TECHNIUM
Romanian Journal of Applied Sciences and Technology
ISSN: 2668-778X

vol. 16 / 2023

The 7th International Conference on Science Technology

organized by
Faculty of Social Science and
Law Universitas Negeri Manado and
Consortium of International Conference
on Science and Technology

The Innovation Breakthrough in Digital and Disruptive Era

PLUS COMMUNICATION
International Communication & PR

Technium: Romanian Journal of Applied Sciences and Technology
Vol. 16, 124-128, October, 2023
ISSN: 2668-778X
www.techniumscience.com

Neural Network Classification in Javanese Handwriting Recognition using Projection Profile Histogram and Local Binary Pattern Histogram

Fetty Tri Anggraeny[1*], *Yisti* Vita Via[1], *Retno* Mumpuni[1], *Heliza* Rahmania Hatta[2], *Narti* Eka Putri[3], and *Joni* Bastian[1]
[1]Department of Informatics, Universitas Pembangunan Nasional Veteran Jawa Timur, Surabaya, 60294, Indonesia
[2]Department of Informatics, Mulawarman University, Samarinda, Indonesia
[3]Department of Information System, Universitas Nagoya, Batam, Indonesia

Abstract. Indonesia consists of various regional tribes, where each tribe has cultural diversity and some even have their own regional letters, like Javanese tribe has Javanese characters. Javanese letters consist of 20 basic letters called Nglegena script. Subject about Javanese language is delivered to elementary student until now aims to preserve Indonesian culture especially the Javanese. In this study, we present two feature extraction methods are Local Binary Pattern (LBP) and Profile Projection (PP). Neural Network (NN) chosen as classifiers for classifying 20 javanese letters Nglegena. Some digital image processing processes are carried out, are image inversion, dilation, denoising and skeletoning. The Javanese script dataset is taken from the Kaggle database with the name Aksara Jawa: Aksara Jawa Custom Dataset, consists of 2154 train images and 480 test images. The experiment were carried out in two models, Projection Profile Histogram - Neural Network (PPH-NN) and Local Binary Pattern Histogram - Neural Network (LBPH-NN). The experiment show that both feature extraction methods have very good performance, 99.98% PPH-NN and 89.6% LBPH-NN on average.

1 Introduction

Handwriting is the result of someone writing by hand. The handwriting produced by each person is unique, meaning that each person produces a different form of handwriting, even though they are written in the same letter. There is even a field of research that examines handwriting to find out the psychological aspects of the writer.

Indonesia consists of various regional tribes, where each tribe has cultural diversity and some even have their own regional letters, one of which is the Javanese tribe which has Javanese characters. Javanese letters consist of 20 letters that come from the Legend of Ajisaka [1]. Javanese script as one of the cultures in Indonesia must be preserved so that it continues to exist. At this time, it is rare to find the use of Javanese script in everyday life, and many people are unable to read and understand Javanese script. With a lack of people who can read and understand Javanese script writing, it will be a threat to the loss of Indonesian culture. One of the efforts of the Indonesian government to overcome this problem is to provide regional language lessons in schools. With current technological developments, it is possible to preserve regional languages by researching and making related applications to make it easier for people to learn local languages.

Computer vision can be one of the solutions to the above problems through computer applications. Computer vision combines digital image processing, pattern recognition and artificial intelligence which allows computers to analyze one or several images [2]. Pattern recognition in computer vision is in the form of image processing with a series of digital image processing techniques and analyzed using artificial intelligence methods. So that it can be known the group of objects in the image. The recognition system is capable of detecting images of Javanese script handwriting. Then use image classification techniques so as to produce identification of Javanese script handwriting.

Research related to Javanese script is very interesting to study, in addition to scientific purposes but also to preserve Indonesian culture, especially the Javanese. The scope of the research is also very broad, including image processing in Javanese documents, applications that are able to convert written Javanese letters into hanacaraka text and vice versa, recognition of Javanese letters and so on. Next we will focus on feature extraction and classification methods

Several classification methods used in previous studies are k-NN [3], [4], neural networks [4]–[7], CNN [7]–[9], and SVM [10]. And several feature extraction in handwritten Javanese letter are zoning [11], ICZ-ZCZ [12]–[14], LBP [10], [15], [16], and HOG [17], [18]. Budhi et.al using ICZ-ZCZ feature extraction comparing three classification machines with the results of accuracy on Basic Hanacaraka letters being 3.17% using Counterpropagation Network (CPN), 58.12% Evolutionary Neural Network (ENN) 1 layer, and 59.31% ENN 2 layer [12]. Rismiyati et.al

* Corresponding author: fettyanggraeny.if@upnjatim.ac.id

used Deep Neural Network (DNN) and Convolutional Neural Network (CNN) with an accuracy rate of 64.65% and 70.22% [9]. In another study, the multiclass SVM method produced an accuracy of 78% which was then improved by adding Local Binary Pattern feature extraction so that it could increase SVM performance to 85% [10]. As for classification using neural networks, the deep learning Convolutional Neural Network (CNN) approach generally shows better results compared to Multi Layer Perceptron (MLP) [7]. This is shown from the 5 trials conducted, CNN has the highest accuracy of the two comparison MLP models, reaching 89%. Putri et al tried to combine the advantages of CNN with the advantages of SVM to form a new Hybrid CNN-SVM method. CNN has feature extraction capabilities and CNN as the classifier. Hybrid CNN-SVM is able to increase accuracy by 1-2% of the basic CNN method [19]. Krisnawati et al compared 4 classifiers namely k-NN, LDA, SVM and Gaussian Naive Bayes. Of the four methods, k-NN has the highest performance rating of the other 3 methods, although k-NN is the simplest method of the other three methods [4].

This study aims to compare the performance of the two feature extraction methods used in the recognition of Javanese script handwriting. Performance testing is measured using the performance of the neural network classification engine.

2 Methods

In this section we will discuss the technical approach used during the research process. The flow of the research scheme can be seen in Figure 1 below.

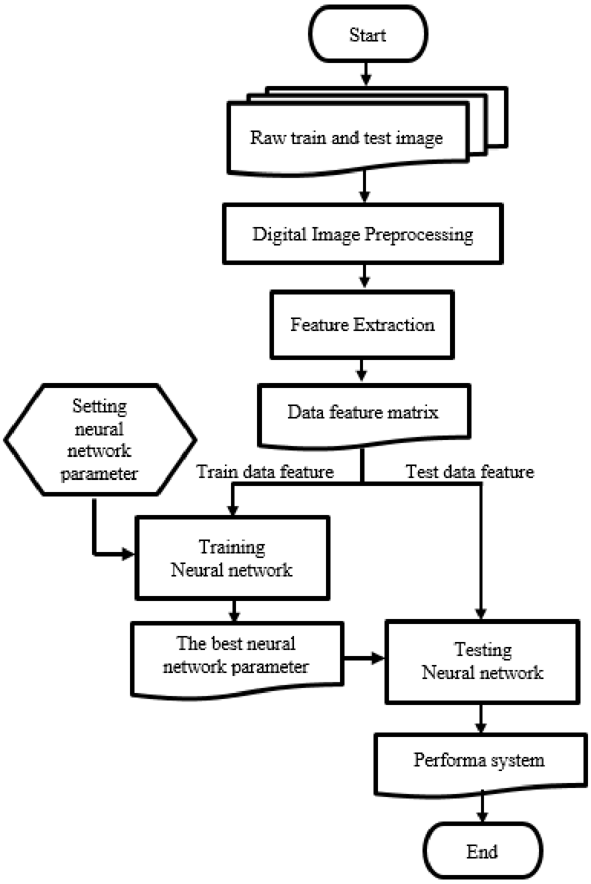

Fig 1. Research Methodology.

The dataset consists of 2 groups, train data and test data. Both go through the same stages of preprocessing and feature extraction to produce a feature matrix. Train data is used in the neural network training process to produce the best neural network structural parameters. The best structure of the neural network is used to test the test data and produce the performance of the system.

2.1 Javanese Script Dataset

This study uses a dataset database from Kaggle with the name Javanese script: Javanese script custom dataset. The dataset is in the form of a basic Javanese script image called Nglegena (see Figure 2). The dataset consists of 2154 train images and 480 test images.

Fig 2. Javanese letters [20].

Technium: Romanian Journal of
Applied Sciences and Technology
Vol. 16, 124-128, October, 2023
ISSN: 2668-778X
www.techniumscience.com

The Javanese characters come from the Legend of Ajisaka which has 2 employees who fight each other until both of them die. Ajisaka immortalized the story in the form of Javanese script or hanacaraka. Ajisaka's writings are as follows [1]:

Hanacaraka, which means '*ana utusan*' or 'there are messengers'.
Datasawala which means '*padha suwala utawa padha regejegan*' or 'fight each other'.
Padhajayanya which means '*padha sektine*' or 'equally strong/ having same supranatural power'.
Magabathanga which means '*Padha dadi bathang*' or 'together to die'.

2.2 Digital Image Preprocessing

In the preprocessing stage, several digital image processing processes are carried out, including image inversion, dilation, denoising using a gaussian kernel and skeletoning.

2.3 Feature Extraction

After the image has gone through a series of preprocessing, the next step is feature extraction. Feature extraction is intended to get more meaningful information from data. In this study, 2 feature extraction techniques were used, namely the projection profile histogram and the local binary pattern.

2.3.1 Local Binary Pattern Histogram (LBPH)

Local Binary Pattern (LBP) was introduced by Ojala et.al in 1996 [21]. The way LBP works involves the concept of neighbor pixels in the kernel with a certain size. In Eq (1), each neighboring pixel ($f_T(x,y)$) is compared to the central pixel ($f_C(x,y)$), if the value is greater than the central pixel ($f_C(x,y)$) then the value is 1 and vice versa is 0. The code is assembled into a binary number and then converted to a decimal number as the LBP value at the center point $LBP(x,y)$.

$$D_T(x,y) = \begin{cases} 1 & if\ f_T(x,y) > f_c(x,y) \\ 0 & otherwise \end{cases} \quad (1)$$

Following the flow of the LBP concept method, each pixel in the image is encoded by LBP. So we get the LBP code matrix which is the same size as the initial image. Then, a histogram of the LBP matrix is made by calculating the variation in LBP values from 0 to 255.

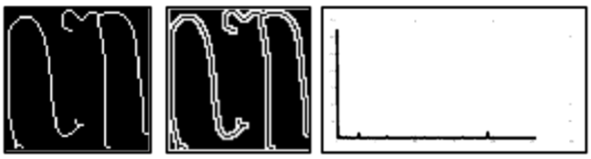

Fig 2. LBP Histogram.

Figure 2 shows the process of forming the LBP histogram. The first step, the image must be converted into an inverse image, then the LBP feature extraction is performed. When using 8-neighbors, an 8-bit binary number with a value in the range 0 to 255 will be formed. From the LBP extraction results, a histogram is made which shows the distribution of the LBP codes. Values on the histogram become input data to the neural network classification machine.

2.3.2 Profile Projection Histogram (PPH)

Projection profiles are generally used for image segmentation. This method has been widely used by previous researchers for line segmentation [22], [23]. The projection profile performs statistical calculations on object pixels according to the horizontal and vertical directions. The process begins by converting the image into a binary image, so that the pixel value is 0 or 1. To produce a projection profile, object pixels are summed in the horizontal or vertical direction. The sum result states the number of pixels of the object in the row horizontal projection profile (HPP) or column vertical projection profile (VPP). In the Eq. (2) and (3), an image has m row and n column pixels, displays a mathematical representation of HPP and VPP [24].

$$VPP(y) = \textstyle\sum_{1 \le x \le m} f(x,y) \quad (2)$$
$$HPP(x) = \textstyle\sum_{1 \le y \le n} f(x,y) \quad (3)$$

Figure 3 shows an illustration of the results of applying the HPP and VPH formulas. The example image consists of 2 Javanese letters located in 1 line. So that the results of the HPP graph are clustered in the middle area according to the position of the character line. Whereas in VPP it appears that there are 2 groups of graphics representing each character in the input image.

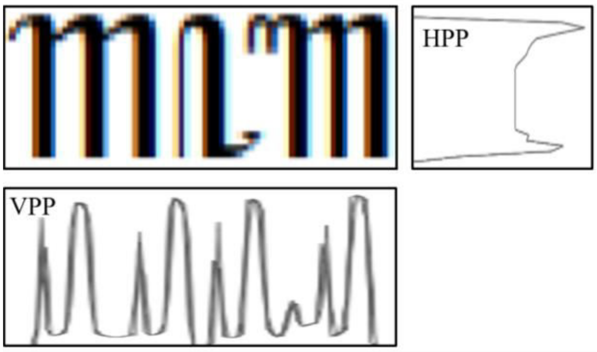

Fig 3. Profile Projection Histogram.

2.4 Neural Network

Neural networks as a machine learning method have proven their reliability in various problems [25]. Neural networks adopt the workings of human nerves. In this study, the neural network is used as a classification engine to test the performance of the 2

TECHNIUM

Technium: Romanian Journal of Applied Sciences and Technology
Vol. 16, 124-128, October, 2023
ISSN: 2668-778X
www.techniumscience.com

feature extraction methods in the previous stage. The neural network structure used consists of 1 hidden layer with a number of nodes between 5 and 50. The learning method used is gradient descent backpropagation and the transfer function used is logsig.

3 Result and Discussion

The experiment were carried out in two models, Projection Profile Histogram - Neural Network (PPH-NN) and Local Binary Pattern Histogram - Neural Network (LBPH-NN). Both use the same NN structure in the hidden layer and the output layer. The number of nodes in the hidden layer varies from 5 to 50 nodes with an increase of 5 nodes. Meanwhile, the input layer adjusts to the amount of data in each model. The number of inputs to the PPH-NN model is 128 nodes and the LBPH-NN model is 256 nodes. Figures 4 and 5 show performance testing graphs on the 2 models.

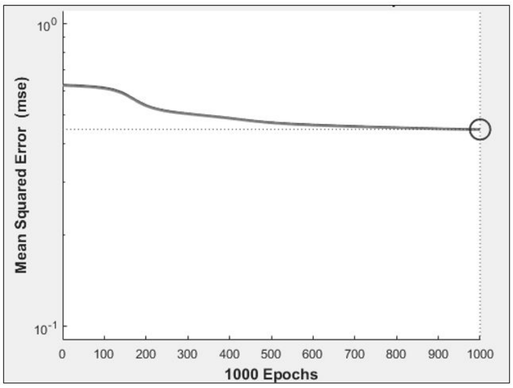

Fig 4. Testing performance of PPH-NN model.

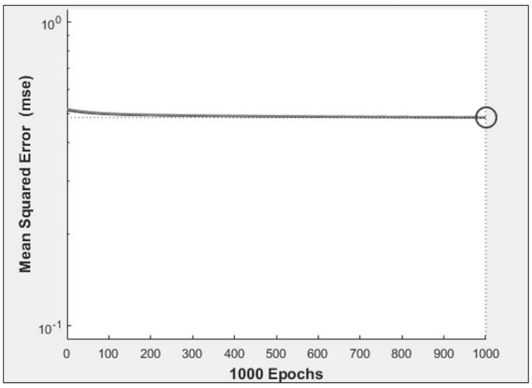

Fig 5. Testing performance of LBPH-NN model.

Table 1 shown the testing results of the research on Neural Network Classification in Javanese Handwriting Recognition using Projection Profile Histogram (PPH) and Local Binary Pattern Histogram (LBPH) as Feature Extraction.

Table 1. Testing performance.

#hidden nodes	**Testing Accuracy (%)**	
	LBPH-NN	**PPH-NN**
5	93.96	100
10	99.58	100
15	72.29	100
20	44.79	100
25	98.13	100
30	98.75	100
35	97.92	100
40	98.33	99.79
45	93.96	100
50	98.33	100

From the table above, almost all accuracy neural network results with PPH feature extraction show overfitting results. It appears that the accuracy reaches 100%. Meanwhile, the Local Binary Pattern Histogram (LBPH) feature extraction shows more optimum results. Based on Table 1, an average accuracy of 99.98% PPH and 89.6% LBPH is obtained.

4 Conclusion

This research has successfully implemented two feature extraction methods, namely projection profile and local binary pattern. Both methods are measured for performance using a neural network classification engine. Experiments that have been done show very good results in both, 99.98% PPH-NN and 89.6% LBPH-NN on average. But the PPH-NN model shows overfitting in the majority of experiments.

In future studies, it is necessary to normalize PPH and LBPH data before being used as input data in a classification machine. Normalization needs to be done with the aim that the data will have a uniform range of values between 0 to 1. Thus the classification process using machine learning will provide more general results for all data.

References

[1] F. Sari, Mr. Supana, and S. Suwandi, “Problem based learning model assisted edmodo,” in *Proceedings of the 1st International Conference on Education Innovation (ICEI 2017)*, Atlantis Press, Feb. 2018, pp. 112–116. doi: 10.2991/icei-17.2018.30.

[2] R. M. Haralick and L. G. Shapiro, “Glossary of Computer Vision Terms,” *Pattern Recognit*, vol. 24, no. 1, pp. 69–93, 1991.

[3] N. D. A. Partiningsih, R. R. Fratama, C. A. Sari, D. R. I. M. Setiadi, and E. H. Rachmawanto, “Handwriting Ownership Recognition using Contrast Enhancement and LBP Feature Extraction based on KNN,” *Proceedings - 2018 5th International Conference on Information Technology, Computer and Electrical Engineering, ICITACEE 2018*, pp. 342–346, 2018, doi: 10.1109/ICITACEE.2018.8576945.

[4] L. D. Krisnawati and A. W. Mahastama, “Building classifier models for on-off javanese character recognition,” in *ACM International Conference Proceeding Series*, Association for Computing Machinery, Dec. 2019. doi: 10.1145/3366030.3366050.

[5] G. Abdul Robby, A. Tandra, I. Susanto, J. Harefa, and A. Chowanda, “Implementation of optical character recognition using tesseract with the

javanese script target in android application," in *Procedia Computer Science*, Elsevier B.V., 2019, pp. 499–505. doi: 10.1016/j.procs.2019.09.006.

[6] A. Mishra and D. Singh, "Handwritten Digit Recognition using Combined Feature Extraction Technique and Neural Network," *International Journal of Current Engineering and Scientific Research (IJCESR)*, vol. 21, no. 2, pp. 80–88, 2017.

[7] C. K. Dewa, A. L. Fadhilah, and Afiahayati, "Convolutional Neural Networks for Handwritten Javanese Character Recognition," *IJCCS (I ndonesian J ournal of Computing and C ybernetics Systems)*, vol. 12, no. 1, pp. 83–94, 2018, Accessed: Aug. 03, 2023. [Online]. Available: https://jurnal.ugm.ac.id/ijccs/article/view/31144/19951

[8] S. Arivazhagan, M. Arun, and D. Rathina, "Recognition of handwritten characters using deep convolution neural network," *J Natl Sci Found*, vol. 49, no. 4, pp. 503–511, 2022, doi: 10.4038/jnsfsr.v49i4.9825.

[9] K. Rismiyati and A. Nurhadiyatna, "Deep Learning for Handwritten Javanese Character Recognition," *Proceedings - 2017 1st International Conference on Informatics and Computational Sciences, ICICoS 2017*, vol. 2018-Janua, pp. 59–63, 2017, doi: 10.1109/ICICOS.2017.8276338.

[10] C. A. Sari, W. S. Sari, V. A. Shelomita, M. R. Kusuma, S. A. Puspa, and M. Bima Gusta, "The Involvement of Local Binary Pattern to Improve the Accuracy of Multi Support Vector-Based Javanese Handwriting Character Recognition," *Journal of Applied Intelligent System*, vol. 8, no. 2, pp. 206–215, Jul. 2023, doi: 10.33633/JAIS.V8I2.8450.

[11] A. N. Handayani, H. W. Herwanto, K. L. Chandrika, and K. Arai, "Recognition of Handwritten Javanese Script using Backpropagation with Zoning Feature Extraction," *Knowledge Engineering and Data Science*, vol. 4, no. 2, p. 117, Dec. 2021, doi: 10.17977/um018v4i22021p117-127.

[12] G. S. Budhi and R. Adipranata, "Comparison of Bidirectional Associative Memory, Counterpropagation and Evolutionary Neural Network for Java Characters Recognition," in *Proceedings - 2014 International Conference on Advanced Informatics: Concept, Theory and Application, ICAICTA 2014*, IEEE, Aug. 2015, pp. 7–10. doi: 10.1109/ICAICTA.2014.7005906.

[13] G. S. Budhi and R. Adipranata, "Handwritten javanese character recognition using several artificial neural network methods," *Journal of ICT Research and Applications*, vol. 8, no. 3, pp. 195–212, Aug. 2015, doi: 10.5614/itbj.ict.res.appl.2015.8.3.2.

[14] E. M. Zamzami, S. Hayanti, and E. B. Nababan, "Diagonal Based Feature Extraction and Backpropagation Neural Network in Handwritten Batak Toba Characters Recognition," *Kinetik: Game Technology, Information System, Computer Network, Computing, Electronics, and Control*, vol. 6, no. 2, pp. 117–126, May 2021, doi: 10.22219/kinetik.v6i2.1212.

[15] A. Susanto, I. U. W. Mulyono, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, "Javanese Script Recognition based on Metric, Eccentricity and Local Binary Pattern," *2021 International Seminar on Application for Technology of Information and Communication (iSemantic) Proceeding*, 2021.

[16] A. Susanto, D. Sinaga, C. A. Sari, E. H. Rachmawanto, and D. R. I. M. Setiadi, "A High Performace of Local Binary Pattern on Classify Javanese Character Classification," *Scientific Journal of Informatics*, vol. 5, no. 1, p. 8, 2018, doi: 10.15294/sji.v5i1.14017.

[17] Rismiyati, Khadijah, and D. E. Riyanto, "HOG and Zone Base Features for Handwritten Javanese Character Classification," *2018 2nd International Conference on Informatics and Computational Sciences, ICICoS 2018*, no. 1, pp. 131–135, 2018, doi: 10.1109/ICICOS.2018.8621781.

[18] A. Susanto, C. A. Sari, I. Utomo, W. Mulyono, and M. Doheir, "Histogram of Gradient in K-Nearest Neighbor for Javanese Alphabet Classification," vol. 8, no. 2, pp. 289–296, 2021, doi: 10.15294/sji.v8i2.30788.

[19] D. U. K. Putri, D. N. Pratomo, and A. Azhari, "Hybrid convolutional neural networks-support vector machine classifier with dropout for Javanese character recognition," *Telkomnika (Telecommunication Computing Electronics and Control)*, vol. 21, no. 2, pp. 346–353, Apr. 2023, doi: 10.12928/TELKOMNIKA.v21i2.24266.

[20] B. Boyolali, "Legenda Ajisaka dan Asal Mula Aksara Jawa," 2015. https://bp3d.boyolali.go.id/index.php/artikel/item/144-legenda-ajisaka-dan-asal-mula-aksara-jawa

[21] T. Ojala, M. Pietikainen, and D. Harwood, "Pergamon A Comparative Study Of Texture Measures With Classification Based On Feature Distributions," *Pattern Recognit*, vol. 29, no. l, pp. 51–59, 1996, doi: https://doi.org/10.1016/0031-3203(95)00067-4.

[22] L. Liliana, S. M. Soephomo, G. S. Budhi, and R. Adipranata, "Segmentation of Hanacaraka Characters using Double Projection Profile and Hough Transform," in *Big Data Technologies and Applications: 8th International Conference, BDTA 2017*, Springer Verlag, 2018, pp. 29–37. doi: 10.1007/978-3-319-98752-1_4.

[23] A. W. Mahastama and L. D. Krisnawati, "Optical Character Recognition for Printed Javanese Script using Projection Profile Segmentation and Nearest Centroid Classifier," *2020 Asia Conference on Computers and Communications, ACCC 2020*, pp. 52–56, 2020, doi: 10.1109/ACCC51160.2020.9347895.

[24] S. Ekici and H. Jawzal, "Breast cancer diagnosis using thermography and convolutional neural networks," *Med Hypotheses*, vol. 137, p. 109542, Apr. 2020, doi: 10.1016/j.mehy.2019.109542.

[25] O. I. Abiodun, A. Jantan, A. E. Omolara, K. V. Dada, N. A. E. Mohamed, and H. Arshad, "State-of-the-Art in Artificial Neural Network Applications: A Survey," *Heliyon*, vol. 4, no. 11, p. e00938, 2018, doi: 10.1016/j.heliyon.2018.e00938.