Akmal Junaidi

Andrian2019-KNN_patrec_batik.pdf

Sources Overview

66%

OVERALL SIMILARITY

1	R Andrian, M A Naufal, B Hermanto, A Junaidi, F R Lumbanraja. "k-Nearest Neighbor (k-NN) Classification for Recognition of the Batik ... CROSSREF	65%
2	Sriwijaya University on 2019-10-26 SUBMITTED WORKS	1%

Excluded search repositories:

None

Excluded from document:

Bibliography

Quotes

Citations

Small Matches (less than 10 words)

Excluded sources:

None

IOP Conf. Series: Journal of Physics: Conf. Series **1338** (2019) 012061 doi:10.1088/1742-6596/1338/1/012061

k-Nearest Neighbor (k-NN) Classification for Recognition of the Batik Lampung Motifs

R Andrian[1,a], M A Naufal[1,b], B Hermanto[1,c], A Junaidi[1,d] and F R Lumbanraja[1,e]

[1]Departmen of Computer Science, Lampung University, Bandar Lampung, Lampung, Indonesia

[a]rico.andrian@fmipa.unila.ac.id; [b]adibnaufal29@gmail.com; [c]bambang.hermanto@fmipa.unila.ac.id; [d]akmal.junaidi@fmipa.unila.ac.id; [e]favorisen.lumbanraja@fmipa.unila.ac.id

Abstract. Batik is a famous name of a traditional fabric from Java. It has been admitted as one if the traditional cultural heritage of Indonesia by UNESCO since October 2nd , 2009. Over the time, Batik is copied and modified by many regions in Indonesia resulting some new unique motifs. Batik Lampung is an sample of them. This paper deals with the k-Nearest Neighbor classification of the motifs (pattern) of the Batik Lampung. The known motifs of Batik Lampung consist of *Jung Agung, Siger Kembang Cengkih, Siger Ratu Agung,* and *Sembagi*. The original image samples are stored in RGB. They are firstly resized into 50 x 50 pixels and then converted to grayscale image. To recognize them, the Gray Level Co-Occurence Matrix (GLCM) feature is extracted and k-Nearest Neighbor (k-NN) with values of k = 3, 5, 7, 9, 11 and orientation angle of 0^0 45^0, 90^0, 135^0 is applied to classify the motifs. The best accuracy is achieved at the rate 97,96% for k = 7 and angle135^0.

1. Introduction

Batik is Indonesia's cultural heritage and has been designated by UNESCO on October 2nd, 2009 as the rights of intellectual culture of Indonesia [1]. Batik has now expanded to the entire archipelago, causing many motifs created in accordance with their respective regional cultures. Various motifs encourage humans to create a machine that can recognize a variety of motives, one of the motifs to be used in this research is Lampung's Batik.

Computer technology has also grown rapidly, one of the computer technology is pattern recognition. Patter recognition is a discipline to classify or describe something based on quantitative measurement features (characteristics) or the main properties of an object [2]. Pattern Recognition has three important stages: first *preprocessing*. This stage is an early stage before the process of classification that *resize* or change the pixel size becomes smaller, and *grayscaling* [3]. The second stage is *feature extraction* or feature extraction which aims to get more clear information about the data in batik image [3]. The third stage is the classification, in this study using *K-Nearest Neighbor* method, which is a method to classify the object based on the data that is most similar (nearest neighbor) with a predetermined number *k* and classifies them into new classes [4]. This study will focus on the ability of a system to classify Lampung Batik motif using *K-Nearest Neighbor*method.

IOP Conf. Series: Journal of Physics: Conf. Series **1338** (2019) 012061 doi:10.1088/1742-6596/1338/1/012061

2. Data and Methods

2.1 Batik
Batik is one of Indonesia's cultural heritages that have been famous all over the world. Batik has been designated by UNESCO on October 2, 2009 as the rights of intellectual culture of Indonesia [1]. The word batik is derived from the combination of two words, namely Java language *amba* and *titik*. *Amba* has the meaning of the fabric and *titik* is how to give a motive on fabric by dots [5].

Lampung's Batik : Lampung province at first did not have a tradition of batik, but there are relics of the so-called first batik worn by people of Lampung, namely as Sembagi fabric [6].

2.2 Pattern Recognition
Pattern recognition is a science in artificial intelligence that aims to classify something based on its measurements [2].

2.3 Classification of K-Nearest Neighbor Method
K-NN classification algorithm predicts the category of test samples according to training samples *k* which is the closest neighbor to the test sample, and inserting it into a category which has the greatest probability category [7]. Near or far distances to neighboring point can be calculated by using *the Euclidean distance*. *Euclidean distance is* represented as follows [8]:

$$D(a,b) = \sqrt{\sum_{k=1}^{d} (a_k - b_k)^2} \tag{1}$$

Description:
D = Distance between points
a = The known point
b = The unknown point
d = Dimention of the point being measured
k = value of neighboring data measured

2.4 Calculation of Accuracy Level
This process is performed as one of the evaluation benchmarks in a system. Measuring the level of accuracy can be used in various ways one of which is using in this reasearch is the *Detction Rate*. *Detection Rate* equation is represented as follows [9]:

$$Detection\ Rate = \frac{TP}{TP + TN} \times 100\% \tag{2}$$

Description:
TP (True Positive) = Amount of correctly identified class
TN (True Negative) = Amount of incorrectly identied class

3. Metodology
This study was conducted based on the structure of that pattern recognition system [2]:

3.1. Collecting Batik's Image
Image of Lampung's Batik of motif obtained from the manufacture of batik cloth of motifs Lampung. Data taken using Canon EOS 600D SLR Camera with Lampung batik motifs taken are *Siger Kembang Cengkih, Jung Agung, Sembagi* and *Siger Ratu Agung*. Image taken for each motif

IOP Conf. Series: Journal of Physics: Conf. Series **1338** (2019) 012061 doi:10.1088/1742-6596/1338/1/012061

is as much as 25 image formats JPEG *(Joint Photographic Experts Group)* with a total of 100 images batik image Lampung.

3.2. Preprocessing
Preprocessing is an initial phase in which the pattern recognition object or image to be processed before proceeding to the next stage. The first step is to cut *(crop)* object or image motif Lampung, then *Grayscaling,* and the last step is changed the image size to 50 x 50 pixels.

3.3. Feature Extraction
Feature Extraction is one of the part of a pattern recognition technique that aims to extract or extract the unique values of an object that distinguishes it from other objects. This study uses *Gray Level Co-occurrence Matrix* (GLCM) is a matrix that describes the frequency with which a pair of two pixels with a certain intensity in the distance d and orientation with a certain angle θ in the image [9]. The features measured are *Angular Second Moment, Contrast, homogeneity, correlation,* with the direction of orientation angle of 0^0 $45^{0,}$ $90^{0,}$ 135^0 and $d = 1$.

3.4. Separate Training Data and Test Data
The amount of training data and the test data used in this study is to perbandigan 70:30 on the entire dataset, which is 70% o ftraining data and 30% of test data.

3.5. Classification Using K-Nearest Neighbor K-NN Method
The last stage in pattern recognition is the classification. *K-Nearest Neighbor* (K-NN) is one method of *supervised learning* algorithm is also called as a category classification algorithms known *output*. KNN works by classifying an object that has the closest resemblance to any other object. K-NN has an attribute that initialize as k, ie the number of neighboring values that made reference to the K-NN classification, the total value of k is a positive integer, small and odd numbered.

4. Discussion

4.1 Data Collection
The data used in this research is Lampung Batik image data and Java batik, Javanese Batik will be used as class "Not Batik Lampung". Lampung batik used is Jung Agung motif, Siger Kembang Cengkih, Sembagi, and Siger Ratu Agung obtained from the collection of Siger Roemah Batik, while for Batik Java is Parang Kusumo and Parang Rusak motif obtained from private collection. Batik image that has been collected totaled 25 for each motif, batik used the next image in the *folder* called 'datasets'.

4.2 Preprocessing
The next stage is preprocessing, while that which is done is *cropping* or cutting batik image according to the characteristic of batik motif. The next process is *scaling* or resize the image to a size smaller. Image resized to 50 x 50 *pixels,* and then change the image becomes grayish.

4.3 Feature Extraction of Batik's Image
Sample results GLCM feature extraction method with a distance of a couple of pixels or d = 1 and the orientation angle $\theta = 0^0$ shown in table 1.

IOP Conf. Series: Journal of Physics: Conf. Series **1338** (2019) 012061 doi:10.1088/1742-6596/1338/1/012061

Tabel 1. Sample Value of Feature Extraction

Contrast	Homogeneity	Correlation	ASM	Class Name
0.3867085427135678	0.84942964824120615	0.86875083105218043	0.21444042259033863	BatikLampungJungAgung
0.38635678391959805	0.84701256281407034	0.84985116272872452	0.15387912016110705	BatikLampungJungAgung
0.36268844221105534	0.86626381909547745	0.86955395213163522	0.28819510586853858	BatikLampungJungAgung
2.0440954773869349	0.67609132767911961	0.75741657145423225	0.10157448580843917	BatikLampungSembagi
1.9264070351758793	0.67987740739898583	0.76840908185104939	0.098129503674149604	BatikLampungSembagi
1.9009296482412061	0.67905672025284791	0.77422130837589975	0.095256777290977493	BatikLampungSembagi

The table 1 shown above feature extraction results represent the value between features in which they have different values to be comparable values among other features. Results of subsequent feature extraction is saved into Microsoft Excel with the file extension CSV (Comma Separated Value).

4.4 Classification Method Using K-Nearest Neighbor

Basic steps in the method of classification K-NN there are 2 that is, determine the nearest neighbor then determine the class based on the nearest neighbor. The nearest neighbor in the K-NN method denoted by k. Determining the value of k will be very influential in the classification process will be. The value of k is usually an odd number in order to see the point of separation between the classes to be determined. A small k value will lead to results that are unstable, while the value of k is too high will cause bias or boundaries between each classification become blurred. K value used in this study was 3, 5, 7, 9 and 11. The experiments were carried out four (4) times based on the orientation of the direction of the corner on the feature extraction Gray Level Co-Occurrence Matrix (GLCM) is 0^0, 45^0 90^0 and 135^0. The class used in this study amounted to 6 and is shown in table 2.

Table 2. Batik Motif used in Research

No.	Name of Batik
1	Batik Lampung Jung Agung
2	Batik Lampung Siger Kembang Cengkih
3	Batik Lampung Sembagi
4	Batik Lampung Siger Ratu Agung
5	Batik Jawa Parang Kusumo (Bukan Batik Lampung)
6	Batik Jawa Parang Rusak (Bukan Batik Lampung)

4.5 Testing and Analysis

The first test was conducted by using a dataset with the feature extraction results orientation angle and neighbors by 0^0, a second trial with the direction of an angle of 45^0, the third attempt by the orientation angle of 90^0 and a fourth experiment with the orientation angle of 135^0. Orientation towards the neighbors on four experiments above are based on feature extraction methods Gray Level Co-occurence Matrix (GLCM). The test results are shown in table 3 of the level of accuracy or recognition rate.

IOP Conf. Series: Journal of Physics: Conf. Series **1338** (2019) 012061 doi:10.1088/1742-6596/1338/1/012061

Table 3. Results of Tests with the orientation angle 0^0, 45^0, 90^0, 135^0

Value of *k*	Accuracy of each angle orientation			
	0^0	45^0	90^0	135^0
3	92,31%	84,62%	95,74%	97,50%
5	91,11%	90,00%	95,65%	97,67%
7	89,47%	89,74%	97,73%	97,96%
9	93,62%	80,49%	87,23%	95,00%
11	88,64%	79,59%	97,56%	93,48%

Table 3 shows the level of accuracy of each test, the table above obtained the highest accuracy in the orientation angle of 135^0 with a value of k = 7 amounted to 97.96%. Lowest accuracy rate obtained in the orientation angle of 45^0 with the number of a neighbor or the value of k = 9 is equal to 80.49%.

Tests that have been performed is obtained in the form of varying levels of accuracy for each value of k and orientation angle 0^0, 45^0, 90^0 and 135^0. The first experiments on the orientation angle 0^0, the accuracy decreases in the value of k = 5 and go up in value k = 9, then back down at k = 11. The second trial on the orientation angle of 45^0, appears to rise in the value of k = 5 and decreases in the value of k = 7 to 11. The third experiment on the orientation angle of 90^0, no pattern changes, the accuracy decreases in the value of k = 5 and 9, and then rise in the value of k = 7 and 11. The fourth experiment on orientation angle of 135^0, appears to rise in the value of k = 5 and 7 and go down in value k = 9 and 11.

Changes in the level of accuracy is unstable and tends to occur in the whole orientation toward the corner, but the accuracy rate changes have little irregularity occurs on the orientation angle of 0^0 and 90^0. This irregularity is caused by test data and training data performed in each fluctuating test, this is done to get the appropriate accuracy level and all the datasets in the research are used in each experiment, so the distribution of testing and data training is evenly distributed.

5. Conclusion

The conclusions obtained based on the research that has been done are as follows:

1. K-Nearest Neighbor classification method has been successfully implemented in the process of pattern recognition of Lampung's Batik.
2. K-Nearest Neighbor classification method has been very good in the process of pattern recognition of Lampung's Batik.
3. The highest accuracy is obtained on testing at the orientation angle of 135^0 at k = 7 amounted to 97.96%. Lowest level of accuracy in testing at the orientation angle of 45^0 with a value of k = 9 is equal to 80.49%.

Reference

[1] Arisandi B Suciati N and Wijaya A Y 2011 Pengenalan Motif Batik dengan Rotated Wavelet Filter dan Neural Network *JUTI* **9** 2 pp 13-19
[2] Putra D 2010 *Pengolahan Citra Digital* (Yogyakarta: CV. Andi Offset)
[3] Yodha J W and Kurniawan A W 2014 Pengenalan Motif Batik Menggunakan Deteksi Tepi Canny dan K-Nearest Neighbor Techno.COM **13** 4 pp 251-262
[4] Harrington P 2012 *Machine Learning in Action* (New York: Manning Publications Co.)
[5] Utami R 2014 Ensiklopedia Batik dan Kain Hias Nusantara (Bandung: CV. Angkasa)
[6] Rudiansyah. 2016. *Sebagi, Riwayat Batik Lampung*. Lampung Post, 18 September 2016.
[7] Suguna N and Thanushkodi K 2010 An Improved k-Nearest Neighbor Classification Using Genetic Algorithm *International Journal of Computer Science* **7** 4 pp 18-21
[8] Pramesti R P A 2013 Identifikasi Karakter Plat Nomor Kendaraan Menggunakan Ekstraksi Fitur ICZ dan ZCZ dengan Metode Klasifikasi K-NN *Skripsi* (Bogor: Institut Pertanian Bogor)

IOP Conf. Series: Journal of Physics: Conf. Series **1338** (2019) 012061 doi:10.1088/1742-6596/1338/1/012061

[9] Norsyafawati F Norwawi N M Sabri M Seman K 2011 Identifying False Alarm Rates for Intrusion Detection System with Data Mining *International Journal of Computer Science and Network Security* **11** 4 pp 95-99