EDUMASPUL
JURNAL PENDIDIKAN
Vol. 7 – No. 1, year (2023), page 147-153
| ISSN 2548-8201 (Print) | 2580-0469) (Online) |

Identification of Learning Javanese Script Handwriting Using Histogram Chain Code

Arif Budiman[1], Abdul Fadlil[2], Rusydi Umar[3]
[1,3](Prodi Magister Informatika, Universitas Ahmad Dahlan, Indonesia).
[2](Prodi Teknik Elektro, Universitas Ahmad Dahlan, Indonesia).

* Corresponding Author. E-mail: [1]arief1708048029@webmail.uad.ac.id

Receive: 05/01/2023	Accepted: 15/02/2023	Published: 01/03/2023

Abstract

One of the wealth of the Indonesian nation is the many tribes with their own languages and scripts. One of the scripts that has existed since long before the independence of the State of Indonesia is the Javanese script, with the use of Latin script used by almost every aspect of life, both official official activities and daily use, the use of traditional scripts, especially Javanese script, is increasingly scarce. To facilitate learning the Javanese script, learning media is needed with the ability to recognize Javanese characters. In this study, pre-processing was used, especially feature extraction using the Histogram Chain Code (HCC) method and classification using artificial neural networks using the Multi Layer Perceptron method. This study compares four research models by setting the number of HCC feature extraction parameters obtained from one intact image and 3 divided images of 4, 9 and 16 parts respectively so that the total parameters of each HCC model are 8, 32, 72 and 128 parameters characteristic. The training and testing process using the Multi Layer Perceptron method uses 2000 handwritten Javanese script image data which is divided into 80%, namely 1600 images for the training process and 400 images for the testing process. This research resulted in different accuracies, namely 57%, 78%, 83% and 76%. The best accuracy is obtained from the HCC model with 72 parameters and the image is divided into 9 sections.

Keywords: *Learning Javanese, Handwriting, Histogram Chain Code*

Introduction

One of the wealth of the Indonesian nation is the many tribes with their own languages and scripts. One of the scripts that has 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 official activities and daily use, the use of traditional scripts, especially Javanese script, is increasingly rare. One of the efforts to preserve the nation's wealth is to introduce this culture to the early generations through technology, especially computers so that learning becomes more efficient and effective [1].

The introduction of the Javanese script to the Indonesian population, especially the Javanese, is carried out through the world of education by including it as a local content subject, especially in the Special Region of Yogyakarta, Central Java and East Java [2]. One of the efforts to preserve the Javanese script is carried out by utilizing information technology, namely through Javanese script handwriting character recognition technology.

Various methods have been developed to introduce Javanese characters. In previous research that has been done is the development of flutter-based Javanese script interactive learning media in Javanese language subjects resulting in the conclusion that flutter-based interactive Javanese script learning media is effective in increasing understanding in learning Javanese script [3]. In addition, in the research on the development of Javanese script interactive multimedia for Javanese language learning, fifth grade students at Sabdodadi Keyongan Bantul Elementary School concluded that the Javanese script interactive multimedia was able to facilitate students' understanding of reading material in Javanese script words using pairs [4].

Character recognition Javanese script handwriting 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. Various methods have been developed from pre-processing, feature extraction to classification of Javanese script handwritten images. The Freeman Chain Code method is a good method for identifying handwritten characters [5]. Several researchers have developed this method to improve the ability to extract features from handwriting such as Differential Chain Code [6], Histogram ChainCode, Vertex Chain Code [7], several other Modified Chain Code.

In this study, the application of HCC feature extraction was carried out on handwritten image objects in Javanese script Ngagena using the Multi Layer Perceptron classification method. researched.

Methods

This study aims to produce a writing identification system for Javanese script using the Percetron Multilayer. To achieve this, systematic steps are needed as shown in Figure 1.

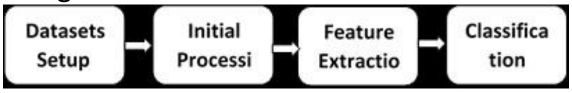

Figure 1. Research Stages

The research object studied was the image of Javanese script handwriting in the amount of 100 data for each letter of the Nglagena Javanese script which was divided into 80 training data and 20 test data. The Javanese script data set used uses a dataset created by phiard version 10 with a public license, the data set acquisition methodology is real human writing. The Javanese script handwritten image files are separated into one file for each Javanese script, so there are a total of 2000 Javanese script image files, with dimensions of 224x224 pixels. Some examples of the datasets used in this study can be seen in Figure 2.

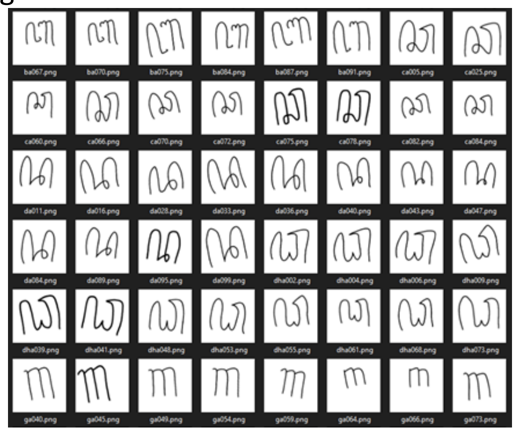

Figure 2. Dataset files

Freeman in 1961 which is used to represent digital curves and Freeman codes. The Frman Chain Code algorithm aims to represent contours including pixels of an object that are interconnected with a wind direction guide. Chain code is a simple way to present an image [8]. The way the chain code algorithm works is by giving a rotation mark that is adjusted to the direction of the wind you want to use. The final result of the chain code is a feature vector containing information on the order of the chain code forming the object [9].

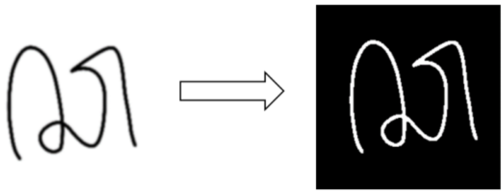

Figure 3. Changing the PNG file format to Binary format

One of the Freeman Chain Code-based feature extraction methods is the Histogram Chain Code (HCC) which is formed by calculating the direction frequency of the chain code vector of the object. So that we get 8 characteristic parameters of the object. This research wants to know the highest classification accuracy by examining 4 HCC models resulting from variations of the intact image which is broken into 1, 4, 9, and 16 parts as shown in Figure 4. The intact image is formed by cropping the parts of the Javanese script, and resizing so that it has the same size of 90 x 150 pixels [10].

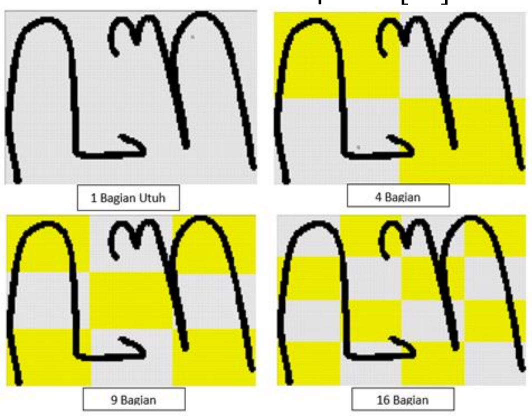

Figure 4. Cutting the image object into 4, 9, and 16 parts

HCC feature extraction for one intact image has 8 parameters, while for divided images, the number of parameters is obtained from the number of parts multiplied by 8. In this study, 4 HCC models were made which were obtained from intact images and cropped images with the number of parameters for each model. in Table 1.

Table 1. HCC Model Table

HCC model	Image Form	Number of Pieces	Number of HCC Feature Parameters
HCC Model A	Citra Utuh	1	8
HCC Model B	Citra potong 2x2	4	32
HCC Model C	Citra potong 3x3	9	72
HCC Model D	Citra potong 4x4	16	128

The MLP stage consists of 2 stages, namely 1) Training, and 2) Testing

a. MLP Training

The training phase uses 80 image data for each letter so that the total images used for the training stage are 1600 image data files

b. MLP testing

The testing phase uses 20 image data for each letter so that the total images used for the training stage are 400 image data files.

The test stages produce accuracy values by comparing the results of the test image classification with the correct classification data.

The overall accuracy of the model is known simply from the number of matched predictions divided by the total number of predictions made [11], in the form of the formula as follows:

$$Accuracy\ (\%) = \frac{Correct\ Prediction}{Number\ of\ Predictions} x\ 100\%$$

Results And Discussion

This study produced some data obtained by conducting 4 repetitions of the research stages, namely the stages of feature extraction and classification. This repetition is in accordance with the design of the 4 HCC models so that it will produce different data.

This stage is carried out by extracting the HCC features from the intact image so as to obtain a number of 8 HCC feature parameters, then the training and testing process is carried out using 1600 training image data and 400 test image data respectively.

Table 2. Model A HCC Test Results

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	LA	NA	JA	RA	KA	DA	TA	DA	WA	TA	PA	DHA	JA	YA	BA	MA	GA	NYA	THA	KA
xx002	YA	NYA	CA	RA	LA	DA	TA	DA	WA	TA	PA	DHA	JA	YA	HA	JA	JA	BA	TA	CA
xx003	LA	NYA	THA	RA	KA	DA	TA	SA	WA	YA	NA	CA	MA	YA	HA	JA	PA	BA	CA	NGA
xx004	LA	SA	DHA	RA	KA	DA	TA	SA	CA	SA	NYA	WA	WA	YA	NYA	CA	DA	NYA	JA	NGA
xx005	TA	NA	CA	RA	KA	DA	TA	DHA	CA	HA	NA	CA	JA	TA	HA	MA	GA	NYA	THA	NGA
xx006	LA	SA	JA	RA	HA	SA	SA	PA	JA	LA	PA	DHA	JA	YA	NGA	MA	GA	PA	NGA	NGA
xx007	HA	NYA	MA	RA	SA	DA	KA	KA	JA	LA	NYA	DHA	JA	YA	NYA	MA	GA	NGA	THA	NGA
xx008	HA	DA	BA	RA	KA	PA	HA	SA	WA	HA	NA	HA	JA	YA	NGA	WA	GA	BA	THA	RA
xx009	NYA	PA	CA	RA	KA	NA	TA	KA	WA	TA	PA	HA	JA	YA	NGA	WA	GA	NGA	THA	NGA
xx010	HA	CA	CA	RA	LA	DA	TA	SA	WA	KA	PA	HA	JA	YA	NYA	MA	GA	NGA	THA	NGA
xx011	SA	NA	WA	RA	TA	PA	HA	NA	WA	HA	PA	CA	NA	YA	NYA	MA	GA	NA	THA	THA
xx012	HA	PA	CA	RA	TA	DA	HA	WA	SA	KA	DA	WA	JA	YA	NYA	MA	GA	BA	THA	WA
xx013	LA	PA	RA	RA	KA	DA	TA	SA	NA	LA	PA	DHA	MA	TA	NYA	MA	GA	HA	THA	NGA
xx014	HA	NA	CA	RA	LA	DA	TA	SA	WA	HA	PA	DHA	JA	YA	BA	MA	GA	BA	THA	NGA
xx015	YA	WA	CA	RA	KA	DA	TA	DHA	WA	HA	PA	DHA	JA	YA	NYA	WA	GA	NA	THA	NGA
xx016	LA	NA	CA	RA	DHA	DA	TA	SA	SA	HA	PA	WA	JA	YA	BA	MA	GA	BA	THA	KA
xx017	LA	MA	DHA	RA	KA	DA	TA	KA	JA	YA	SA	DHA	JA	YA	PA	MA	GA	SA	THA	DHA
xx018	HA	NA	MA	RA	LA	SA	TA	PA	JA	HA	PA	DHA	MA	YA	GA	MA	GA	BA	THA	NGA
xx019	LA	SA	CA	RA	KA	NA	YA	PA	PA	BA	PA	NA	MA	TA	SA	MA	GA	PA	THA	NGA
xx020	TA	NA	CA	RA	THA	DA	CA	SA	WA	YA	PA	WA	JA	YA	NYA	MA	GA	NYA	THA	NGA

From Table 2 the results of the HCC Model A test are transformed into a confusion matrix making it easier to find the value of classification accuracy in testing with the HCC Model A method. The confusion matrix is shown in Table 3.

Table 3. Confusion Matrix HCC Model A

		RIIL																			
	Huruf	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
PREDIKSI	HA	6	0	0	0	1	0	3	0	0	7	0	3	0	0	3	0	0	1	0	0
	NA	0	7	0	0	0	2	0	1	1	0	3	1	1	0	0	0	0	2	0	0
	CA	0	1	10	0	0	0	1	0	2	0	0	3	0	0	0	1	0	0	1	1
	RA	0	0	1	20	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	KA	0	0	0	0	10	0	1	3	0	2	0	0	0	0	0	0	0	0	0	2
	DA	0	1	0	0	0	14	0	2	0	0	1	0	0	0	0	0	1	0	0	0
	TA	2	0	0	0	2	0	13	0	0	3	0	0	0	3	0	0	0	0	1	0
	SA	1	3	0	0	1	2	1	8	2	1	1	0	0	0	1	0	0	1	0	0
	WA	0	1	1	0	0	0	0	1	10	0	0	4	1	0	0	3	0	0	0	1
	LA	8	0	0	0	4	0	0	0	0	3	0	0	0	0	0	0	0	0	0	0
	PA	0	3	0	0	0	2	0	3	1	0	13	0	0	0	1	0	1	2	0	0
	DHA	0	0	2	0	1	0	0	2	0	0	0	9	0	0	0	0	0	0	0	1
	JA	0	0	2	0	0	0	0	0	4	0	0	0	14	0	0	2	1	0	1	0
	YA	2	0	0	0	0	0	1	0	0	3	0	0	0	17	0	0	0	0	0	0
	NYA	1	3	0	0	0	0	0	0	0	0	2	0	0	0	8	0	0	4	0	0
	MA	0	1	2	0	0	0	0	0	0	0	0	0	4	0	0	14	0	0	0	0
	GA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	1	0	17	0	0	0
	BA	0	0	1	0	0	0	0	0	0	1	0	0	0	0	3	0	0	7	0	0
	THA	0	0	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0	16	1
	NGA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	0	0	3	1	13

Based on the confusion table, accuracy will be obtained with the following calculations:

$$Accuracy\ (\%) = \frac{6+7+10+20+10+14+13+8+10+3+13+9+14+17+8+14+17+7+16+13}{400} x\ 100\%$$

$$Accuracy\ (\%) = \frac{229}{400} x\ 100\% = 57\%$$

From the calculation, the accuracy is obtained by 57%.

The next stage is to repeat the process of extracting the HCC features from the intact image which is divided into 4 parts so as to obtain a total of 32 HCC feature parameters, then the training and testing process is carried out again with the same data, namely 1600 training image data and 400 test image data. The test results are presented in Table 4.

From the table of HCC Model B test results it is transformed into a confusion matrix making it easier to find the classification accuracy value in testing with the HCC Model B method. The confusion matrix for Model B is presented in Table 5.

Table 4. Model B HCC Test Results

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	BA	PA	CA	RA	KA	DA	TA	SA	WA	LA	SA	HA	JA	YA	NYA	MA	WA	NGA	THA	NGA
xx002	BA	THA	CA	CA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	NYA	NYA	MA	GA	HA	THA	NGA
xx003	HA	KA	CA	RA	KA	DA	TA	SA	CA	LA	PA	HA	JA	NYA	BA	MA	THA	BA	THA	NGA
xx004	KA	NA	WA	RA	KA	DA	TA	SA	WA	LA	PA	CA	LA	YA	NYA	MA	HA	BA	THA	NGA
xx005	HA	KA	CA	RA	KA	DA	TA	SA	WA	SA	THA	DHA	JA	NYA	NYA	MA	GA	BA	THA	NGA
xx006	HA	KA	RA	RA	KA	DA	JA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	HA	NGA	NGA
xx007	HA	NA	CA	CA	BA	DA	TA	SA	CA	LA	PA	DHA	JA	YA	NYA	MA	GA	NYA	THA	NGA
xx008	HA	KA	CA	RA	NYA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx009	HA	NA	CA	RA	NYA	PA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	HA	THA	NGA
xx010	TA	NA	CA	GA	KA	PA	TA	NA	WA	LA	DA	DHA	JA	YA	JA	MA	GA	HA	MA	MA
xx011	HA	NA	DHA	RA	KA	DA	DA	SA	PA	LA	THA	DHA	JA	LA	NYA	MA	GA	DHA	THA	NGA
xx012	HA	NA	CA	RA	KA	PA	TA	SA	WA	LA	THA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx013	BA	NA	CA	RA	KA	DA	BA	DA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	PA	NGA
xx014	HA	NA	CA	RA	LA	DA	TA	SA	WA	LA	PA	DHA	JA	TA	NYA	MA	GA	BA	THA	NGA
xx015	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	WA	JA	YA	NYA	MA	GA	BA	THA	CA
xx016	JA	NA	CA	RA	NA	DA	HA	NGA	PA	LA	PA	DHA	JA	YA	TA	MA	GA	BA	THA	NGA
xx017	JA	NA	CA	RA	TA	DA	CA	SA	WA	PA	PA	DHA	JA	YA	HA	MA	RA	MA	THA	NGA
xx018	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	THA	DHA	JA	YA	TA	MA	GA	BA	NGA	NGA
xx019	HA	YA	CA	RA	KA	NA	KA	SA	WA	DA	PA	DHA	JA	YA	NGA	MA	RA	BA	THA	NGA
xx020	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	HA	NYA	MA	GA	BA	PA	NGA

Table 5. Confusion Matrix Model B

		RIIL																			
	Huruf	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
PREDIKSI	HA	13	0	0	0	0	0	1	0	0	0	0	2	0	1	1	0	1	4	0	0
	NA	0	13	0	0	1	1	0	1	0	0	0	0	0	0	0	0	0	0	0	0
	CA	0	0	17	2	0	0	1	0	2	0	0	1	0	0	0	0	0	0	0	1
	RA	0	0	1	17	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	0
	KA	1	4	0	0	14	0	1	0	0	0	0	0	0	0	0	0	0	0	0	0
	DA	0	0	0	0	0	16	1	1	0	1	1	0	0	0	0	0	0	0	0	0
	TA	1	0	0	0	1	0	14	0	0	0	0	0	0	1	2	0	0	0	0	0
	SA	0	0	0	0	0	0	0	17	0	1	1	0	0	0	0	0	0	0	0	0
	WA	0	0	1	0	0	0	0	0	16	0	0	1	0	0	0	0	1	0	0	0
	LA	0	0	0	0	1	0	0	0	0	17	0	0	1	1	0	0	0	0	0	0
	PA	0	1	0	0	0	3	0	0	2	1	14	0	0	0	0	0	0	0	2	0
	DHA	0	0	1	0	0	0	0	0	0	0	0	16	0	0	0	0	0	1	0	0
	JA	2	0	0	0	0	0	1	0	0	0	0	0	19	0	1	0	0	0	0	0
	YA	0	1	0	0	0	0	0	0	0	0	0	0	0	14	0	0	0	0	0	0
	NYA	0	0	0	0	2	0	0	0	0	0	0	0	0	3	14	0	0	1	0	0
	MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	20	0	1	1	1
	GA	0	0	0	1	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0
	BA	3	0	0	0	1	0	1	0	0	0	0	0	0	0	1	0	0	12	0	0
	THA	0	1	0	0	0	0	0	0	0	0	4	0	0	0	0	0	1	0	15	0
	NGA	0	0	0	0	0	0	0	1	0	0	0	0	0	0	1	0	0	1	2	18

Confusion matrix Model B is used to calculate its accuracy. Accuracy is calculated in the manner shown below:

$$Accuracy\ (\%) = \frac{13+13+17+17+14+16+14+17+16+17+14+16+19+14+14+20+15+12+15+18}{400} x\ 100\%$$

$$Accuracy\ (\%) = \frac{311}{400} x\ 100\% = 78\%$$

From the calculation results, the accuracy is 78%.

The next stage is to repeat the HCC feature extraction process from the intact image which is divided into 9 parts to obtain a total of 72 HCC feature parameters. From the HCC feature parameters, a training and testing process was carried out again with the same data, namely 1600 image data to train and 400 image data to test it.

Table 6. Table of Model C test results

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	LA	DHA	JA	YA	NYA	MA	GA	HA	THA	NGA
xx002	BA	NA	CA	RA	KA	DA	TA	SA	WA	LA	HA	DHA	JA	YA	NYA	MA	GA	HA	THA	THA
xx003	BA	DA	CA	RA	KA	DA	TA	PA	WA	LA	PA	DHA	JA	NYA	NYA	MA	GA	BA	THA	NGA
xx004	DHA	NA	CA	RA	KA	JA	TA	SA	NGA	LA	PA	DHA	JA	YA	NYA	SA	GA	BA	THA	NGA
xx005	TA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	THA	YA	NYA	PA	GA	HA	THA	NGA
xx006	NYA	NA	CA	RA	KA	RA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	RA	BA	THA	NGA
xx007	HA	NA	LA	RA	NA	DA	HA	SA	WA	LA	HA	DHA	JA	YA	NYA	MA	GA	NYA	THA	NGA
xx008	HA	DA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	DHA	GA	BA	NGA	NGA
xx009	HA	NA	CA	RA	KA	SA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	TA	GA	BA	THA	NGA
xx010	PA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx011	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	KA	NYA	MA	RA	BA	THA	NGA
xx012	HA	NA	CA	RA	KA	DA	PA	SA	NGA	LA	THA	LA	JA	YA	NYA	MA	GA	HA	THA	GA
xx013	HA	NA	SA	RA	KA	DA	TA	SA	JA	LA	PA	DHA	JA	GA	NYA	MA	GA	BA	THA	PA
xx014	SA	NA	CA	RA	KA	DA	TA	SA	THA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx015	HA	NA	SA	RA	KA	NA	TA	SA	TA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx016	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	THA	DHA	JA	YA	NYA	MA	GA	BA	NGA	NA
xx017	HA	NA	CA	RA	KA	DA	TA	SA	WA	BA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	PA
xx018	HA	NA	NA	RA	TA	DA	TA	SA	WA	GA	PA	DHA	JA	NYA	BA	MA	RA	BA	THA	THA
xx019	HA	NA	CA	RA	KA	THA	KA	SA	WA	BA	JA	DHA	JA	NYA	NYA	MA	GA	BA	THA	NGA
xx020	HA	NA	CA	RA	KA	NA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA

The results presented in Table 6 are then transformed into a confusion matrix making it easier to find the classification accuracy value in testing with the HCC Model C Method. The confusion matrix resulting from this transformation is presented in Table 7.

Table 7. Confusion Matrix Model C

		RIIL																			
	Huruf	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
PREDIKSI	HA	13	0	0	0	0	0	2	0	0	0	2	0	0	0	0	0	0	4	0	0
	NA	0	18	1	0	1	2	0	0	0	0	0	0	0	0	0	0	0	0	0	1
	CA	0	0	16	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	RA	0	0	0	20	0	1	0	0	0	0	0	0	0	0	0	0	3	0	0	0
	KA	0	0	0	0	18	0	1	0	0	0	0	0	0	1	0	0	0	0	0	0
	DA	0	2	0	0	0	14	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	TA	1	0	0	0	1	0	16	0	1	0	0	0	0	0	0	1	0	0	0	0
	SA	1	0	2	0	0	1	0	19	0	0	0	0	0	0	0	1	0	0	0	0
	WA	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0	0	0	0	0	0
	LA	0	0	1	0	0	0	0	0	0	17	1	1	0	0	0	0	0	0	0	0
	PA	1	0	0	0	0	0	1	1	0	0	14	0	0	0	0	1	0	0	0	2
	DHA	1	0	0	0	0	0	0	0	0	0	0	19	0	0	0	1	0	0	0	0
	JA	0	0	0	0	0	1	0	0	1	0	1	0	19	0	0	0	0	0	0	0
	YA	0	0	0	0	0	0	0	0	0	0	0	0	0	15	0	0	0	0	0	0
	NYA	1	0	0	0	0	0	0	0	0	0	0	0	0	3	19	0	0	1	0	0
	MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	16	0	0	0	0
	GA	0	0	0	0	0	0	0	0	0	1	0	0	0	1	0	0	17	0	0	1
	BA	2	0	0	0	0	0	0	0	0	2	0	0	0	0	1	0	0	15	0	0
	THA	0	0	0	0	0	1	0	0	1	0	2	0	1	0	0	0	0	0	18	3
	NGA	0	0	0	0	0	0	0	0	2	0	0	0	0	0	0	0	0	0	2	13

The accuracy of the transformation whose result is the Confusion Matrix presented in Table 7 can be calculated from the following formula:

$$Accuracy\ (\%) = \frac{\begin{matrix}13+18+16+20+18+14+16+19+15+17+14+19+\\19+15+19+16+17+15+18+13\end{matrix}}{400} x\ 100\%$$

$$Accuracy\ (\%) = \frac{331}{400} x\ 100\% = 83\%$$

From the calculation results, the accuracy is 83%.

The next step is to repeat the HCC feature extraction process from the intact image which has been divided into 16 sections to obtain 128 HCC feature parameters. From the HCC feature parameters, a training and testing process was carried out again with the same data, namely 1600 image data to train and 400 image data to test it.

The test results above are presented in the following tabulation Table 8:

Table 8. Table of Model D test results

FileNo	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx001	HA	NA	CA	RA	HA	DA	TA	PA	WA	SA	WA	DHA	NA	YA	NYA	MA	GA	BA	THA	THA
xx002	HA	NA	CA	RA	KA	DA	KA	PA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx003	HA	NA	CA	RA	KA	DA	TA	SA	CA	LA	PA	NA	JA	YA	NYA	MA	GA	BA	THA	NGA
xx004	HA	NA	CA	RA	KA	NA	TA	SA	GA	LA	SA	WA	JA	YA	NYA	MA	HA	BA	THA	NGA
xx005	KA	NA	CA	RA	KA	KA	TA	PA	DHA	LA	GA	DHA	JA	YA	NYA	THA	GA	BA	THA	BA
xx006	YA	NA	CA	RA	TA	NA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx007	HA	NA	CA	RA	KA	NA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	THA	GA	BA	THA	THA
xx008	HA	NA	DHA	RA	HA	SA	HA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	THA
xx009	HA	NA	NA	RA	BA	DA	TA	SA	WA	LA	DA	CA	JA	YA	NYA	MA	GA	BA	NGA	THA
xx010	HA	NA	CA	RA	TA	DA	TA	SA	CA	BA	DA	CA	JA	YA	NYA	MA	BA	BA	NGA	NGA
xx011	HA	NA	WA	RA	TA	PA	TA	SA	WA	LA	DA	WA	JA	YA	NYA	MA	GA	GA	NGA	NGA
xx012	HA	NA	CA	RA	KA	DA	JA	SA	DHA	LA	PA	CA	JA	YA	NYA	MA	GA	BA	THA	THA
xx013	HA	NA	CA	RA	TA	DA	TA	PA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	NYA	THA	NGA
xx014	HA	NA	WA	RA	NA	NA	KA	SA	NA	LA	PA	DHA	JA	HA	NYA	MA	GA	BA	THA	THA
xx015	HA	NA	CA	RA	KA	NA	TA	SA	CA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	TA
xx016	HA	NA	CA	GA	KA	DA	TA	SA	WA	LA	DA	DHA	JA	YA	NYA	MA	GA	HA	THA	NGA
xx017	PA	NA	WA	RA	HA	DA	WA	SA	WA	LA	DA	DHA	JA	YA	YA	MA	GA	BA	THA	THA
xx018	HA	NA	CA	RA	TA	DA	TA	SA	WA	LA	PA	CA	KA	YA	NYA	MA	GA	BA	THA	THA
xx019	YA	NA	DHA	RA	KA	DA	TA	SA	WA	LA	GA	DHA	HA	YA	NYA	MA	GA	BA	THA	NGA
xx020	LA	NA	DHA	RA	YA	NA	TA	SA	WA	LA	PA	DHA	JA	HA	NYA	MA	GA	BA	THA	NGA

The results presented in Table 8 are then transformed into a confusion matrix to make it easier to find the classification accuracy value in testing with the HCC Model D Method. The confusion matrix resulting from this transformation is presented in Table 9.

Table 9. Confusion Matrix Model D

		RIIL																			
	Huruf	HA	NA	CA	RA	KA	DA	TA	SA	WA	LA	PA	DHA	JA	YA	NYA	MA	GA	BA	THA	NGA
PREDIKSI	HA	15	0	0	0	3	0	3	0	0	0	0	0	1	2	0	0	1	1	0	0
	NA	0	20	1	0	1	6	0	0	1	0	0	1	1	0	0	0	0	0	0	0
	CA	0	0	13	0	0	0	0	0	3	0	0	4	0	0	0	0	0	0	0	0
	RA	0	0	0	19	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0
	KA	1	0	0	0	9	1	2	0	0	0	0	0	1	0	0	0	0	0	0	0
	DA	0	0	0	0	0	11	0	0	0	0	5	0	0	0	0	0	0	0	0	0
	TA	0	0	0	0	5	0	13	0	0	0	0	0	0	0	0	0	0	0	0	1
	SA	0	0	0	0	0	1	0	16	0	1	1	0	0	0	0	0	0	0	0	0
	WA	0	0	3	0	0	0	1	0	13	0	1	2	0	0	0	0	0	0	0	0
	LA	1	0	0	0	0	0	0	0	0	18	0	0	0	0	0	0	0	0	0	0
	PA	1	0	0	0	0	1	0	4	0	0	11	0	0	0	0	0	0	0	0	0
	DHA	0	0	3	0	0	0	0	0	2	0	0	13	0	0	0	0	0	0	0	0
	JA	0	0	0	0	0	0	1	0	0	0	0	0	17	0	0	0	0	0	0	0
	YA	2	0	0	0	1	0	0	0	0	0	0	0	0	18	1	0	0	0	0	0
	NYA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	19	0	0	1	0	0
	MA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	18	0	0	0	0
	GA	0	0	0	1	0	0	0	0	1	0	2	0	0	0	0	0	18	1	0	0
	BA	0	0	0	0	1	0	0	0	0	1	0	0	0	0	0	0	1	17	0	1
	THA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	2	0	0	17	10
	NGA	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	0	3	8

The accuracy of the transformation whose result is the Confusion Matrix presented in Table 9 can be calculated from the following formula:

$$Accuracy\ (\%) = \frac{\begin{matrix}15+20+13+19+9+11+13+16+13+18+11+13+\\17+18+19+18+18+17+17+8\end{matrix}}{400} x\ 100\%$$

$$Accuracy\ (\%) = \frac{330}{400} x\ 100\% = 76\%$$

From the calculation results, the accuracy is 76%.

This study has extracted 4 models of HCC features. The recapitulation of the four models is presented in Table 10 below:

Table 10. Accuracy of the HCC Model Feature Extraction Test

HCC Feature Extraction Model	Number of Chain Code Segments	Number of HCC Feature Parameters	accuracy
A	1	8	57%

B	4	32	78%
C	9	72	83%
D	16	128	76%

Table 10 shows the lowest accuracy obtained from Model A extraction, which is 57%. The highest accuracy was obtained from Model C extraction, namely 83%. Models with intact images, namely those that are not divided, produce the lowest accuracy value compared to images that are divided. However, making a larger number of images does not always increase the accuracy value. This can be seen from the division of the image into 16 parts which only produces an accuracy value of 76%. This result is smaller than the image which is divided into 9 parts with an accuracy value of 83%.

Conclusion

This research studies the application of the Histogram Chain Code (HCC) method as an image processing method for Old Javanese letters. The number of image parameters is increased by dividing the whole image into several parts with the number of divisions of 4, 9 and 16. The division of the image with the 3 divisions is compared with the intact image (not divided). The accuracy of the computer's ability to recognize these images was experimentally compared between the divided and the intact images. From the experiment, the HCC model that divides the image into 3x3 pieces or divides it into 9 parts, which obtains 72 feature parameters produces the highest accuracy compared to images that are not divided or divided into other numbers. The characteristics of the Javanese script which have special characteristics between one letter and another cause the division of the image into several parts to have the best number of image divisions compared to the division with other numbers. This study also found that dividing an image into more parts does not always increase the level of accuracy compared to those with fewer divisions. Javanese script letters that have unique shapes such as Ra, Sa, La, Ja, and Nya have a good level of accuracy compared to scripts that look similar to one another. This requires improved methods that can capture the special features of the Javanese script to improve its accuracy. This study also suggests that future research should compare other chain code methods, for example Differential Chain Code and Vertext Chain Code.

References

[1] Rais, N. S. R., Dien, M. M. J., & Dien, A. Y. (2018). Kemajuan teknologi informasi berdampak pada generalisasi unsur sosial budaya bagi generasi milenial. Jurnal Mozaik, 10(2), 61-71.

[2] Pergub, D. I. Y. (2013). Peraturan Gubernur Daerah Istimewa Yogyakarta Nomor 64 Tahun 2013 tentang Mata Pelajaran Bahasa Jawa Sebagai Muatan Lokal Wajib di Sekolah/Madrasah. Yogyakarta: Pergub DIY.

[3] Aldian, D. (2022). Pengembangan Media Pembelajaran Interaktif Aksara Jawa Berbasis Flutter Pada Mata Pelajaran Bahasa Jawa (Doctoral dissertation, Universitas Duta Bangsa Surakarta).

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

[5] Freeman, H., & Davis, L. S. (1977). A corner-finding algorithm for chain-coded curves. IEEE Transactions on computers, 26(03), 297-303.

[6] Mawaddah, S., & Suciati, N. (2020). Pengenalan Karakter Tulisan Tangan Menggunakan Ekstraksi Fitur Bentuk Berbasis Chain Code. Jurnal Teknologi Informasi dan Ilmu Komputer, 7(4), 683-692.

[7] Salem, A. B. M., Sewisy, A. A., & Elyan, U. A. (2005). A vertex chain code approach for image recognition. ICGST International Journal on Graphics, vision and Image processing, 5(3), 1-8.

[8] Hadi, F. A., Wirayuda, T. A. B., & Ramadhani, K. N. (2015). Pengenalan

Angka Tulisan Tangan Dengan Penerapan Freeman Chain Code yang Dimodifikasi. eProceedings of Engineering, 2(2).

[9] Dewi, G. L., & Armanto, H. (2015). Analisa berbagai jenis huruf komputer menggunakan algoritma berbasis chain code dalam bentuk run length encoding. In Seminar Nasional Inovasi dalam Desain dan Teknologi Ideatech (pp. 2089-1121).

[10] Prima, G. R. (2021). Analisa Perbandingan Nilai K Terbaik Untuk Clustering K-Means Menggunakan Pendekatan Elbow dan Silhouette Pada Citra Aksara Jawa. Universitas Sanata Dharma..

[11] Hastuti, K. (2012). Analisis komparasi algoritma klasifikasi data mining untuk prediksi mahasiswa non aktif. Semantik, 2(1).