah satu tumbuhan monokotil, yang termasuk dalam suku Arecaceae. Tumbuhan ini memiliki sifat yang sangat unik serta keragaman yang cukup tinggi, baik dilihat dari corak maupun bentuk. Namun, sebagian masyarakat tidak mengetahui jenis palem yang ada di sekitarnya karena keragaman tersebut. Identifikasi tumbuhan secara digital penting dilakukan karena efektif dan efisien. Salah satu alat yang dapat digunakan untuk mengidentifikasi tumbuhan yaitu menggunakan AutoML Vision, karena alat itu mudah dikembangkan dan dapat mengidentifikasi dengan cukup akurat. Identifikasi tumbuhan dilakukan dengan melihat bentuk daun, karena tiap jenis palem memiliki daun dengan fitur yang unik. Fitur daun tersebut membawa informasi penting yang dapat membantu manusia mengenali dan mengklasifikasikan tanaman yang dilihatnya secara digital. Penerapan AutoML Vision dilakukan di lingkungan kampus UIN Syarif Hidayatullah Jakarta. Hasil penelitian menunjukkan tingkat akurasi sebesar 90%, dan untuk nilai threshold 0.9 didapat nilai presisi 100%.

Available online at http://dx.doi.org/10.36055/tjst.v18i1.14599

1. Introduction

The campus of UIN Syarif Hidayatullah is a habitat for various kinds of plants, both monocots and dicots [1]. Some of these plants belong to the Fabaceae [2] and Arecaceae tribes. These plants are used for ornamental plants, road shade, noise suppression, or dust absorbers [2][3]. One of the monocotyledonous plants of the Arecaceae tribe around the campus is a palm tree, also known as a multipurpose plant. The plant has a unique nature and a reasonably high diversity, widely used as an ornamental plant that is usually in the yard or the office because of its beautiful leaves and stems.

However, some people do not know the species of palms around them. The identification of palms has problems when only vegetative characters (stems and leaves) are found. At this time, plant identification is often made manually by matching the plant characters encountered with the characters in plant taxonomic references, so it takes a long time. Therefore, digital technology is needed to speed up the identification process [4-6]. The plant classification method widely used in research is identification based on leaf shape [7][8] because the leaf shape of each plant has different features [7]. These features carry essential information that can help humans recognize and classify plants [9].

Digital identification with good accuracy has been carried out on legumes (Papilionaceae) using a forward chaining approach to an expert system. This process is done through image processing, determining aspect ratio, form factor, leaf tip, and diagonal [10]. In addition, the identification of real-time objects based on Android has also been carried out through the convolutional neural network method using AutoML Vision to create a model and the 10-fold cross-validation method to determine the validation value of the test results [11]. AutoML Vision can help developers understand image content by encapsulating powerful machine learning models with easy-to-use interfaces, REST APIs. In addition, AutoML Vision can classify images into thousands of categories, detect objects, detect human faces, and quickly find and read words printed in images.

This study uses AutoML Vision to identify palms because it is fast and accurate. AutoML Vision used in this research is equipped with RAD system development. Rapid application development (RAD) is an object-oriented approach for rapid system development through iteration [12][13], so it takes a short time. AutoML Vision is applied at the Syarif Hidayatullah State Islamic University Jakarta campus by considering the evaluation results of training data, including precision, recall, and accuracy based on the threshold value.

2. Research Methodology

This research was conducted in April-October 2020 at UIN Syarif Hidayatullah Jakarta, and the identification of palms was carried out at the Integrated Central Computer Laboratory, UIN Syarif Hidayatullah Jakarta. The study was conducted by taking four individuals from each species of palm, then taking pictures with a mobile phone camera. Some of these images were used as research samples.

Palm identification is made with AutoML Vision using the RAD approach. The RAD system development method has three stages: requirements planning, workshop design, and implementation. The requirements planning stages include

a. Collect valid data through literature study, interviews, and observations
b. Examine the current system methods so that researchers can formulate better solutions
c. Identify the features that will be applied to the system that will be made for research, and
d. Design applications with hardware and software to design applications.

The second stage is the design workshop, which is carried out with a design and build system. The design and coding process is carried out by correcting design discrepancies. Writing code using android studio and java programming language, then creating a dataset. The stage goes through the steps below:

a. Design a model to classify with the AutoML Vision model using the ML Kit feature from firebase to make it easier to create a model to classify images.
b. Design architecture of the system
c. Design process on the system with UML using Visio 2016 software by making three kinds of diagrams: use case diagrams, activity diagrams, and class diagrams.
d. Design system user interface to facilitate system use.

The last stage is the implementation or installation stage on android devices and application testing using the black box method.

3. Results and Discussion

3.1. Problem Identification

Identification of Arecaceae plants at UIN Syarif Hidayatullah Jakarta Campus through leaf shape is made using digital image processing and machine learning for the classification stage. There are eight species of palm: raja, yellow, red, waregu, bamboo, fishtail, fan, and bottle. Photos of each palm plant were used as research samples. The system needed for identification is the AutoML Vision Edge feature, which can be run offline. Machine learning features are available on the www.firebase.com website to train, deploy, and serve datasets. The application features used in this research are camera input and library input. The camera input feature is used to enter images using the camera, while the library input feature is used to enter images in the library.

Before coding or creating an application, prepare a model that will be used in the application. Models are created using AutoML in the firebase console. The steps taken in the AutoML process are collecting training data, training the model, and evaluating the model. The collection of training data or datasets is done using machine learning, with datasets available on the internet or creating datasets. The dataset can be in the form of images of Arecaceae plants growing in the area of UIN Syarif Hidayatullah Jakarta or download images from the palmpedia.net site, planet.org. The dataset contains more than 100 images on each label or species of palm plant. The dataset used to train the model in AutoML Vision Edge must meet the following criteria:

a. Has the formats: JPEG, PNG, GIF, BMP, ICO.
b. The maximum image size is 30 MB.

AutoML Vision Edge may degrade the quality of most images during initial processing but generally does not affect accuracy. Increasing the value of accuracy can be done by collecting at least 100 sample images of each species of palm, and each image has a variety of angles, resolutions, and different backgrounds. In addition, the training data should be as close as possible to the data on which the predictions are based. If a low-resolution, a blurry image is present (such as an image from a security camera), the training data must consist of a low-resolution blurry image. After setting up training images, prepare them for import into Firebase using the structured zip archive. The import of images is done by creating a directory named according to the species of palm, and each appropriate image is inserted into it, then compressing the directory structure into a zip archive as shown in Figure 1.

The next step is to create a model using AutoML found in the Firebase console at www.firebase.com. Model evaluation can be done by clicking on the model on the dataset details page to view performance metrics on the model. One important use of the page is to determine the most appropriate score threshold for the model. The score threshold is the minimum confidence a model must have to assign a label to an image. Changes in the threshold score can be used to influence the threshold in influencing the model's performance. Model performance is measured using two metrics, namely precision, and recall. In image classification, precision is the ratio between the number of correctly labeled images and the number of images labeled by the model based on the selected threshold. The model that is the least mislabeled is called the high-precision model (fewer false positives). Recall is the ratio between the number of correctly labeled images and the number of images that have content that the model should be able to label. A model with high recall rarely fails to assign any labels (fewer false negatives). The results of this study support research [14], which states that AutoML can be used for tuning many models.

Comparison of model performance on each label can use a confusion matrix. The ideal model has a high diagonal value, and the other values will be low. It shows that the image identified is correct. If there is another high value, it shows how the model classifies its test image, as shown in Figures 2. Table 1 shows the confusion matrix, which shows how often the model correctly classifies each label in blue and which labels are most frequently confused for that label in yellow.

```
dataset.zip
 |----raja
 | |----01.jpg
 | |----02.jpg
 | |----03.jpg
 |----kuning
 | |----kn-01.jpg
 | |----kn-02.jpg
 | |----kn-03.jpg
 |----merah
 | |----mr-01.jpg
 | |----mr-02.jpg
 | |----mr-03.jpg
```

Figure 1. Dataset in zip.

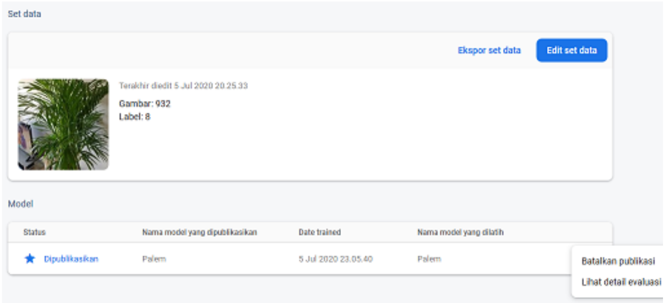

Figure 2. Performance matrix.

3.2. Plant Identification Stages

The process of identifying plants digitally in this research consists of several processes, namely training dataset, preprocessing (input, Conv, Relu, Pool, Fully-Connected), and identification. The complete process is shown in Figure 3. The training image and test image will enter the preprocessing process. In the training process, the dataset creation will be directly processed in AutoML Vision contained in the firebase console and generate a model that will be used as training data. Next, the training data is downloaded and stored in the application folder (app/assets). The model data is converted into Javascript Object Notation (JSON). The test image will be sent to AutoML Vision Edge contained in the application by using image images that are on the sending android and displaying the data that is most similar or has the closest weight value to the test image along with the mention of the image label, which is shown in Figure 4.

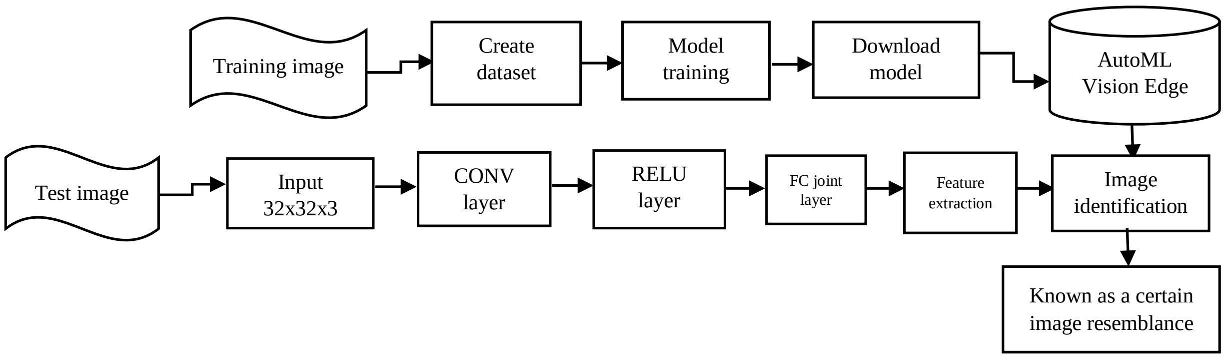

Figure 3. Overview of digital plant identification system.

The stages of the dataset training process using Firebase AutoML Vision are carried out by entering the image sent to AutoML into the appropriate dataset based on the image label. If the number of images exceeds 100, it will continue with the modeling process. If the image data is less than 100, it will be asked to enter additional images. The stages of sending and retrieving data from AutoML in the form of JSON, which will be used for comparison data or image identification that will display labels and weight values, are shown in Figure 5.

3.3. Implementation

After the model training process is complete, there is an evaluation menu that will display the results of the training model evaluation in the form of details of total images containing many training images. The threshold value serves as a comparison of training results. If the threshold value is low, the model will classify more images, but it risks classifying some images inaccurately in the classification process. If the threshold is high, the model will classify fewer images but have a lower risk of classifying images incorrectly. A confusion matrix, which shows true positive and true negative comparisons, is used to measure how well the model performs on each label. In an ideal model, all values on the diagonal will be high, and all other values will be low, which indicates that the image identified is correct. If another high value is found, this shows how the model classifies the test image. It is shown in Figures 6 and Table 1.

Based on the evaluation results of the training data model, the results of precision, recall, and accuracy values are obtained based on the threshold value with a value between 0–1, which is shown in Table 2. The highest accuracy value is found at the threshold value of 0.4, with an accuracy value of 0.66, precision of 0.67, and recall value of 0.65. The threshold value of 0.5 is an accuracy value of 0.65, precision of 0.81, and recall with a value of 0.54. The lowest accuracy value is found at the threshold value of 0.9, with an accuracy value of 0.12, precision of 1.0, and recall value of 0.062. The precision, recall, and accuracy value on each label with a threshold value of 0.4. The values obtained are the average accuracy value of 0.64, the average precision value of 0.64, and the average recall value of 0.69. The lowest accuracy value is in the bamboo image with an accuracy value of 0.4, and the highest is found in the red image with an accuracy value of 0.80, shown in Table 3.

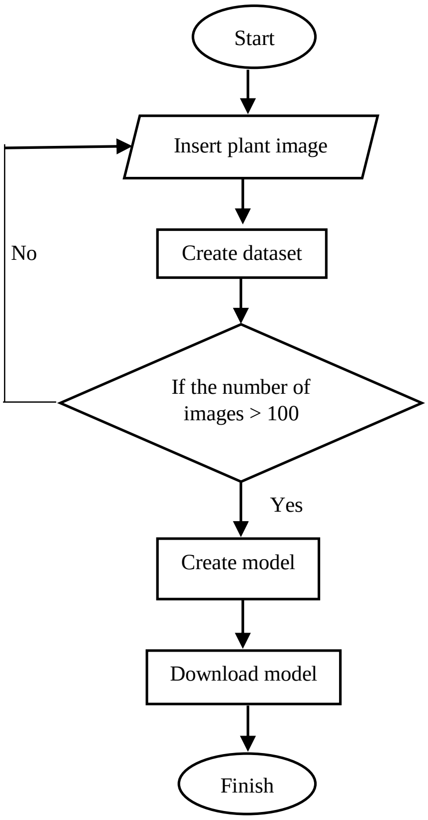

Figure 4. **Stages of the dataset training process.**

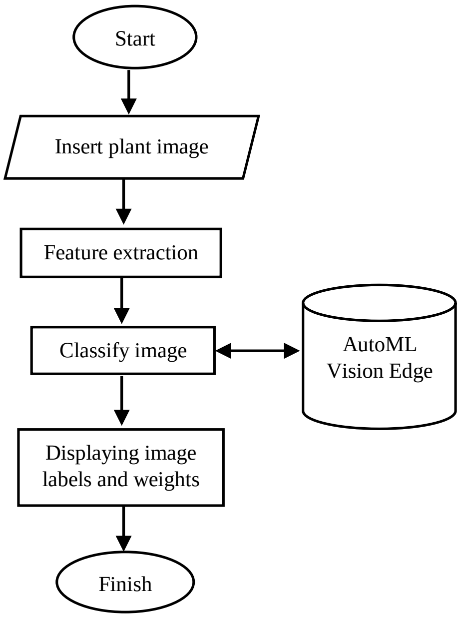

Figure 5. **Stages of the system process in the application.**

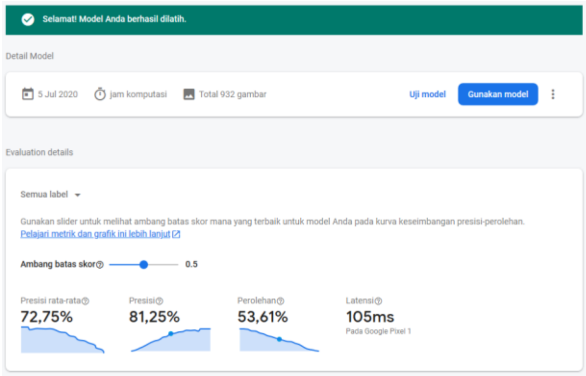

Figure 6. **Evaluation results.**

Table 1. Confusion matrix.

Predicted label / **Label true**	**Bamboo**	**Waregu**	**Bottle**	**Fishtail**	**Fan**	**Red**	**Yellow**	**Raja**
Bamboo	45.5%	18.2%	-	-	-	9.1%	27.3%	-
Waregu	-	50.0%	-	-	-	10.0%	30.0%	10.0%
Bottle	-	8.3%	66.7%	-	16.7%	-	8.3%	-
Fishtail	-	-	-	91.7%	-	-	-	8.3%
Fan	-	18.2%	-	9.1%	45.5%	-	18.2%	9.1%
Red	-	-	-	-	-	91.7%	8.3%	-
Yellow	5.9%	17.6%	5.9%	-	-	-	64.7%	5.9%
Raja	-	-	8.3%	8.3%	-	-	8.3%	75.0%

Table 2. Recapitulation of precision, recall, and accuracy values based on each threshold.

Threshold	Precision	Recall	F1-score
0.0	0.13	1	0.23
0.1	0.23	0.98	0.37
0.2	0.38	0.89	0.53
0.3	0.53	0.77	0.63
0.4	0.67	0.65	0.66
0.5	0.81	0.54	0.65
0.6	0.87	0.43	0.57
0.7	0.94	0.35	0.51
0.8	0.95	0.20	0.33
0.9	1	0.062	0.12
1.0	0	1	0

Table 3. Recapitulation of precision, recall, and accuracy values based on threshold 0.4.

Label	Precision	Recall	F1-score
Fishtail	0.83	0.83	0.83
Bamboo	0.75	0.27	0.4
Yellow	0.57	0.7	0.63
Waregu	0.5	0.6	0.54
Raja	0.62	0.83	0.71
Red	0.77	0.83	0.8
Fan	0.62	0.45	0.52
Bottle	0.87	0.58	0.7
Average	0.69	0.64	0.64

The values of precision, recall, and accuracy for each label with a threshold value of 0.5 are shown in Table 4. The values obtained are the average accuracy value of 0.60, the average precision value of 0.73, and the recall average value of 0.53. The lowest accuracy value is in the bamboo image with an accuracy value of 0, and the highest is in the red image with an accuracy value of 0.81. This precision value is similar to research [15] which used Auto ML to classify coffee varieties. The values obtained are the average accuracy value of 0.38, the average precision value of 1, and the average recall value of 0.34. The lowest accuracy value is in the image of bamboo, yellow, waregu, and fan with an accuracy value of 0, and the highest value is found in the image of red, fishtail, and bottle with an accuracy value of 0.9. Based on the test on the threshold value above, it can be concluded that the average value of precision, recall, and accuracy is considered quite good because it passes 0.5 with a threshold value of 0.5, shown in Table 5.

Table 4. Recapitulation of precision, recall, and accuracy values based on threshold 0.5

Label	Precision	Recall	F1-score
Fishtail	0.82	0.75	0.78
Bamboo	0	0	0
Yellow	0.77	0.59	0.67
Waregu	0.71	0.5	0.59
Raja	0.82	0.75	0.78
Red	1	0.67	0.80
Fan	0.71	0.45	0.55
Bottle	1	0.5	0.67
Average	0.73	0.53	0.60

Table 5. Recapitulation of precision, recall, and accuracy values based on threshold 0.9.

Label	Precision	Recall	F1-score
Fishtail	1	0.83	0.90
Bamboo	1	0	0
Yellow	1	0	0
Waregu	1	0	0
Raja	1	0.25	0.4
Red	1	0.83	0.90
Fan	1	0	0
Bottle	1	0.83	0.90
Average	1	0.34	0.38

4. Conclusions

Digital image processing using AutoML Vision Edge can speed up the translation of palm plant characters. The important character to identify the species of palm plant on the UIN Syarif Hidayatullah Campus, Jakarta, based on the new testing data, there are as many as 80 images, and each class has ten images tested. The identification of palm plants in the testing data resulted in a new level of accuracy of 90%. Based on the threshold value of 0.5, the precision value is 81.25%, the recall is 53.51%, and the accuracy value is 64.59%. At the threshold value of 0.9, the precision value is 100%, the recall is 6.19%, and the accuracy value is 11.66%.

Acknowledgement

We thank the Center for Research and Publishing, UIN Syarif Hidayatullah Jakarta, for providing research funds in 2020.

References

[1] Spoerer, C. J., McClure, P., & Kriegeskorte, N. (2017). Recurrent convolutional neural networks: a better model of biological object recognition. *Frontiers in psychology*, vol. 8, no. 1551, pp. 1-14.

[2] Priyanti, & Irsyam, A. S. D. (2017). Suku Fabaceae di Kampus Universitas Islam Negeri Syarif Hidayatullah, Jakarta, Bagian 2: Tumbuhan polong berperawakan terna. *Al-Kauniyah*, vol. 10, no. 1, pp. 42-47.

[3] Herdiyeni, H., Adisantoso, J., Damayanti, K. E., Zuhud, E. A. M., Nurfadhila, E., & Paskianti, K. (2013). Pemanfaatan teknologi tepat guna identifikasi tumbuhan obat berbasis citra. *Jurnal Ilmu Pertanian Indonesia (JIPI)*, vol. 18, no. 2, pp. 85-91.

[4] Akmalia, T. S. (2014). Identifikasi tumbuhan Angiospermae dengan kunci determinasi berbasis Flash sebagai media belajar untuk siswa kelas X SMA/MA. [*Final Project*]. Yogyakarta: UIN Sunan Kalijaga Yogyakarta.

[5] Budiharto, W., & Suhartono, D. (2014). *Artificial Intelligence Konsep dan Penerapannya*. Yogyakarta: Andi.

[6] Pressman, R. S. (2015). *Software Engineering: A Practitioners Approach* (*8th Edition*). New York: McGraw-Hill Education.

[7] Amrullah, I. N., & Sutojo, T. (2014). Identifikasi daun berdasarkan faktor kekompakan dan faktor kebundaran bentuk daun. *Techno. Com*, vol.13, no. 4, pp. 198-205.

[8] Danukusumo, K. P. (2017). Implementasi deep learning menggunakan convolutional neural network untuk klasifikasi citra candi berbasis GPU [*Dissertation*]. Yogyakarta: Universitas Atma Jaya Yogyakarta.

[9] Britton, C., & Doake, J. (2001). *Object-Oriented System Development: A Gentle Introduction*. New York: McGraw-Hill Publishing.

[10] Priyanti. (2008). Tanaman peneduh di Kampus I UIN Syarif Hidayatullah Jakarta. *Al-Kauniyah*, vol. 2. no. 2, pp. 93-100.

[11] Priyanti. (2008). Tanaman monokotil di Kampus I dan II UIN Syarif Hidayatullah Jakarta. *Al-Kauniyah*, vol. 2, no. 1, pp. 29-36.

[12] Abdillah, F. (2016). Penggunaan deep learning untuk prediksi churn pada jaringan telekomunikasi mobile. *E-Proceeding of Engineering*, vol. 3, no. 2. pp. 3882-3885.

[13] Taufik, Priyanti, & Amrizal, V. (2010). Identifikasi tumbuhan kacang-kacangan (Papilionaceae) menggunakan sistem pakar dengan pendekatan forward chaining. [*Final Project*]. Jakarta: UIN Syarif Hidayatullah Jakarta.

[14] Kendall, K. E., & Kendall, J. E. (2010). *Analisis dan Perancangan Sistem*. Jakarta: PT Indeks Kelompok Gramedia.

[15] Zahro, H. Z. (2016). Analisis tekstur untuk identifikasi tumbuhan obat menggunakan klasifikasi Support Vector Machine. *Industri Inovatif: Jurnal Teknik Industri*, vol. 6, no. 2, pp. 33-40.