Accredited Ranking SINTA 2
Decree of the Director General of Higher Education, Research, and Technology, No. 158/E/KPT/2021
Validity period from Volume 5 Number 2 of 2021 to Volume 10 Number 1 of 2026

Published online on: **http://jurnal.iaii.or.id**

JURNAL RESTI
(Rekayasa Sistem dan Teknologi Informasi)
Vol. 7 No. 1 (2023) 193 - 200 ISSN Media Electronic: 2580-0760

Character Recognition of Handwriting of Javanese Character Image using Information Gain Based on the Comparison of Classification Method

Irham Ferdiansyah Katili[1], Mochamad Arief Soeleman[2], Ricardus Anggi Pramunendar[3]
[1,2,3] Program Studi Magister Teknik Informatika, Fakultas Ilmu Komputer, Universitas Dian Nuswantoro Semarang
[1]irhamferdi759@gmail.com, [2]m.arief.soeleman@dsn.dinus.ac.id, [3]ricardus.anggi@dsn.dinus.ac.id

Abstract

Indonesia is a country rich in a variety of regional cultures. Regional airspace needs to be preserved so as not to become extinct. One of them is the local culture of Central Java Province, namely Javanese Character. In this modern era, globalization is growing in every country. The impact of globalization is increasingly widespread and developing in society. One effect of globalization is local people prefer foreign language skills to learn local languages. This study, applies the method of character recognition using a new combination workflow that contains Local Binary Pattern (LBP) and Information Gain. Then compare Support Vector Machine (SVM), k-Nearest Neighbor and Naïve Bayes. The LBP method is used to obtain an image's texture or shape characteristics. Information Gain is used for the feature selection algorithm, whereas SVM, k-Nearest Neighbor and Naïve ayes is used for the classification method. From previous research, the information gain method succeeded in increasing the accuracy by 2%. This research compares the SVM classification with another classification method, and the result shows that our proposed can improve classification performance. The best accuracy result using SVM classification gets 87,86%, at ten folds and cell size 64x64.

Keywords: *character recognition; javanese character; information gain; LBP*

1. Introduction

Indonesia is a country rich in regional culture. Regional culture needs to be preserved so that its existence is not extinct. Local languages are one of the regional cultures that need to be kept. The impact of globalization that is continuously developing makes people prefer to learn and develop other language skills compared to learning local languages. One of the well-known regional cultures of Central Java Province is Javanese Character. Javanese Character has a historical story originating from the land of Java [1].

Along with the changing times, the globalization influence is growing and entering most countries including Indonesia. The impact resulting from globalization is the loss of regional culture and being replaced by foreign cultures. The rapid growth of globalization has caused society to be unable to filter its effects. That way it can quickly eliminate the local culture. Globalization also has a positive influence, one of which is information technology [2]. That way, the preservation of regional culture, especially Javanese Characters can be done. The preservation was carried out by way of introducing Javanese Character characters. Character recognition is a branch of the field of Digital Image Processing. In digital image processing, character recognition is also called Optical Character Recognition [3]. In this research, pattern recognition is applied to handwritten images of Javanese characters.

Several research have been conducted [4]- [8]. Then, in research [8], the local binary pattern feature extraction method produces a large number of features so that feature simplification is required. In this research, the author will conduct experiments by applying the LBP method to extract texture features. The LBP method is a suitable method of doing texture analysis [9], [10]. The Local Binary Pattern method works by creating a new binary pattern derived from comparing the median pixel value with the neighbor pixel value [11]. LBP produces a large number of features so it needs to be simplified. Information feature selection is used to simplify irrelevant features. This study applies the Information Gain method as a feature selection method. Information Gain works based on the ranking of the attributes [12].

Research on Information Gain in Image has been carried out, among others. Conducted by Adel Hamdan et al [13], about Comparing Information Gain and Gain

Ratio on three classifications. Result show the feature selection methods done in the pre-processing stage, and by implementing the gain ratio method, the average level of precision is 0.88 for the Naive Bayes method, 0.71 for the C4.5 method and 0.89 for the SVM method. For the final result, the SVM method gives a little bit better than naïve Bayes and C4.5 results with gain ratio.

Subsequent research was carried out by Asma Elmaizi et al [14], about classification and dimensional reduction using the Information Gain method. The method based on hyperspectral bands selection is adopted for picking the most informative bands and deleting the irrelevant and noisy ones. The research compares with the MRMR, MIFS and MIFB, and the result is IGBS achieves 96.83% accuracy with 80 bands, which is higher than MRMR by 2% and higher more than 1.5% from MIFS and MIFB. The conclusion, the Information Gain method outperforms the reproduced method.

Research conducted by Xinzheng Wang et al [15], about a novel feature selection method using Feature Set Equivalence and Mutual Information Gain Maximization. The proposed method in this research, MIGM achieves the highest accuracy with two different classifiers. The overall accuracy of MIGM over ten datasets is 75.37%. Meanwhile, the MIGM method obtains the highest average accuracy on seven out of ten data sets. The conclusion, sparse autoencoder, and stacked sparse autoencoder can work properly.

Research conducted by Ajib Susanto et al [8], about the Character Recognition of Javanese Characters using K-Nearest Neighbor method based on Local Binary Pattern. This research uses a Local Binary Pattern as a feature extraction method and a k-Nearest Neighbor as a classification method. The testing was carried out using 160 image datasets. The test results show that the combination of cell size parameters measuring 64x64 and parameter k=3 gives the highest accuracy of 82,5%

Another research conducted by Wibowo et al [16], about Javanese character feature extraction based on Shape Energy. This research uses the Shape Energy method for feature extraction, which is one of the methods with the basic idea of how the character can be distinguished simply through its skeleton. The testing was carried out using 387 data with 19 labels by using K-Nearest Neighbor as a classification method. The result, generate through the Cross-Validation get an accuracy value of 81.90% with an angel of 20 degrees.

With the background and several journals that have been discussed, especially in journals [8]which has a problem with the amount of binary data resulting from the feature extraction stage. So, the main contribution of this paper is adding a feature selection method using information gain, the result of feature extraction can be simplified and then can affect the performance of the classification. Then, the authors compare three algorithms Support Vector Machine algorithm, K-Nearest Neighbors algorithm and Naive Bayes algorithm for classification.

2. Research Methods

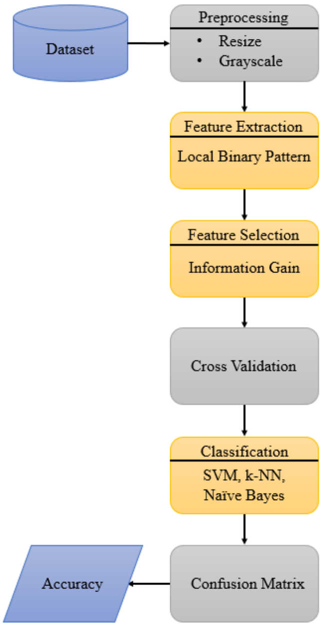

Figure 1. Flow Chart of Research Method

2.1 Dataset

The dataset used in this research uses handwritten Javanese character as many as 160 records, which is divided into 120 training data and 40 test data. The data is the same as previous research [8]

2.2 Preprocessing

From Figure 1. it is known that the first step in testing is to preprocess the dataset used. At the preprocessing stage using the Matlab application. There are 2 processes carried out at the preprocessing, resize and grayscale. Resize, the image dataset will equalize the image size of all images, 640 x 480 pixels. Grayscale, then the image color will be changed to grayish as shown in Figure 2.

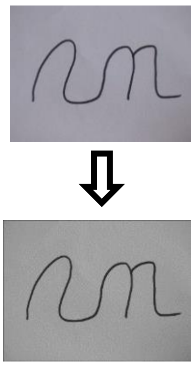

Figure 2. RGB image to Grayscale Image

DOI: https://doi.org/10.29207/resti.v7i1.4488

After that, the binary numbers from the grayscale image will be used for the next step.

2.3 Feature Extraction

After get the preprocessing results, it will proceed to the feature extraction stage. At the feature extraction stage, the proposed method uses a Local Binary Pattern. The researcher will use three cell sizes at this stage, namely 32x32, 64x64, and 128x128. However, this method is sensitive to image noise and cannot capture macro-structural information.

This method works by calculating one pixel's value with neighboring pixels (a_n) divided into a_0 to a_7 shown in Figure 3. The values resulting from pixel calculation are then compared and get a new binary pattern. These binary patterns are features that can represent pixels.

a_0	a_1	a_2
a_7	a_x	a_3
a_6	a_5	a_4

⇩

170	206	190
231	195	215
192	189	223

Figure 3. Determination of neighbor pixel value

In Figure 4 after getting the neighbor pixel value, the next step is the image threshold with a limit value equal to the a_x value. Then, give each Neighbor a weight with a weighted value of 2^n.

0	1	0
1		1
0	1	1

⇩

1	2	4
128		8
64	32	16

Figure 4. LBP calculation on image-based the result of the threshold

The result of the threshold determines the value of LBP. If the value is 0, the weight value is 0, whereas if the value is 1 then the weight value is maximum. The result of the threshold :

0	1	0
1		1
0	1	1

Threshold = 01011101

The value of LBP :

1	2	4
128		8
64	32	16

LBP = 0 + 2 + 0 + 8 + 16 + 32 + 0 + 128 = 186

Form the calculation of LBP, then the result of LBP for the block image is 186. So, the value of a_x is 186.

2.4 Feature Selection

From the research [8], the LBP result has large feature. The main contribution of the research in this step. The results from feature extraction will be simplified at the feature selection stage. The method used at the feature selection stage is Information Gain Weighted.

Information gain is a feature selection algorithm, that works by ranking attributes and reduces noise, and it may cause irrelevant features [12]. Determination of the best attributes is a stage of Information Gain, aims to eliminate irrelevant features. The best attribute is done by calculating the entropy value. Entropy is an abstract measure in its simplest form. To calculate the entropy value, shown in the equation (1).

$$entropy(S) = \sum_{i}^{c} - P_i \log_2 P_i \qquad (1)$$

From equation (1), c is the number of classes in the dataset. While P_i is the number of the sample from class - i

After the entropy value is obtained, then calculate the weight gain value with equation (2)

$$gain(S, A) = entropy(S) - \sum_{Value(A)} \frac{|S_v|}{|S|} Entropy(S_v) \qquad (2)$$

From equation (2), S is the sample data value. A is an attribute in the dataset. $|S_v|$ is the number of samples of the importance of v, while v is the possible value of attribute A.

The information gain method works by removing features that are not relevant to the classification process. Features that meet the weighting criteria will be used in the next process.

FBT	FBU	FBV	FBW
0	0,9992	0,0841	0
0	0	0,995	0,135
0	0,9991	0,0035	0,4521

⇩

FBT	FBU	FBV	FBW
0	0	0,0841	0
0	0	0,995	0
0	0	0,0035	0

Figure 5. Feature Selection Information Gain applied in feature data

Figure 5, illustrates the reduction in the number in cells in the feature which is the result of information gain. In this process, the results from feature extraction will be simplified and will be weighted.

2.5 Classification

The value of the weights obtained will then be tested in the Cross-Validation process. The dataset will be divided into training data and random test data in this process. Then it will be classified using the SVM, k-NN and Naïve Bayes Method.

Support Vector Machine (SVM) is a learning algorithm that uses a hypothetical space in a linear function in a

DOI: https://doi.org/10.29207/resti.v7i1.4488

feature space with high dimensions. SVM is designed to minimize the error and maximize the separation margin among different classes.

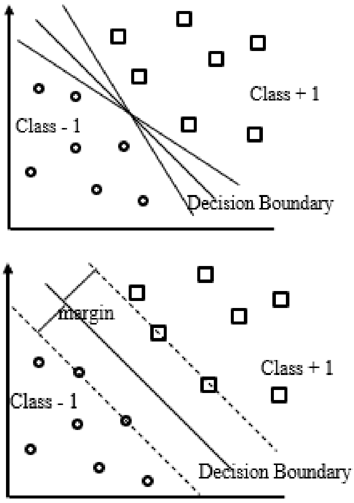

Figure 6. Hyperplane boundary separating two class

Figure 6, illustrates some of the data that are members of 2 classes, namely class -1 and class +1. Circle data goes to class -1, while square data goes to class +1. The best hyperplane between the two classes can be found by measuring the margins hyperplane and searching for its maximum point. Margin is the distance between the hyperplane with the closest data from each category. The data most relative to the hyperplane is called the support vector. The equation of SVM can be show in equation (3), (4), and (5).

$$f_{svm}(x) = w.x + b \tag{3}$$

Where,

$$w = \sum_{i \in N} \alpha_i y_i \Phi(x_i) \tag{4}$$

With

$$K(x_i, x_j) = \Phi(x_i)\Phi(x_j) \tag{5}$$

Where Kernel is function, and b is bias value.

The support Vector Machine algorithm is divided into two types: linear and non-linear SVM. In linear, the algorithm works on data that can be linearly separated. At the same time, the data that cannot be separated linearly can use non-linear SVM. On non-linear SVM using the kernel on the data feature, the kernel's function is mapping features from lower dimension to higher dimension. The equation of the kernel can be shown in equation (6) and (7).

$$\Phi = D^r \rightarrow D^q \tag{6}$$

$$x \rightarrow \Phi(x) \tag{7}$$

Where Φ is the kernel mapping, *D* is the data train, *r* is the feature data before mapping, and *q* is the feature set after mapping. While *x* is training data, x_1, x_2, and $x_n \in D^r$ is a mapping feature.

K-Nearest Neighbor (K-NN) is a simple classification algorithm that categorizes objects based on similar classes. This algorithm is at k, and if the majority of k-data are included in the feature space, then the new data will fall into that category [17]. When the data is evenly distributed, it will be challenging to determine the distance of the Neighbor because the distance between the testing data and training data is the same. The equation (8) is to calculate the euclidean distance.

$$d(u,v) = \sqrt{\sum_{i=1}^{N} \left|u_i - v_j\right|^2} \tag{8}$$

From the equation, *d* is euclidean distance, u_i is testing data *i*-th and v_j is training data *j*-th.

Naive Bayes is a probability algorithm whose principle is that an attribute value cannot influence other attributes. In conditional probability, the thing that needs to be considered in the Naive Bayes algorithm is the observed object features. This feature is intended to assess the conditional probability of an object entering each class.

This algorithm works well even when the independent assumptions are invalid [18].

The probability value in the Naive Bayes algorithm can be calculated by equation (9).

$$P(H|X) = \frac{p(H|X).p(H)}{p(X)} \tag{9}$$

From The equation, *P(H|X)* is the posterior probability *H* based on the condition of *X* obtained from probability *H(P(H))* multiplied by probability *X* based on the hypothesis condition *H(P(X|H)* and divided by the probability X(P(X)).

3. Results and Discussions

3.1 Pre-processing Result

This test uses RGB image seen in Figure 7 and Grayscale image seen in Figure 8 measuring 640x480 pixels.

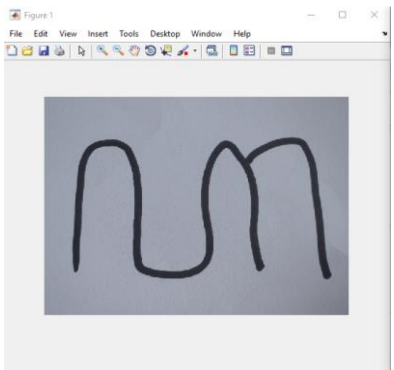

Figure 7. The sample image of RGB image

DOI: https://doi.org/10.29207/resti.v7i1.4488

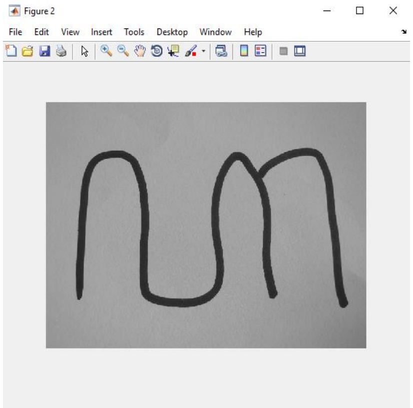

Figure 8. The sample image of grayscale image

3.2 Feature Extraction Result

In this research using Local binary Pattern for Feature Extraction step. The example results can be seen in Table 1, 2 and 3.

Table 1. The result of LBP Feature Extraction [32 32] pattern

A	B	C	D	…	ROP
0	0	0	0	…	0
4,00e-05	0	0,0012702	0	…	0
0	0	0	0	…	0,0003627
0	0	0	0	…	0
0	0	0	0	…	0

Table 2. The result of LBP Feature Extraction [64 64] pattern

A	B	C	D	…	ROP
0	0	0	0	…	0,008453
0	0	0	0	…	0,027778
0	0	0	0	…	0,003556
0	0	0	0	…	0
0	0	0	0	…	0,12445

Table 3. The result of LBP Feature Extraction [128 128] pattern

A	B	C	D	…	AHA
0	0	0	0	…	0,006524
0	0	0	0	…	4,03E-01
0	0	0	3,18E-01	…	0,0001550
0	0	0	0	…	0,0002665
0	0	0	0	…	0,0074961

3.3 Feature Selection Result

In this research uses Information Gain for feature selection step. The example results can be seen in Table 4, 5 and 6.

Table 4. The result of Information Gain with [32 32] cell size

att1	att2	att3	att4	…	att17700
0	0	0	0	…	0,0000
0	0	0	0	…	0,0425
0	0	0	0	…	0,1451
0	0	0	0	…	0,04545
0	0	0	0	…	0,0018284

Table 5. The result of Information Gain with [64 64] cell size

att1	att2	att3	att4	…	Att4130
0	0	0	0	…	0,00845
0	0	0	0	…	0,02778
0	0	0	0	…	0,00356
0	0	0	0	…	0,00000
0	0	0	0	…	0,12445

Table 6. The result of Information Gain with [128 128] cell size

att1	att2	att3	att4	…	Att885
0	0	0	0	…	0,00652
0	0	0	0	…	0,00004
0	0	0	0	…	0,00016
0	0	0	0	…	0,00027
0	0	0	0	…	0,007496

3.4 Classification Result

In previous studies [8], testing was carried out using the Local binary pattern algorithm as a feature extraction algorithm and k-Nearest Neighbor as a classification algorithm. The highest level of accuracy is obtained using a cell size of 64x64 and k = 3, with an accuracy rate of 82.5%.

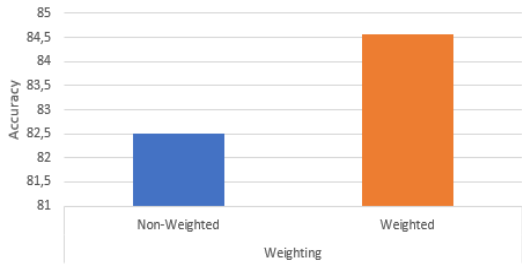

Figure 9. Comparison between with Information Gain and non-Information Gain testing in k-NN

Table 7. Comparison of accuracy and testing time

No	With or without Information Gain	Number of k	Accuracy	Test Time
1	Without Information Gain	3	82,5%	0,428 sec
2	With Information Gain	3	84,5%	0,223 sec

Figure 9. and Table 7, explained that in the first experiment implementing the Local Binary Patten method with Cell Size [64 64] and k-Nearest Neighbor with k = 3 without using the Information Gain weighting method and with applying Information Gain. Information gain weighting is done by 50% for all test data features. Data features that are less relevant to the test will be removed.

For the result, without using the Information Gain weighting method get average test time was 0.428 seconds and achieved an accuracy rate of 82.5 %. While the second experiment implements the Local Binary Pattern method with Cell Size [64 64] and then applies the weighting method with Information Gain and the same classification method as the study using k-Nearest

DOI: https://doi.org/10.29207/resti.v7i1.4488

Neighbor, obtaining a testing time of 0.223 seconds and achieving an accuracy level of 84, 5%.

The two tests concluded that the use or implementation of the Information Gain weighting method can increase the computational process by 0.2 seconds and increase the level of accuracy by 2%.

3.4.1 Comparison Classification Method with Information Gain

Table 8. Evaluation Result of Character Recognition with Information Gain

Cell Size	Classifi-cation	k-Folds Cross-Validation	Accuracy
[32 32]	k-NN, k=3	10	67,14 %
		8	65,36 %
		6	63,57 %
	k-NN, k=5	10	65,00 %
		8	63,93 %
		6	64,29 %
	k-NN, k=7	10	63,93 %
		8	62,50 %
		6	62,50 %
	Naive Bayes	10	73,93 %
		8	73,21 %
		6	74,64 %
	SVM (rbf)	10	57,14 %
		8	61,07 %
		6	55,71 %
	SVM (linear)	10	77,86 %
		8	77,14 %
		6	76,79 %
[64 64]	k-NN, k=3	10	84,57 %
		8	82,14 %
		6	78,50 %
	k-NN, k=5	10	76,43 %
		8	77,50 %
		6	75,36 %
	k-NN, k=7	10	73,93 %
		8	74,29 %
		6	72,86 %
	Naive Bayes	10	72,50 %
		8	74,64 %
		6	73,93 %
	SVM (rbf)	10	52,14 %
		8	57,50 %
		6	56,80 %
	SVM (linear)	10	87,86 %
		8	85,36 %
		6	86,43 %
[128 128]	k-NN, k=3	10	74,86 %
		8	76,07 %
		6	77,43 %
	k-NN, k=5	10	75,36 %
		8	73,21 %
		6	75,71 %
	k-NN, k=7	10	74,64 %
		8	73,93 %
		6	75,03 %
	Naive Bayes	10	67,86 %
		8	67,14 %
		6	66,97 %
	SVM (rbf)	10	58,57 %
		8	53,57 %
		6	55,00 %
	SVM (linear)	10	82,73 %
		8	78,57 %
		6	75,35 %

In Figure 9, the dataset was tested using feature extraction Local Binary Pattern and weighting has been done with the Information Gain algorithm. The value of the weighted results is classified by the classification algorithm. Then, this research compared three classification algorithms, namely k-Nearest Neighbor, Support Vector Machine and Naive Bayes.

In the next test, two other algorithms were compared Support Vector Machine and Naive Bayes. In Table 8, the results of testing by applying the cell size 32x32, 64x64, and 128x128. For the parameter of K-Nearest Neighbor, this research uses the same parameters used in research [8]. While the kernel of Support Vector Machine using a kernel that produces a high level of accuracy, according to the kernel used in research [19]. After the effects of the feature extraction, feature selection is carried out and weighted.

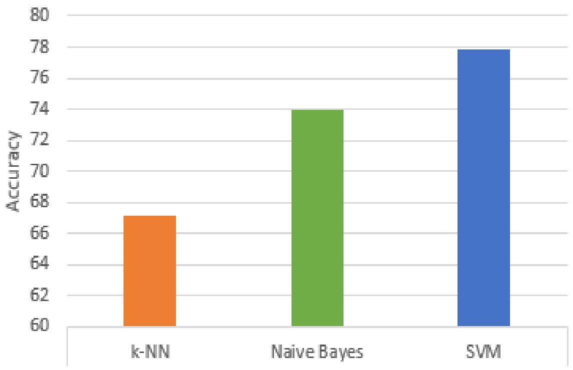

Figure 10. Comparison between k-NN, Naïve Bayes and SVM in [32 32] cell size with Information Gain

Figure 10, shows the comparison of the three classification algorithms with Information Gain. Uses 32x32 cell size in feature extraction. For k-NN, k=3 get an accuracy of 67,14 %, Naive Bayes get 73,93%, and Support Vector Machine get 77,86%.

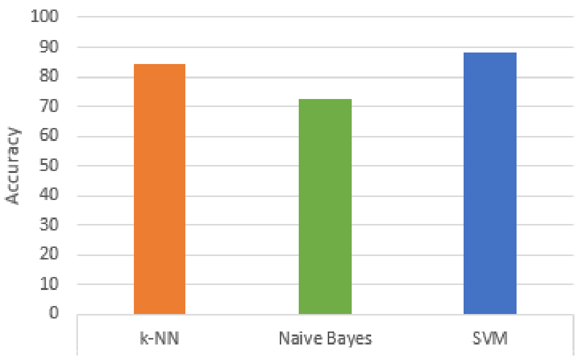

Figure 11. Comparison between k-NN, Naïve Bayes and SVM in [64 64] cell size with Information Gain

Figure. 11, shows the comparison of the three classification algorithms with Information Gain. Uses 64x64 cell size in feature extraction. For k-NN, k=3 get an accuracy of 84,57 %, Naive Bayes get 72,50 %, and Support Vector Machine get 87,86 %.

DOI: https://doi.org/10.29207/resti.v7i1.4488

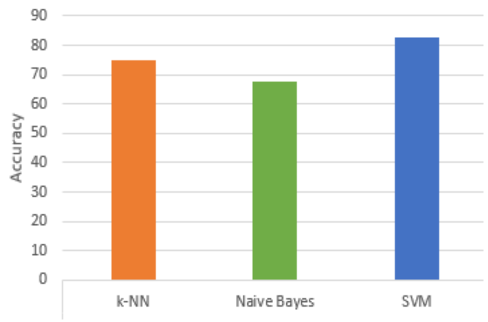

Figure 12. Comparison between k-NN, Naïve Bayes and SVM in [128 128] cell size with Information Gain

Figure 12, shows the comparison of the three classification algorithms with Information Gain. Uses 128x128 cell size in feature extraction. For k-NN, k=3 get an accuracy of 74,86 %, Naive Bayes get 67,86 %, and Support Vector Machine get 82,73 %.

From the comparison, in Table 9 and 10 the Support Vector algorithm can provide the best performance in cell size 64x64 and tenth folds, compared to the k-Nearest Neighbor and Naive Bayes algorithms.

Table 9. Result of Accuracy of Each Classification Method without Information Gain

No	Algorithm Classification	Average Accuracy
1.	K-Nearest Neighbor	74,72%
2.	Naive Bayes	69,41%
3.	Support Vector Machine	81,21%

Table 10. Result of Accuracy of Each Classification Method with Information Gain

No	Algorithm Classification	Average Accuracy
1.	K-Nearest Neighbor	75,52%
2.	Naive Bayes	71,43%
3.	Support Vector Machine	82,81%

The Support Vector Machine algorithm obtains the best accuracy result is 87,86 % at cell size 64x64 and tenth folds with Information Gain.

4. Conclusion

This research has implemented the Information Gain algorithm as a feature selection and weighting is carried out on the feature selection results. The information Gain algorithm works well and can increase accuracy in testing the k-Nearest Neighbor classification. From the previous research, the accuracy rate obtained is 82.5%. After implementing the selection feature Information Gain, the accuracy rate increased to 84.5%.

The simplification of features on the dataset by the Information Gain Algorithm and weighting are also applied to the Naive classification algorithm and the Support Vector Machine.

Acknowledgment

This research was supported by the Magister Faculty of Computer Science, Universitas Dian Nuswantoro.

Reference

[1] F. R. Nur Awalin, “DUNIA BATIN JAWA: Aksara Jawa Sebagai Filosofi dalam Memahami Konsep Ketuhanan,” *Jurnal Ilmu-Ilmu Ushuluddin*, vol. 5, no. 2, pp. 290–209, Dec. 2017, doi: 10.21274/kontem.2017.5.2.289-309.

[2] M. Diqi and M. F. Muhdalifah, “Design and Building Javanese Script Classification in The State Museum of Sonobudoyo Yogyakarta,” vol. 1, no. 2, 2019.

[3] F. Alotaibi, M. T. Abdullah, R. B. H. Abdullah, R. W. B. O. K. Rahmat, I. A. T. Hashem, and A. K. Sangaiah, “Optical Character Recognition for Quranic Image Similarity Matching,” *IEEE Access*, vol. 6, no. c, pp. 554–562, 2017, doi: 10.1109/ACCESS.2017.2771621.

[4] S. bin Ahmed, S. Naz, S. Swati, and M. I. Razzak, “Handwritten Urdu character recognition using one-dimensional BLSTM classifier,” *Neural Comput Appl*, vol. 31, no. 4, pp. 1143–1151, 2019, doi: 10.1007/s00521-017-3146-x.

[5] S. Tian *et al.*, “Multilingual scene character recognition with co-occurrence of histogram of oriented gradients,” *Pattern Recognit*, vol. 51, pp. 125–134, 2016, doi: 10.1016/j.patcog.2015.07.009.

[6] Y. Sari, M. Alkaff, and R. A. Pramunendar, “Iris recognition based on distance similarity and PCA,” *AIP Conf Proc*, vol. 1977, no. June, 2018, doi: 10.1063/1.5042900.

[7] S. Santosa, R. A. Pramunendar, D. P. Prabowo, and Y. P. Santosa, “Wood types classification using back-propagation neural network based on genetic algorithm with gray level co-occurrence matrix for features extraction,” *IAENG Int J Comput Sci*, vol. 46, no. 2, pp. 149–155, 2019.

[8] A. Susanto, D. Sinaga, E. H. Rachmawanto, and D. R. I. M. Setiadi, “Unjuk Kerja K-Nearest Neighbors Pada Pengengalan Karakter Jawa Berbasis Local Binary Pattern,” *Prosiding SNATIF Ke -5 Tahun 2018*, no. 5, pp. 25–30, 2018, doi: 10.2298/PAN0903301G.

[9] N. Gupta, G. v. Pillai, and S. Ari, “Change detection in Landsat images based on local neighbourhood information,” *IET Image Process*, vol. 12, no. 11, pp. 2051–2058, 2018, doi: 10.1049/iet-ipr.2018.5524.

[10] L. Xu, Y. Wang, X. Li, and M. Pan, “Recognition of Handwritten Chinese Characters Based on Concept Learning,” *IEEE Access*, vol. 7, no. 1, pp. 102039–102053, 2019, doi: 10.1109/ACCESS.2019.2930799.

[11] S. Jia, B. Deng, J. Zhu, X. Jia, and Q. Li, “Local Binary Pattern-Based Hyperspectral Image Classification with Superpixel Guidance,” *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 2, pp. 749–759, 2018, doi: 10.1109/TGRS.2017.2754511.

[12] Kurniabudi, D. Stiawan, Darmawijoyo, M. Y. bin bin Idris, A. M. Bamhdi, and R. Budiarto, “CICIDS-2017 Dataset Feature Analysis with Information Gain for Anomaly Detection,” *IEEE Access*, vol. 8, pp. 132911–132921, 2020, doi: 10.1109/ACCESS.2020.3009843.

[13] T. K. Hazra, D. P. Singh, and N. Daga, “Optical character recognition using KNN on custom image dataset,” *2017 8th Industrial Automation and Electromechanical Engineering Conference, IEMECON 2017*, pp. 110–114, 2017, doi: 10.1109/IEMECON.2017.8079572.

[14] A. Elmaizi, H. Nhaila, E. Sarhrouni, A. Hammouch, and C. Nacir, “A novel information gain based approach for classification and dimensionality reduction of hyperspectral images,” *Procedia Comput Sci*, vol. 148, pp. 126–134, 2019, doi: 10.1016/j.procs.2019.01.016.

[15] X. Wang, B. Guo, Y. Shen, C. Zhou, and X. Duan, “Input Feature Selection Method Based on Feature Set Equivalence and Mutual Information Gain Maximization,” *IEEE Access*,

vol. 7, pp. 151525–151538, 2019, doi: 10.1109/ACCESS.2019.2948095.

[16] G. H. Wibowo, R. Sigit, and A. Barakbah, “Javanese Character Feature Extraction Based on Shape Energy,” *EMITTER International Journal of Engineering Technology*, vol. 5, no. 1, pp. 154–170, 2017, doi: 10.24003/emitter.v5i1.175.

[17] Z. Duan, “Characters recognition of binary image using KNN,” *ACM International Conference Proceeding Series*, pp. 116–118, 2018, doi: 10.1145/3198910.3234651.

[18] D. DIvyamary, S. Gopika, S. Pradeeba, and M. Bhuvaneswari, “Brain Tumor Detection from MRI Images using Naive Classifier,” *2020 6th International Conference on Advanced Computing and Communication Systems, ICACCS 2020*, pp. 620–622, 2020, doi: 10.1109/ICACCS48705.2020.9074213.

[19] W. Widayani and H. Harliana, “Analisis Support Vector Machine Untuk Pemberian Rekomendasi Penundaan Biaya Kuliah Mahasiswa,” *Jurnal Sains dan Informatika*, vol. 7, no. 1, pp. 20–27, Jun. 2021, doi: 10.34128/jsi.v7i1.268.

DOI: https://doi.org/10.29207/resti.v7i1.4488