**Scientific Journal of Informatics**
Vol. 5, No. 1, May 2018
p-ISSN 2407-7658 http://journal.unnes.ac.id/nju/index.php/sji e-ISSN 2460-0040

# A High Performace of Local Binary Pattern on Classify Javanese Character Classification

**Ajib Susanto[1], Daurat Sinaga[2], Christy Atika Sari[3], Eko Hari Rachmawanto[4], De Rosal Ignatius Moses Setiadi[5]**
[1,2,3,4,5]Computer Science Faculty, Universitas Dian Nuswantoro Semarang, Indonessa
Email: [1]ajib.susanto@dsn.dinus.ac.id, [2]dauratsinaga@dsn.dinus.ac.id, [3]atika.sari@dsn.dinus.ac.id, [4]eko.hari@dsn.dinus.ac.id, [5]moses@dsn.dinus.ac.id

**Abstract**

The classification of Javanese character images is done with the aim of recognizing each character. The selected classification algorithm is K-Nearest Neighbor (KNN) at K = 1, 3, 5, 7, and 9. To improve KNN performance in Javanese character written by the author, and to prove that feature extraction is needed in the process image classification of Javanese character. In this study selected Local Binary Patter (LBP) as a feature extraction because there are research objects with a certain level of slope. The LBP parameters are used between [16 16], [32 32], [64 64], [128 128], and [256 256]. Experiments were performed on 80 training drawings and 40 test images. KNN values after combination with LBP characteristic extraction were 82.5% at K = 3 and LBP parameters [64 64].

**Keywords**: Optical Character Recognition, Javanese Character, Local Binary Pattern

## 1. INTRODUCTION

The process of image recognition can be done by classification process either in real time or not. The usual image classification is done by image processing on unsupervised learning and supervised learning techniques [1]. In supervised learning techniques, some commonly used algorithms are Neural Network (NN) [2], K-Nearest Neighbor (KNN) [3], Support Vector Machine (SVM) [4], Decision Tree, and others. The use of supervised learning algorithms is usually used when researchers already have data train and clear variables to classify data. While the object of research in the form of images that are often used that can be either static images or dynamic images. Static images are images obtained from other media such as books making it easier to do the detection. For dynamic images, the images are drawn or written by humans on a piece of paper for example and the shape of the image is not always the same as the printed one in the book.

In this paper will be used static image model as the object of research in the form of image character of Javanese script. The character of Javanese script is in addition to the ancestral heritage, is also a means of learning for the community to get to know the culture. Javanese writing is known to be difficult to learn, because it has a variety of fractional characters, including numeric characters, mudra, sandangan and others. In this paper, classification of Javanese character using K-Nearest Neighbor (KNN) supervised learning algorithm will be conducted. KNN itself is an algorithm with a fairly high accuracy. However, to improve the accuracy of the static image form used, we do a preprocessing of Local Binary Pattern (LBP) calculations to detect some of the images written in more slanted form than the original. From several studies that

have been done by similar research in Javanese character classification, no one has used preprocessing in the form of LBP algorithm. We hope this paper will contribute to similar research. Whereas another paper only proposed segmentation pattern [5] or Backpropagation Neural Network (BNN) [6].

## 2. METHODS

### 2.1. Local Binary Pattern (LBP)

Local Binary Pattern (LBP) is one method of texture analysis [7]. This method can be used in both classification model and data clustering on supervised learning and unsupervised learning. LBP deals with the various fields in the preprocessing process of digital image processing, especially on the type of feature extraction. Figure 1 is a representation of fields related to the LBP.

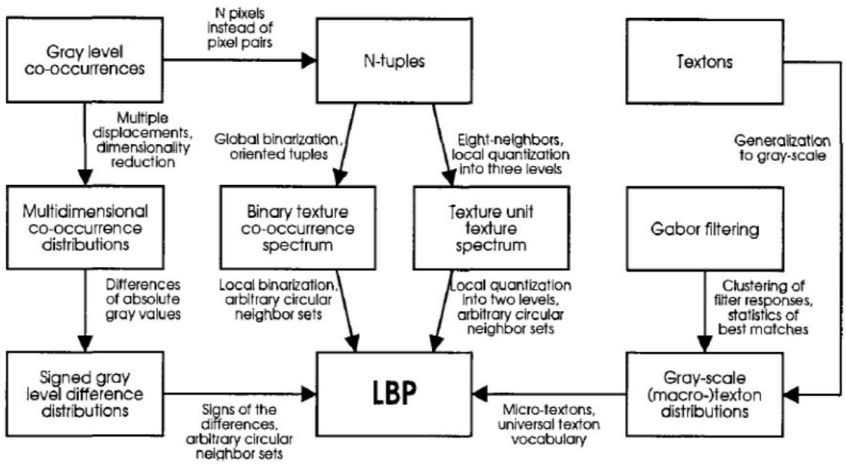



Figure 1. LBP in the field of texture analysis operators [8]

According to Cheng in the book of pattern recognition, the LBP operator can be seen as a truly unifying approach to the traditionally divergent statistical and structural models of texture analysis. Texture is described in terms of micro-primitives (textons) and their statistical placement rules. Optionally, the primitives may be coupled with a complementary measure of local image contrast, which measures the strength of the primitives. The LBP operator is relatively invariant with respect to changes in illumination and image rotation. It can even resist changes in texture scale. Still more work needs to be done to make the operator invariant against 3-D distortions, often present in natural scenes, for example. The promising results achieved in view based recognition and in detecting moving objects indicate that this problem can be solved with the LBP. Furthermore, the success in detecting and recognizing faces suggests that the LBP may be useful in many object recognition tasks that have not previously been considered texture analysis problems. Perhaps the most important property of the LBP operator in real-world applications is its tolerance against illumination changes. Equally important is its computational simplicity, which makes it possible to analyze images in challenging real time settings.

### 2.2. K-Nearest Neighbor (KNN)

K-Nearest Neighbor (K-NN) is one of algorithm which has capability to classify object especially using images [9]. KNN as machine learning and supervised learning, very popular because has simple operation and easy to learn [10]. KNN has a high accuracy for classification and regression. The basic of KNN is calculate the distance between new testing data and training data, then it will classify through class. One of popular distance calculate using Euclidean Distance as follow in Equation 1.

Here, an equation of Euclidean Distance [11] in KNN algorithm:

$$d\,(x\,,\,y) = \sqrt{\sum_{i=1}^{n}(xi - yi)^2} \quad (1)$$

where,

d = distance
x = training data
y = testing data
i = total of neighbor
n = total pixels of image

KNN stages as follow [12]:

1) Define total of K, which used as reference from nearest class.
2) Calculate the distance between training data and testing data.
3) Sorting and define the distance of nearest neighbor based on K value as reference distance.
4) Check output or labels for every nearest class.
5) Classify testing image into majority of nearest class.

## 3. RESULTS AND DISCUSSION

Experiment was done using character Javanese character as image with size 512x512 pixels. Here, javanese character taken from manual writing by author, then we take images. We did not use character image form another source because we must do some investigate about an accuracy. In this paper, we use 160 dataset that separate into 120 data training and 40 data testing as follow in Figure 2 and Figure 3.

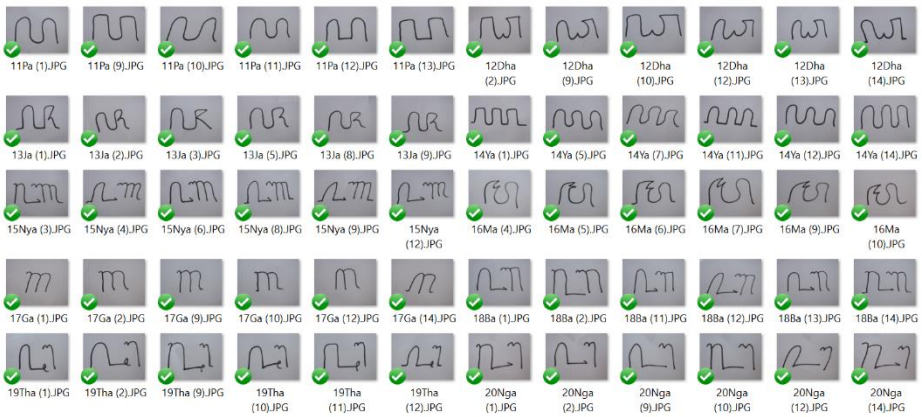



Figure 2. Training Data Size in 512x512 Pixels


Figure 3. Testing Data Size in 512x512 Pixels

Due to enhance an accuracy, we use Local Binary Pattern (LBP) as feature extraction in Javanese character. Images process by preprocessing using supervise learning, and choose K-Nearest Neighbor. The stages of experiment done by Figure 4.

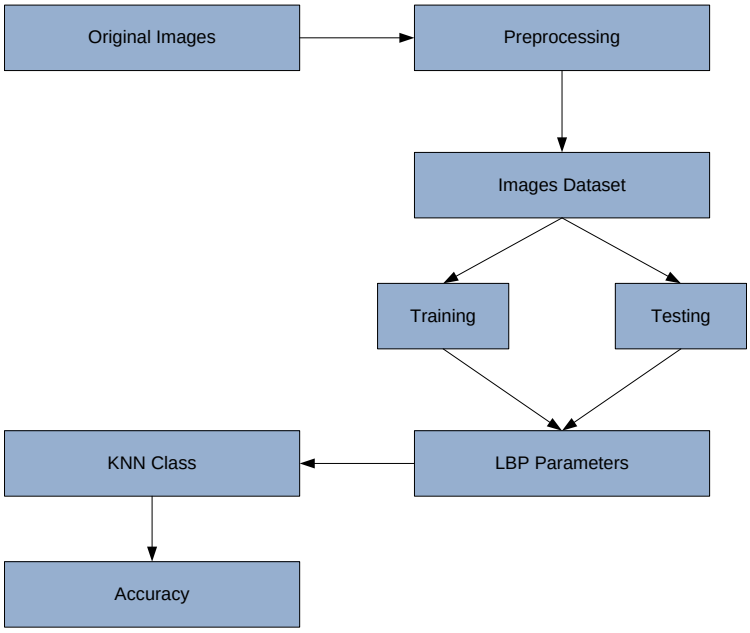


Figure 4. Stages of Classification Javanese Character using LBP-KNN

Due to achieve a good accuracy, there are several steps as follow:

1) Write every Javanese character into paper.
2) The result is manually photographed, then in the crop with matlab, originally 640 x480 into 512x512 size using the function `imresize`.
3) The image that has been cropped then calculated `cellvalue` value using LBP parameters. The parameters we use in the identification of this Javanese script are [16 16], [32 32], [64 64], [128 128] and [256 256]. With this parameter the image will be converted to grayscale and in accordance with the LBP parameters used. Value LBP can be seen in the pseudocode we have made following.

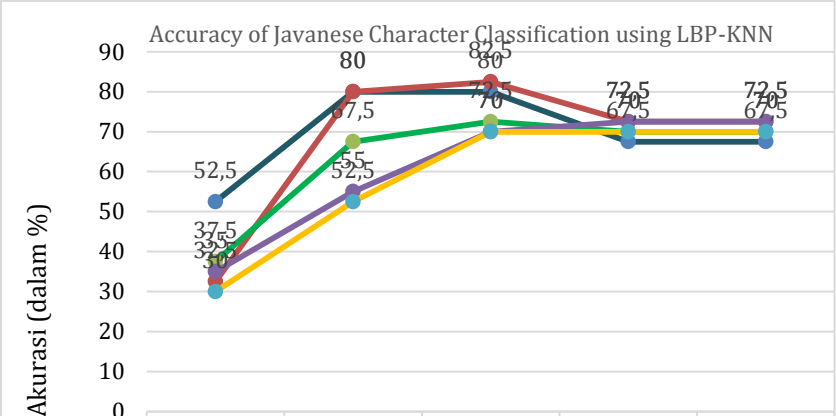
```
menu = get(handles.menu_LBP,'Value');
if menu==1
    ukuran_sel = [16 16];
else if menu==2
        ukuran_sel = [32 32];
    else if menu==3
            ukuran_sel = [64 64];
        else
            ukuran_sel = [128 128];
        end
    end
end
```

4) After going through the extraction feature, the image is processed by KNN for the classification process with the value of K = 1,3,5,7,9 as in the following pseudocode.

```
k = get(handles.menu_k,'Value');
if k==1
    k = 1;
else if k==2
        k = 3;
    else if k==3
            k = 5;
        else if k==4
                k = 7;
            else
                k = 9;
            end
        end
    end
end
```

5) The javanese character matching process will be performed, and the LBP-KNN performance assessment is calculated with accuracy.

The data matching results with the value of K = 1 and the LPB parameters on [16 16] which are dried up to K = 9 and the LBP [128 128] parameter can be seen in Figure 5.




| Parameter LBP | [16 16] | [32 32] | [64 64] | [128 128] | [256 256] |
|---|---|---|---|---|---|
| K=1 | 52,5 | 80 | 80 | 67,5 | 67,5 |
| K=3 | 32,5 | 80 | 82,5 | 72,5 | 72,5 |
| K=5 | 37,5 | 67,5 | 72,5 | 70 | 70 |
| K=7 | 35 | 55 | 70 | 72,5 | 72,5 |
| K=9 | 30 | 52,5 | 70 | 70 | 70 |

Figure 5. Accuracy of Javanese Character Classification using LBP-KNN

The best results are obtained at K = 3 and LBP parameters [64 64] according to Table 1.

Table 1. Characters Classification using K = 3 and LBP Parameters

| Character | LBP Parameters | | | | |
|---|---|---|---|---|---|
| | [16 16] | [32 32] | [64 64] | [128 128] | [256 256] |
| Ha_1 | Ra | Ta | Ha | Ha | Ha |
| Ha_2 | Ca | Ca | Ha | Ha | Ha |
| Na_1 | Ra | Na | Na | Na | Na |
| Na_2 | Na | Na | Na | Na | Na |
| Ca_1 | Ra | Ca | Ca | Ca | Ca |
| Ca_2 | Ra | Ca | Ca | Ca | Ca |
| Ra_1 | Ra | Ra | Ra | Ra | Ra |
| Ra_2 | Ra | Ra | Ra | Ra | Ra |
| Ka_1 | Ka | Ka | Ka | Ka | Ka |
| Ka_2 | Ka | Ka | Ka | Ka | Ka |
| Da_1 | Ka | Ka | Ka | Ta | Ta |
| Da_2 | Ka | Ka | Da | Da | Ta |
| Ta_1 | Ta | Ta | Ta | Ta | Ta |
| Ta_2 | Ta | Ta | Sa | Ta | Ta |
| Sa_1 | Wa | La | La | La | La |
| Sa_2 | Ra | Ra | Pa | Ra | Ra |
| Wa_1 | Ra | Wa | Wa | Wa | Wa |
| Wa_2 | Ra | Ra | Sa | Sa | Sa |
| La_1 | Ra | La | La | La | La |
| La_2 | Sa | Sa | Sa | Pa | Pa |
| Pa_1 | Pa | Pa | Pa | La | La |
| Pa_2 | Pa | Pa | Dha | Pa | Pa |
| Dha_1 | Pa | Pa | Dha | Dha | Dha |
| Dha_2 | Dha | Dha | Dha | Pa | Pa |
| Ja_1 | Dha | Ja | Ja | Dha | Dha |
| Ja_2 | Pa | Ja | Ja | Ya | Ya |
| Ya_1 | La | Ja | Ja | Ja | Ja |
| Ya_2 | Pa | Ya | Ya | Ja | Ja |
| Nya_1 | Ga | Nya | Nya | Nya | Nya |
| Nya_1 | Ha | Ha | Nya | Nya | Nya |
| Ma_1 | Ma | Ma | Ma | Ma | Ma |
| Ma_2 | Ra | Ma | Ma | Ma | Ma |
| Ga_1 | Ga | Ga | Ga | Ga | Ga |
| Ga_2 | Ga | Ga | Ga | Ga | Ga |
| Ba_1 | Ga | Ba | Ba | Ba | Ba |
| Ba_2 | Ga | Ba | Ba | Ba | Ba |
| Tha_1 | Pa | La | Tha | Tha | Tha |
| Tha_1 | Ra | Tha | Tha | Tha | Tha |
| Nga_1 | Ra | Nga | Nga | Nga | Nga |
| Nga_1 | La | Nga | Nga | Nga | Nga |
| Accuracy | 32,5% | 80% | 82,5% | 72,5% | 72,5% |

According on Table 1, the test data used amounted to 40 pieces where each character is written each 2 times with different writing styles. Our experimental results have been tested on K between 1 and 9, and LBP parameters between [16 16], [32 32], [64 64], [128 128], and [256 256]. The existence of limited presentation places allows us to represent only the highest performance, which is obtained at K = 3 and LPB parameters [64 64]. The second highest value is obtained in the parameter [32 32].

## 4. CONCLUSION

Based on our experiments, K-Nearest Network (KNLN) shows optimal performance with preprocessing Local Binary Patter (LBP). KNN has been tested with a value of K that is 1,3,5,7,9 and LBP is implemented at values [16 16], [32 32], [64 64], [128 128], [256 256]. The highest KNN performance was obtained at K = 3 and the LBP parameter [64 64] which yielded 82.5% using the training data of 80 images and test data of 40 images. The accuracy value is certainly not as high as the experiments that used the dataset derived from the book. To further improve the accuracy, in the next research can be used a reminder on the dataset used.

## 5. REFERENCES


[1] González, A., Bergasa, L. M., Yebes, J. J., & Bronte, S. (2012, November). A character recognition method in natural scene images. In *Pattern Recognition (ICPR), 2012 21st International Conference on* (pp. 621-624). IEEE.

[2] Yadav, D., Sánchez-Cuadrado, S., & Morato, J. (2013). Optical character recognition for Hindi language using a neural-network approach. *JIPS*, *9*(1), 117-140.

[3] Misnadin, R. A., Mola, S. A., & Fanggidae, A. (2016). Pengenalan Pola Tulisan Tangan dengan Metode K–Nearest Neighbor. *Jurnal Komputer dan Informatika (JICON)*, *2*(1), 65-72.

[4] Fikri, R., Arnia, F., & Muharar, R. (2016). Pengenalan Karakter Tulisan Tangan Jawi Menggunakan Metode New Relative Context dan SVM. *Jurnal Nasional Teknik Elektro dan Teknologi Informasi (JNTETI)*, *5*(3), 1-6.

[5] Munggaran, L. C., Nuryuliani, N., & Choyrivanie, U. (2010) Klasifikasi Karakter Tulisan Tangan berdasarkan Pola Segmen. *Jurnal Ilmiah Program Ilmu Komputer*.

[6] Singh, G., & Lehri, S. (2012). Recognition of Handwritten Hindi Characters using Backpropagation Neural Network. *International Journal of Computer Science and Information Technologies*, *3*(4), 4892-4895.

[7] I. B. Kusuma & A. Kartika. (2017). Image Spoofing Detection Using Local Binary Pattern. *J. Masy. Inform. Indones.* 2(1), 49–54.

[8] Hau, C. C. (Ed.). (2015). *Handbook of pattern recognition and computer vision*. World Scientific.

[9] Akbar, R., & Sarwoko, E. A. (2016). Studi Analisis Pengenalan Pola Tulisan Tangan Angka Arabic (Indian) menggunakan Metode K-Nearest Neighbors dan Connected Component Labeling. *Dinamika Rekayasa*, *12*(2), 45-51.

[10] Indriani, O. R., Kusuma, E. J., Sari, C. A., & Rachmawanto, E. H. (2017, November). Tomatoes classification using K-NN based on GLCM and HSV color space. In *Innovative and Creative Information Technology (ICITech), 2017 International Conference on* (pp. 1-6). IEEE.

[11] Abdullah, S., Huda, S. N., Bohani, F. A., Nayef, B. H., Sahran, S., Al Akash, O., ... & Ismail, F. (2016). Round Randomized Learning Vector Quantization for Brain Tumor Imaging. *Computational and mathematical methods in medicine*, *2016*.

[12] Syam, R. M. (2013). Pengenalan Aksara Jawa Tulisan Tangan dengan Menggunakan Ekstraksi Fitur Zoning dan Klasifikasi K-Nearest Neighbour.