Improving The Results of Learning Nglegena Javanese Handwriting Using Backpropagation Artificial Neural Network

Arif Budiman[1]; Abdul Fadlil[2]; Rusydi Umar[3]

[1,3]Program Studi Magister Informatika, Universitas Ahmad Dahlan, Indonesia
[2]Program Studi Teknik Elektro, Universitas Ahmad Dahlan, Indonesia
[1]Corresponding Email: arief1708048029@webmail.uad.ac.id, Phone Number : 0823 xxxx xxxx

Article History:
Received: Nov 30, 2022
Revised: Dec 28, 2022
Accepted: Jan 06, 2023
Online First: Jan 10, 2023

Keywords:
Javanese Script, Backpropagation Artificial Neural Network, Histogram Chain Code

Kata Kunci:
Aksara Jawa, Backpropagation Artificial Neural Network, Histogram Chain Code

How to cite:
Budiman, A., Fadlil, A., & Umar, R. (2023). Improving The Results of Learning Nglegena Javanese Handwriting Using Backpropagation Artificial Neural Network. *Edunesia: Jurnal Ilmiah Pendidikan*, 4(1), 259-269.

Abstract: The Nglagena Javanese script is one of the cultural assets of the Indonesian nation that needs to be preserved. Various efforts have been made to preserve this script, one of which is using information technology as a learning medium for the Nglagena Javanese script. Information Technology allows the Javanese script to be introduced interactively to students. To support this need, one of which is the ability of information technology to classify Javanese script. Classification of Javanese script is carried out using the Backpropagation Artificial Neural Network (BANN) method. Twenty primary Javanese characters are classified as classes using the Backpropagation Artificial Neural Network (BANN) method. The stages of this research are initial processing, feature extraction, model training, and model testing. Initial processing is carried out to prepare image data so that it is ready for the feature extraction process. The feature extraction method is the Histogram Chain Code (HCC) to obtain the main characteristics of each data class or character of the Nglegena Javanese script. This study compares three research models by adjusting the ratio between the training image data and the test image so that the model that produces the highest accuracy value is produced. The model training and testing process uses 2000 image data, with the percentage distribution of training image data and test images, namely 20%, 80%, second 50%, 50%, and third 80%, 20%, resulting in different levels of accuracy. The results are to produce successive accuracy of 66%, 72%, and 88%.

Abstrak: Aksara Jawa Nglagena adalah salah satu kekayaan budaya Bangsa Indonesia yang perlu dilestarikan, berbagai upaya untuk melestarikan aksara ini telah dilakukan salah satunya penggunaan teknologi informasi sebagai media belajar Aksara Jawa Nglagena. Teknologi Informasi memungkinkan Aksara Jawa dikenalkan secara interaktif kepada pelajar. Untuk mendukung keperluan tersebut salah satunya dibutuhkan kemampuan Teknologi informasi dalam mengklasifikasikan Aksara Jawa. Klasifikasi Aksara Jawa dilakukan mengan metode *Backpropagation Artificial Neural Network* (BANN). Dua puluh Aksara Jawa Dasar digunakan sebagai kelas yang diklasifikasikan dengan metode *Backpropagation Artificial Neural Network* (BANN). Tahapan penelitian ini pemrosesan awal, Ekstrasi ciri, Pelatihan model dan pengujian model. Pemrosesan awal dilakukan untuk mempersiapkan data citra sehingga siap untuk dilakukan proses ekstraksi ciri. Metode ekstrasi ciri yang digunakan yaitu *Histogram Chain Code* (HCC) untuk memperoleh ciri-ciri utama dari masing-masing kelas data atau karakter Aksara Jawa Nglegena. Penelitian ini membandingkan tiga model penelitian dengan mengatur rasio antara jumlah data citra latih dan citra uji sehingga dihasilkan model yang mengeluarkan nilai akurasi yang paling tinggi. Proses Pelatihan dan pengujian model menggunakan 2000 data citra, dengan pembagian prosentase perbandingan data citra latih dan citra uji yaitu 20%, 80%, kedua 50%, 50% dan ketiga 80%, 20%, sehingga menghasilkan tingkat akurasi yang berbeda-beda. Hasil yang diperoleh yaitu menghasilkan akurasi berturut-turut sebesar 66%, 72% dan 88%.

https://doi.org/10.51276/edu.v4i1.339

A. Introduction

One wealth of the Indonesian nation is the many tribes with their languages and scripts. One of the scripts that have existed since long before the independence of the State of Indonesia is the Javanese script, with the use of Latin script, which is used by almost every aspect of life, both official activities and daily use, the use of traditional scripts, especially Javanese script, is increasingly rare. Argues that script is a cultural product that has significance for the development of human life (Rahardjo, 2019). One of the efforts to preserve the nation's wealth is introducing this culture to the early generations through technology, especially computers.

The introduction of the Javanese script to the Indonesian population, especially the Javanese, especially those in three provinces, namely DIY, East Java, and Central Java, through the world of education by incorporating Javanese local content subjects. One of the efforts to preserve the existence of the Javanese script is by utilizing information technology so that learning the Javanese script is more effective and efficient (Surjono, 2017). One of the technologies developed in this regard is handwritten character recognition technology (Chandra, 2022), especially recognizing written characters. Javanese script is needed to make Javanese script learning media more interactive because it allows Javanese script learning media to provide direct feedback on students' understanding of Javanese script.

The Javanese script, or Hanacaraka, is one of the traditional scripts in Indonesia. The use of the Javanese script itself is more common in the Yogyakarta and Surakarta areas. It can be seen that many street names use Javanese script in their writing. Currently, writing Javanese scripts has been forgotten a lot. Based on observations made by researchers on several students at SMA Negeri 1 Kalasan, students find it difficult to understand writing with Javanese script.

The Javanese script (carakan), which is used in the spelling of the Javanese language, basically consists of twenty main characters, syllabic (syllabic). The Javanese script consists of 20 scripts that are still legendary or have not been attached to sandhangan. The order of the Javanese script still legit is called dentawyanjana, which comes from dental (tooth) and wyanjana (sound). Ordinary Javanese script is also given the meaning carakan, namely the sequence of Javanese script starting from the letters ha to nga (Adyningsih et al., 2022).

Javanese script handwriting character recognition technology can be used to make interactive Javanese script learning media because, according to (Fatima, 2020) and the use of interactive learning media in the teaching and learning process can increase students' understanding of learning Javanese script and (Prihatin, 2015) learning media Javanese script in the form of interactive multimedia can facilitate understanding of reading Javanese script using a partner.

Media comes from Latin, the plural form of medium, which means intermediary or delivery of messages from the sender to the recipient (Banat et al., 2022). In line with this, Suparmi (2018) argues that media comes from the Latin medius, which means middle, intermediary, or intermediary. Media learning resources are helpful tools in teaching and

https://doi.org/10.51276/edu.v4i1.339

learning activities. The aids can represent something the teacher cannot convey through words or sentences (Alfan & Sulistiyo, 2015).

Various methods have been developed, from initial processing and feature extraction to classifying Javanese script handwritten images. The Histogram Chain Code (HCC) method was used in this study because this method is good enough to be used in feature extraction of the Davnagari character (Arora et al., 2010), whose writing follows a lot of curved curves like Javanese characters. The classification method chosen is Backpropagation Artificial Neural Network (BANN) because Neural Networks can recognize handwritten characters efficiently and reliably (Aqab & Tariq, 2020).

This study compares the percentage of image data for the training process and testing the Javanese script handwriting recognition model Nglagena with Histogram Chain Code (HCC) feature extraction and the Backpropagation Artificial Neural Network (BANN) Classification method because similar research has never been done before. This research aims to produce a character recognition model for the Javanese script using the Histogram Chain Code (HCC) feature extraction method and the Backpropagation Artificial Neural Network (BANN) Classification method with reasonable accuracy.

B. Method

The stages of character recognition of the Javanese script are divided into several stages: initial processing, feature extraction, model training, and model testing. Image data, as shown in Figure 1, includes 20 classes of nglagena Javanese Script Characters, namely Ha, Na, Ca, Ra, Ka, Da, Ta, Sa, Wa, La, Pa, Dha, Ja, Ya, Nya, Mo, Ga, Ba, Tha, Nga.

Figure 1. Characters of the Nglagena Javanese script

The total image data used in this study is 100 for each letter class, so the total image used is 2000. The entire research process is described in a flowchart, as shown in Figure 2.

https://doi.org/10.51276/edu.v4i1.339

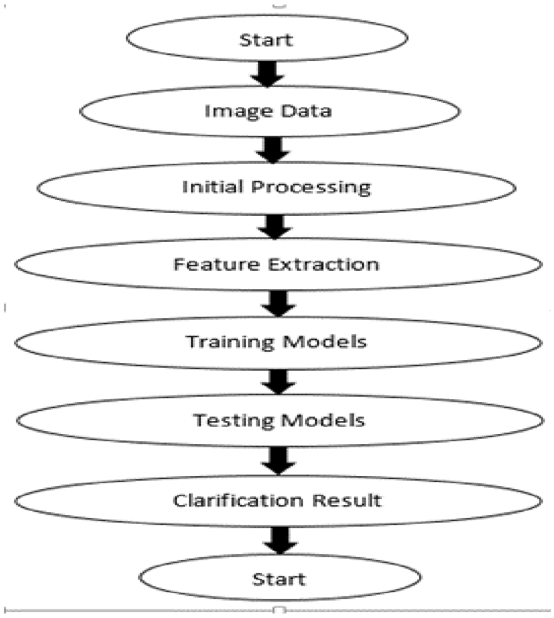

Figure 2. Research flowchart

This study uses the method. Namely, the data obtained is pre-processed. After that, the feature extraction process is carried out, and the training process is carried out. The testing process is carried out to determine the best accuracy of the variations. The system is built using the Matlab programming language script. According to (Sugianela & Suciati, 2019) the chain code feature extraction method is often used in research. Histogram Chain Code is a feature extraction method based on the freeman chain code by calculating the frequency obtained from each image block (Qian et al., 2013).

The process of testing the character recognition of Javanese script handwritten images through the initial processing stages, feature extraction, and entering image data into the Backpropagation Artificial Neural Network (BANN) network model that has been created through the training model produces data in the form of predictions of Javanese script character classes. The testing model stages are carried out with the number of test images of 80%, 50%, and 20%, respectively. Accuracy calculation is done by using equation 1.

$$Accuracy\ (\%) = \frac{predictions}{Correct} x\ 100\% \quad (1)$$

C. Result and Discussion

This study produced data including image data used, determination of data classes, distribution of data percentages for the training model and testing model, and the accuracy of the tests performed.

https://doi.org/10.51276/edu.v4i1.339

Table 1. Data on Javanese Script Character Image

No	Alphabet	Amount of data
1	Ha	100
2	Na	100
3	Ca	100
4	Ra	100
5	Ka	100
6	Da	100
7	Ta	100
8	Sa	100
9	Wa	100
10	La	100
11	Pa	100
12	Dha	100
13	Ja	100
14	Ya	100
15	Nya	100
16	Mo	100
17	Ga	100
18	Ba	100
19	Tha	100
20	Nga	100
Jumlah		2000

The classification process using the Backpropagation Artificial Neural Network (BANN) method in this study consisted of 20 classes according to the number of Javanese Nglageno characters. The amount of data for each character class is 100, so the total data is 2000. Each character class is assigned a numeric code to differentiate between character classes, as shown in Table 2.

Table 2. Javanese Script Class Data

No	Character	Class Code
1	Ha	1
2	Na	2
3	Ca	3
4	Ra	4
5	Ka	5
6	Da	6
7	Ta	7
8	Sa	8
9	Wa	9
10	La	10
11	Pa	11
12	Dha	12
13	Ja	13
14	Ya	14

https://doi.org/10.51276/edu.v4i1.339

15	Nya	15
16	Mo	16
17	Ga	17
18	Ba	18
19	Tha	19
20	Nga	20

This study produced three research models with variations in the comparison of the amount of training image data and the amount of test image data, which were produced by adjusting the percentage of the amount of data used for the training model and testing model to the actual data, the percentage comparison is shown in table 3.

Table 3. Image Data Distribution

Total Training Data	Total Test Data	Model Name
20%	80%	Model A
50%	50%	Model B
80%	20%	Model C

Table 4. Data for the Javanese Script Class Model A

Nama Kelas	Data Latih	Data Uji
Ha	20	80
Na	20	80
Ca	20	80
Ra	20	80
Ka	20	80
Da	20	80
Ta	20	80
Sa	20	80
Wa	20	80
La	20	80
Pa	20	80
Dha	20	80
Ja	20	80
Ya	20	80
Nya	20	80
Mo	20	80
Ga	20	80
Ba	20	80
Tha	20	80
Nga	20	80
Jumlah	400	1600

Table 5. Data from Model A Analysis

Nama Kelas	Nilai Benar	Nilai Prediksi
Ha	80	37
Na	80	58

https://doi.org/10.51276/edu.v4i1.339

Ca	80	52
Ra	80	60
Ka	80	66
Da	80	65
Ta	80	49
Sa	80	39
Wa	80	19
La	80	45
Pa	80	31
Dha	80	57
Ja	80	74
Ya	80	55
Nya	80	67
Mo	80	69
Ga	80	78
Ba	80	39
Tha	80	46
Nga	80	40
Jumlah	1600	1048

Table 6. Model B Javanese Script Class Data

Nama Kelas	Data Latih	Data Uji
Ha	50	50
Na	50	50
Ca	50	50
Ra	50	50
Ka	50	50
Da	50	50
Ta	50	50
Sa	50	50
Wa	50	50
La	50	50
Pa	50	50
Dha	50	50
Ja	50	50
Ya	50	50
Nya	50	50
Mo	50	50
Ga	50	50
Ba	50	50
Tha	50	50
Nga	50	50
Jumlah	1000	1000

https://doi.org/10.51276/edu.v4i1.339

Table 7. Data From the Analysis of Model B

Nama Kelas	Nilai Benar	Nilai Prediksi
Ha	50	34
Na	50	40
Ca	50	34
Ra	50	47
Ka	50	34
Da	50	39
Ta	50	38
Sa	50	21
Wa	50	28
La	50	44
Pa	50	32
Dha	50	33
Ja	50	44
Ya	50	36
Nya	50	31
Mo	50	44
Ga	50	44
Ba	50	29
Tha	50	43
Nga	50	20
Jumlah	1000	715

Table 8. Model C Javanese Script Class Data

Nama Kelas	Data Latih	Data Uji
Ha	80	20
Na	80	20
Ca	80	20
Ra	80	20
Ka	80	20
Da	80	20
Ta	80	20
Sa	80	20
Wa	80	20
La	80	20
Pa	80	20
Dha	80	20
Ja	80	20
Ya	80	20
Nya	80	20
Mo	80	20
Ga	80	20
Ba	80	20
Tha	80	20

https://doi.org/10.51276/edu.v4i1.339

Nga	80	20
Jumlah	1600	400

Table 9. Data From the Analysis of Model C

Nama Kelas	Nilai Benar	Nilai Prediksi
Ha	20	13
Na	20	19
Ca	20	16
Ra	20	19
Ka	20	18
Da	20	18
Ta	20	18
Sa	20	18
Wa	20	15
La	20	18
Pa	20	18
Dha	20	18
Ja	20	19
Ya	20	16
Nya	20	17
Mo	20	19
Ga	20	19
Ba	20	18
Tha	20	18
Nga	20	16
Jumlah	100	350

Discussion

In model A, 400 training image data are used for the training process, while the testing process uses 1,600 test image data. In Model A, the training image data is set to be less than the test image data with a ratio of 1:4.

The difference in setting the amount of data is meant to see the highest accuracy of the various variations of the data previously set. Based on the test results of the total number of test image data of 1600 data, it produces 1048 images whose predictions are correct.

The prediction accuracy of Model A class classification is equal to:

$$Accuracy\ (\%) = \frac{1048}{1600} x\ 100\%$$
$$Accuracy\ (\%) = 66\%$$

Model B uses a ratio of training image data and tests image data of 50% / 50% of the total data. In Model B, 1000 training image data and 1000 test image data are used. The number of image data used in Model B has the same ratio of training images and test images comparison of the number of training image data and test image data is 1:1. The prediction accuracy of Model B class classification is equal to:

https://doi.org/10.51276/edu.v4i1.339

$$Accuracy\ (\%) = \frac{715}{1000} x\ 100\%$$
$$Accuracy\ (\%) = 72\%$$

Model C uses a ratio of training image data and tests image data of 80% / 20% of the total data. In Model C, 1600 training image data and 400 test image data are used. In Model C, the test data is smaller than the training data. The comparison between training image data and test image data is 4:1. The prediction accuracy of Model C class classification is equal to:

$$Accuracy\ (\%) = \frac{350}{400} x\ 100\%$$
$$Accuracy\ (\%) = 88\%$$

Based on the calculations carried out on the results of the C model test, the highest accuracy value is 88%. This value is the highest accuracy value in this study.

D. Conclusion

Based on the analysis of the three models, there are still many data classes where the difference between the predicted value and the correct value is quite significant, which means that the predictive ability of the system in this data class is small. Suggestions for further research are that feature extraction capabilities in capturing the uniqueness of Javanese script data should be maximized, such as data classes Ha, Ca, Wa, Ya, and Nga.

References

Adyningsih, M., Rusmawati, R. D., & Nurjati, N. (2022). Pengembangan Buku Ajar Cara Cepat Membaca Aksara Jawa dengan Metode Al-Barqy di Tingkat Sekolah Menengah Pertama. *Tafhim Al-'Ilmi*, *14*(1), 81-110.

Alfan, M., & Sulistiyo, E. (2015). Perbandingan media pembelajaran (AutoPlay Media Studio) sebagai alat bantu pembelajaran memperbaiki CD Player siswa Kelas XI di SMK Negeri 3 Surabaya. *Jurnal Pendidikan Teknik Elektro*, *4*(1).

Aqab, S., & Tariq, M. U. (2020). Handwriting recognition using artificial intelligence neural network and image processing. *International Journal of Advanced Computer Science and Applications*, *11*(7).

Arora, S., Bhattacharjee, D., Nasipuri, M., Basu, D. K., & Kundu, M. (2010). Application of Statistical Features in Handwritten Devnagari Character Recognition. *arXiv preprint arXiv:1006.5911*. https://doi.org/10.48550/arXiv.1006.5911

Banat, A., Febrianti, M., Martiani, M., Juwita, J., & Gustini, G. (2022). Pendampingan Penggunaan Teknologi Media dan Internet Bagi Pengurus Bumdes Teratai Indah Desa Nanti Agung Ilir Talo Kabupaten Seluma. *Jurnal Dehasen Untuk Negeri*, *1*(1), 33-

https://doi.org/10.51276/edu.v4i1.339

36

Chandra, A. (2022). *Handwriting Recognition with ML (An In-Depth Guide)*. Retrieved from https://nanonets.com/blog/handwritten-character-recognition/

Fatima, W. Q., Khairunisa, L., & Prihatminingtyas, B. (2020). Metode Pembelajaran Berbasis Game untuk Meningkatkan Ketrampilan Membaca dan Menulis Aksara Jawa. *Inteligensi: Jurnal Ilmu Pendidikan*, *3*(1), 17-22. https://doi.org/10.33366/ilg.v3i1.1766

Prihatin, K. (2015). Pengembangan Multimedia Interaktif Aksara Jawa Untuk Siswa Kelas V Sdn Sabdodadi Keyongan Bantul. *Basic Education*, *4*(7).

Qian, Y., Xichang, W., Huaying, Z., Zhen, S., & Jiang, L. (2013). Recognition method for handwritten digits based on improved chain code histogram feature. *3rd International Conference on Multimedia Technology, ICMT-13*, 431-438. Atlantis Press.

Rahardjo, T., Degeng, I. N. S., & Soepriyanto, Y. (2019). Pengembangan Multimedia Interaktif Mobile Learning Berbasis Anrdroid Aksara Jawa Kelas X Smk Negeri 5 Malang. *Jurnal Kajian Teknologi Pendidikan*, *2*(3), 195-202. https://doi.org/10.17977/um038v2i32019p195

Sugianela, Y., & Suciati, N. (2019). Ekstraksi fitur pada pengenalan karakter Aksara Jawa berbasis Histogram of Oriented Gradient. *JUTI: Jurnal Ilmiah Teknologi Informasi*, *17*(1), 64-72. https://doi.org/10.12962/j24068535.v17i1.a819

Suparmi, S. (2018). Penggunaan Media Komik Dalam Pembelajaran IPA di Sekolah. *Journal of Natural Science and Integration*, *1*(1), 62-68. https://doi.org/10.24014/jnsi.v1i1.5196

Surjono, H. D. (2017). Multimedia Pembelajaran Interaktif Konsep dan Pengembangan. Yogyakarta: Universitas Negeri Yogyakarta.