ABSTRACT

Patchouli has various varieties with almost the same physical characteristics. This often makes it difficult to recognize varieties with a high PA (Patchouli Alcohol) content. In this study an improvisation was introduced in the identification of patchouli varieties using leaf images using Extreme Learning Machine (ELM). However, problems occur in ELM if the data used is not balanced where the training process can not able to recognize data in the minority class well. Therefore, this study conducted a process to balance the composition of the data using the Synthetic Minority Over-sampling Technique (SMOTE) method. The test results of 93 data on the imbalanced composition with a comparison of 70% of training data and 30% of test data obtained an average accuracy of 93.57%. After implementing SMOTE in the Tetraploid, Patchoulina and Sidikalang classes where the amount of data in each class becomes 58 data, an average accuracy of 96.00% is achieved. This shows the existence of an increase in the process of identification with ELM when new data generation with SMOTE is carried out to balance the composition of the data.

CCS CONCEPTS
• Computing methodologies → Machine learning → Machine learning approaches → Neural networks

KEYWORDS

Patchouli varieties, leaf image, imbalanced data, ELM, SMOTE

ACM Reference format:


1 Introduction

Patchouli has a lot of varieties with content of Patchouli Alcohol and resistance to diseases that are also different. To get a quality patchouli harvest, of course, seeds must be selected from superior varieties in the cultivation process. A farmer or expert who has been dealing with patchouli, can recognize some of these varieties by observing the morphology and shape of patchouli leaves. However, as more and more varieties appear with similar characteristics, it becomes more difficult in the identification process. Therefore we need an effort that is able to identify patchouli varieties based on the physical characteristics of patchouli leaves with low cost and fast time.

Several studies on the use of leaf image for plant identification have been carried out. Among them is the identification of plants through leaf shapes by counting the number of leaf shape patterns [16, 21]. The others are leaf identification using a combination of Deep Belief Networks and Multi-feature [18], an android application for identification of plant species based on leaf images [37], identification of plant leaves based on leaf bones [35] and identification of plant species based on leaf texture [22]. Other studies related to the identification of plant leaves has been done
using three features, namely shape, color and texture features with identification accuracy was 94% [14]. It has also been carried out the identification of medicinal plants by combining the three features i.e. morphology, texture, and shape. The maximum accuracy reaching was 74.67% [12].

These studies mainly identify plants that have different species which each species usually has different leaf characteristics and shapes, so it is quite easy to distinguish plants based on these characteristics. However, if the plants have the same species with different varieties, they tend to have almost the same physical characteristics. However, if the plants have the same species with different varieties, they tend to have almost the same physical characteristics and making it difficult for the identification. For this reason, it is necessary to identify specific features in the identification of patchouli varieties. Based on physical characteristics have been observed on patchouli varieties, this study uses a combination of morphological, texture and shape features of leaves as input in the recognition process. The features used are somewhat different from previous studies where mostly use morphological features [2, 4, 11, 19] and shape features [17]. The use of morphological features was done by [32] resulting in a fairly high accuracy of 90.3%. The study by [13] showed that the use of shape features extracted with Convex Hulls and Fast Fourier Transform (FFT) on leaf veins resulted in accuracy reaching 97.19%. Considering these results, our paper uses a combination of morphological features and shape features extracted with Convex Hulls.

In addition, the observations showed that some patchouli varieties have different leaf texture. Therefore, this study added texture features extracted using the wavelet feature analysis method. Several studies have shown that the wavelet method is quite effective in obtaining texture features. It can yield good recommendations for the detection of copper deposits [1], can increase the effectiveness of the weed detection process [3] and can improve performance in the recognition process [20, 27].

The features that have been obtained are then used as input in the classification. This study implements the Extreme Learning Machine (ELM) classification algorithm. ELM is a method that has been proven to produce a high accuracy with a short training time when compared to other methods [7, 5, 8, 24]. However, ELM performance is not optimal if the data is in an imbalance class [32, 28]. Classification of imbalance class will tend to ignore classes that have a small number of samples so that it can adversely affect the performance of the classification algorithm [34] because features in minority classes will usually be difficult to identify [15]. To overcome this problem, a process is added to handle imbalanced data by generating new data for minority classes so that the amount is equal to the amount of data from the majority class. The method used to handle imbalanced data is Synthetic Minority Over-Sampling Technique (SMOTE). This method is quite widely used in various studies with optimal results [10, 23, 31, 36, 10].

Based on the description that has been explained, the purpose of this study is to identify patchouli varieties using leaf images and the ELM algorithm by handling imbalanced data. Specifically, the objectives of this study are: 1) find the optimal number of ELM hidden neuron; 2) analyze the performance of ELM on imbalanced class; 3) analyze the performance of ELM on balanced class after implementing SMOTE.

2 Method

In general, the stages of identification patchouli varieties using leaf images are shown in Figure 1. Input in the form of patchouli leaf image measuring 400x500 pixels consisting of four classes, namely Sidikalang, Diploid, Tetraploid and Patchoulina. The feature extraction process then is performed to obtain features that will be used as input from the classification process. The features extracted were 6 morphological features, 6 texture features extracted using wavelet feature analysis and 2 shape features extracted using convex hulls. Then the process is carried out to balance the data in each class by generating new data in the minority class using SMOTE.

The resulting features are then used as input vectors in the classification process with the ELM algorithm. The classification phase begins with the ELM training process to get the best weight with an optimal ELM parameter. The best weights obtained during the training process are then used as weights in the ELM testing process. Classification performance in the recognition process is calculated using accuracy to get the number of classifications that are recognized correctly for the entire data tested.

![Figure 1. Flow diagram of patchouli varieties identification](image)

2.1 Data

The data used were taken in several regions namely Kesamben, Brawijaya University (UB) and Trenggalek. Data taken is an image of Diploid patchouli leaves, Patchoulinia, Sidikalang and Tetraploid. The total data used was 91 with the number of the data in each class was 58 Diploid, 9 Tetraploid, 9 Patchoulinia and 15 Sidikalang. An example of patchouli leaf image is shown in Figure 2.
Leaf image is taken indoors using the iPhone 4S camera with specifications of 8 MP, f/2.4, 35mm, autofocus, LED flash. Leaves to be taken are placed on a white pedestal in an upright position a distance of 20-25 cm from the camera.

### 2.2 Features Extraction

This study uses morphological, texture and shape features of leaves as input in the recognition process. These three features were extracted using different techniques namely morphology to extract morphological feature, wavelet feature analysis to extract texture features and convex hulls to extract shape features.

#### 2.2.1 Morphological Feature Extraction

Morphological characteristics can be divided into two, namely the basic characteristics and derivative characteristics [14]. The basic characteristics include diameter (D), physiological length (Lp), physiological width (Wp), area (A), and perimeter (P).

This study uses 6 derivative features as input features of the classification process namely aspect ratio, form factor, rectangularity, narrow factor, perimeter ratio of diameter (PD) and perimeter ratio of physiological length & width (PLW). This characteristic is obtained from the calculation of the ratio between the basic characteristics of the leaf. The formula used to calculate derivative features is shown in equation 1 to equation 6.

\[
\text{Aspect Ratio} = \frac{\text{Lp}}{\text{Wp}}
\]

\[
\text{Form Factor} = \frac{4\pi A}{P^2}
\]

\[
\text{Rectangularity} = \frac{\text{LpWp}}{A}
\]

\[
\text{Narrow Factor} = \frac{D}{\text{Lp}}
\]

\[
PD = \frac{P}{D}
\]

\[
PLW = \frac{P}{(\text{Lp+Wp})}
\]

#### 2.2.2 Wavelet Feature Extraction

Wavelet transform convert a signal into a series of wavelets. In digital image processing, the signal is represented by a sequence of discrete signals so-called 2D discrete wavelet transform (DWT-2D) [33]. DWT-2D calculations are performed using low-pass and high-pass filters of pixel image values. The DWT-2D image can be decomposed up to n levels. At each level, high-pass filters produce detailed pixel formation of images while low-pass filters produce rough estimates of images [26]. This transformation can be calculated using equation (7).

\[
X_{WT}(i,j) = \sum_{t=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} x(t)\psi_{i,j}^t\]

\[
\text{Denotes:}
\]

\[
X_{WT} = \text{wavelet transform function}
\]

\[
\psi = \text{mother wavelet function}
\]

\[
x(t) = \text{reverse transformation}
\]

\[
t = \text{time and output of the function}
\]

\[
i,j = \text{pixel coordinates}
\]

While the reverse transformation of DWT-2D is shown in (8)

\[
X(t) = \sum_{t=-\infty}^{\infty} \sum_{s=-\infty}^{\infty} X_{t,s} \psi_t
\]

\[
\text{Denotes:}
\]

\[
X(t) = \text{reverse transformation}
\]

\[
\psi = \text{mother wavelet function}
\]

\[
t = \text{time and output of the function}
\]

The results of the filter using the low-pass and high-pass obtained 4 subsections, namely low pass filter images in horizontal and vertical sections (LL), high pass filter images in horizontal and vertical sections (HH), low pass filter images in the horizontal direction and high pass filter images in the vertical direction (LH) and low pass filter image in the vertical direction (HL).

The application of the wavelet transform has a variety of algorithms with different wavelet coefficients. This study uses the Daubechies wavelet transformation. Daubechies has the advantage of not having redundancy, the same amount of computation as ordinary wavelets, and symmetrical interactions so that it is easy to use for edges of images [29]. The wavelet feature uses first energy normalization (L1) and second energy normalization (L2) of high-frequency sub-bands (HH) and feature extraction is performed for each level of decomposition using equation (9) and equation (10).

\[
L1 = \sum_{HH} \frac{|HH|}{MN}
\]

\[
L2 = \sum_{HH} \frac{HH^2}{MN}
\]

\[
\text{Denotes:}
\]

\[
HH = \text{high sub band of wavelet}
\]

\[
MN = \text{width x height of the image}
\]

This study extract six wavelet features for each high-sub band i.e. L1(HH3), L1(HH2), L1(HH1), L2(HH3), L2(HH2) and L2(HH1).

#### 2.2.3 Convex Hull Frature Extraction

The Convex Hull feature is obtained from the ratio between the image area of the Convex Hull and the area of leaf image [17].
Two features extracted from Convex Hull are convexity and solidity. Convexity is formed from a convex set, which includes all the points connecting two points in the set. The convexity value is obtained from the ratio of the perve length of the convex hull surrounding the object to the object's perimeter length. It can be represented using equation (11).

\[
\text{Convexity} = \frac{\text{ConvexPerimeter}}{\text{ObjectPerimeter}} \tag{11}
\]

Solidity is comparison of the area of an object compared to its convex hull, by utilizing the pixels that make up the convex hull. Mathematically, solidity formulated in (12).

\[
\text{Solidity} = \frac{\text{ObjectArea}}{\text{ConvexArea}} \tag{12}
\]

2.3 Balanced Data using Synthetic Minority Over-Sampling Technique (SMOTE)

The SMOTE method increases the amount of minor class data to be equivalent to the major class by generating artificial or synthetics data [6]. The artificial data is made based on k-nearest neighbors. The SMOTE method is explained as follows [5].

Choose two samples, \(x_1\) and \(x_2\), from the given minority sample set randomly, where each sample has \(n\) attributes. For \(x_1\) and \(x_2\), calculate the difference on the \(i\)-th attribute; that is, \(diff_i = x_{2i} - x_{1i}\). Then obtain the \(i\)-th attribute value of the new target sample according to (13).

\[
x_{12i} = \text{rand} \[0, 1\] * diff_i \tag{13}
\]

where \(\text{rand}[0, 1]\) means a random number between 0 and 1. So the final synthetic sample of \(x_1\) and \(x_2\) is calculated using (14).

\[
x_{12} = \text{rand} \[0, 1\] * diff \tag{14}
\]

This study generates synthetics data for Patchoulina class of 49 data, Tetraploid class of 49 data and Sidikalang class of 43 data. Thus, the total amount of balanced data is 232.

2.4 Classify Dataset using Extreme Learning Machine (ELM)

ELM have one or more neuron in hidden layers that work in one iteration [30]. Therefore, ELM is faster than other Neural Network methods that use the concept of Backpropogation learning [9].

This study uses ELM architecture introduced by Huang et al [13]. The method has a structure consisting of 3 layers, namely the input layer, hidden layer, and output layer. Each node in each layer is interconnected with weights. The steps of ELM Algorithm can be explained as follows [25].

Suppose there are \(N\) data samples \((x_i, t_i)\), where \(x_i = [x_{1i}, \ldots, x_{im}] \in \mathbb{R}^m\) and \(t_i = [t_{1i}, \ldots, t_{iq}] \in \mathbb{R}^q\). A single layer feedforward neural network with \(N\) hidden nodes and activation function \(g(x)\) is shown as (15).

\[
\tilde{\beta}_{i}g_i(x_j) = \sum_{i=1}^{N} \beta_{i} g(w_{i}, x_j + b_i) = 0, \ j = 1, \ldots, N \tag{15}
\]

In this equation, \(w_i = [w_{i1}, \ldots, w_{im}]^T\) is the weight vector of the connectors from input nodes to the \(i\)-th hidden node, and \(\beta_{i} = [\beta_{i1}, \ldots, \beta_{iq}]^T\) is the weight vector of the connectors between the \(i\)-th hidden node and the output nodes. The variable \(b_i\) is the threshold of the \(i\)-th hidden node. By approximating the samples with zero error as in (16).

\[
\sum_{j=1}^{N} \|0_i - t_i\| = 0 \tag{16}
\]

means that there exist \(w_i, \beta_{i}\), and \(b_i\) such that (17),

\[
\sum_{i=1}^{N} \beta_{i} g(w_{i}, x_j + b_i) = t_j, \ j = 1, \ldots, N \tag{17}
\]

By using the following substitutions (18):

\[
H(w_1, \ldots, w_N, b_1, \ldots, b_N, x_1, \ldots, x_N) = \begin{bmatrix}
g(w_{1}, x_1 + b_1) & \cdots & g(w_{N}, x_1 + b_N) \\
\vdots & \ddots & \vdots \\
g(w_{1}, x_N + b_1) & \cdots & g(w_{N}, x_N + b_N)
\end{bmatrix} \tag{18}
\]

Where \(\beta = [\beta_{1}^T, \ldots, \beta_{N}^T]\) and \(T = [t_1^T, \ldots, t_N^T]\), all \(N\) can be written as (19).

\[
HB = T \tag{19}
\]

3 Result and Discussion

This study conduct testing the use of balanced and imbalanced data to determine the effect of SMOTE on the classification results of the ELM. Testing for each data is done to find out the best neuron number of ELM. The number of hidden neurons tested is multiple of 5 ranging from 5 to 50. Each hidden neuron testing is done 10 times with a comparison of 70% training data and 30% testing data. The separation of training and test data is carried out randomly by paying attention to variations in data for each class.

3.1 Result of Testing on Imbalanced Class

The test results for the ELM Hidden neuron parameters on imbalanced data are shown in Table 1.

<table>
<thead>
<tr>
<th>Hidden Neuron</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>85.71</td>
<td>96.43</td>
<td>93.57</td>
</tr>
<tr>
<td>10</td>
<td>88.21</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>15</td>
<td>88.21</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>20</td>
<td>89.29</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>25</td>
<td>89.29</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>30</td>
<td>88.21</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>35</td>
<td>88.21</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>40</td>
<td>89.29</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>45</td>
<td>89.29</td>
<td>96.43</td>
<td>96.43</td>
</tr>
<tr>
<td>50</td>
<td>89.29</td>
<td>96.43</td>
<td>96.43</td>
</tr>
</tbody>
</table>
Based on 10 times tests conducted on each hidden neuron, there are variations in the accuracy obtained between the minimum and maximum value with the highest accuracy of 93.57% obtained at hidden neurons 20. If we look at Table 1, the minimum and maximum accuracy ranges are quite high, between 7.14 to 25.0. Therefore, this study uses average accuracy, where the value is between 78.57% to 93.57%. The test results also show an increase in the average accuracy starting from the number of neurons 5 to 20. However, after that the accuracy value tends to decrease with the increase in the number of hidden neurons.

It can be concluded that the number of hidden neurons is very influential on the accuracy generated, especially on the stability of accuracy in each experiment. The more stable and the higher accuracy produced, the parameter number of hidden neurons will be increasingly considered in subsequent testing.

Although the resulting accuracy value is quite high, however, there are still errors in recognition. This prediction error can be determined using a Confusion matrix. An example here is the Confusion matrix for the tests with the number of hidden neurons 20 as shown in Figure 3. The code in the table shows the patchouli class, namely Diploid (D), Patchoulinea (P), Sidikalang (S) and Tetraploid (T). From the confusion matrix, it can be seen that of the four classes, there are 2 data that are recognized incorrectly, namely Patchoulinea recognized by Sidikalang and Tetraploid recognized by Diploid. Although the accuracy and precision obtained were quite high, namely 92.86% for accuracy and 94.44% for precision, but the recall obtained was much lower, which is 81.25%. This occurs because the number of test data is too few in classes P, S and T so that if there is a missing value even though only 1 data (target P is recognized by S), it causes the recall value to be low.

<table>
<thead>
<tr>
<th>Prediction</th>
<th>D</th>
<th>P</th>
<th>S</th>
<th>T</th>
</tr>
</thead>
<tbody>
<tr>
<td>True</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>P</td>
<td>0</td>
<td>3</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>S</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>T</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
</tbody>
</table>

Figure 3: Confusion matrix of imbalanced data at hidden neuron 20

### 3.2 Result of Testing on Balanced Class

The process with SMOTE is done to generate artificial data in the Tetraploid, Patchoulinea and Sidikalang classes so that each class has 58 data. The results of testing the number of ELM hidden neurons in balanced data are shown in Table 2.

<table>
<thead>
<tr>
<th>Hidden Neuron</th>
<th>Minimum</th>
<th>Maximum</th>
<th>Average</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>74.29</td>
<td>90.00</td>
<td>81.71</td>
</tr>
<tr>
<td>10</td>
<td>87.14</td>
<td>98.57</td>
<td>93.00</td>
</tr>
<tr>
<td>15</td>
<td>85.71</td>
<td>98.57</td>
<td>93.43</td>
</tr>
<tr>
<td>20</td>
<td>85.71</td>
<td>98.57</td>
<td>94.71</td>
</tr>
<tr>
<td>25</td>
<td>91.43</td>
<td>98.57</td>
<td>95.86</td>
</tr>
<tr>
<td>30</td>
<td>87.14</td>
<td>98.57</td>
<td>96.00</td>
</tr>
<tr>
<td>35</td>
<td>92.86</td>
<td>98.57</td>
<td>96.00</td>
</tr>
<tr>
<td>40</td>
<td>90.00</td>
<td>98.57</td>
<td>95.71</td>
</tr>
<tr>
<td>45</td>
<td>90.00</td>
<td>98.57</td>
<td>95.71</td>
</tr>
<tr>
<td>50</td>
<td>91.43</td>
<td>98.57</td>
<td>95.14</td>
</tr>
</tbody>
</table>

As with imbalanced data, testing with balanced data also shows that there are prediction errors and can be observed using a Confusion matrix. The example shown in this discussion is the Confusion matrix for the test with 35 hidden neurons (Figure 4). The confusion matrix shows that of the four existing classes, there is only 1 data that is recognized incorrectly, namely Tetraploid (T) recognized by Diploid (D). The accuracy, precision and recall values are also very high, namely 98.57% for accuracy, 98.75% for precision and 98.21 for recall. This happens because of the increase in the number of test data in classes P, S and T so that if only 1 data is identified as incorrect (target T is recognized by D), it will not have much effect on reducing the value of precision and recall.
3.3 Comparison of Test Result on Balanced and Imbalanced Class

The comparison of the average accuracy of 10 tests on the best number of neurons in the balanced and imbalanced data is shown in Figure 5. From the figure, it is known that the accuracy value on the balanced data test is above the accuracy value on the imbalanced data. The exception occurs in the third test where the accuracy value on the balanced data is lower because at this time the accuracy of the imbalanced data is at the highest value. If we pay attention to the accuracy value on 10 tests with balanced data is also very high, which is always above 92%. Furthermore, by referring to the average accuracy value as shown in Table 1 and Table 2, there is an increase in accuracy by 2.6%. Although this increase is not very significant, the stability to get an accuracy above 92% is quite high. These can occur because the more training data, the variation of data recognized in the training process becomes more diverse. Thus, the ability to recognize test data will be better. Improvements in the recognition process mainly occurred in classes other than Diploid, where after duplicating data using SMOTE obtained data with feature values that are evenly distributed between the lower limit to the upper limit. This will certainly be very influential if the initial feature value before being repaired with SMOTE has a fairly high range.

![Figure 5: Comparison of imbalanced and balanced accuracy for 10 iteration](image)

4 Conclusion

In this research, an improvement method is used to improve the performance of ELM in the recognition process if the training data is not large and has an imbalanced composition. The step taken by creating the artificial data of classes that have a small number of data using SMOTE. The test results of the imbalanced class obtained the best average accuracy is 93.57% at number of neuron 20, while the balanced data obtained the best average accuracy is 96.00% at number of neuron 30 and 35. This result shows an increase in accuracy although it is not significant. But, the stability and probability of the test giving a correct response is quite high. Therefore the methods offered can describe solutions in handling imbalanced class in the identification process using ELM.

However, the variation of data for each class in this study is still low, as is the number of varieties recognized. Thus, for further research, it needs to be re-tested by adding variations in the data of each class and adding recognizable varieties. Besides that, it is also necessary to explore feature that are reliable and resistant to the uses of various mobile camera in image capture.

ACKNOWLEDGMENTS

We would like to thank to Faculty of Computer Science, University of Brawijaya for the funding of this research.

REFERENCES

Improve Performance of Extreme Learning Machine in Classification of Patchouli Varieties with Imbalanced Class


