

# Comparative analysis of support vector machine and k-nearest neighbors with a pyramidal histogram of the gradient for sign language detection

Imantoko <sup>1\*</sup>, Arief Hermawan <sup>2</sup>, Donny Avianto <sup>3</sup>

<sup>1,2,3</sup> Magister Teknologi Informasi, Universitas Teknologi Yogyakarta, Indonesia

\*Corresponding Author: [imantoko@student.uty.ac.id](mailto:imantoko@student.uty.ac.id)

**Abstract:** The communication method using sign language is very efficient considering that the speed of information delivery is closer to verbal communication (speaking) compared to writing or typing. Because of this, sign language is often used by people who are deaf, speech impaired, and normal people to communicate. To make sign language translation easier, a system is needed to translate symbols formed from hand movements (in the form of images) into text or sound. This study aims to compare performance such as accuracy and computation time of Support Vector Machine (SVM) and K-Nearest Neighbors (KNN) with Pyramidal Histogram of Gradient (PHOG) for feature extraction, to know which one is better at recognizing sign language. Yield, both combined methods PHOG-SVM and PHOG-KNN can recognize images from hand movements that form certain symbols. The system built using the SVM classification produces the highest accuracy of 82% at PHOG level 3, while the system built with the KNN classification produces the highest accuracy of 78% at PHOG level 2. The total computation time of the fastest training and testing by the SVM model is 236.53 seconds at PHOG level 3, while the KNN model is 78.27 seconds at PHOG level 3. In terms of accuracy, PHOG-SVM is better, but in terms of computation time, PHOG-KNN takes the place.

**Keywords:** Support Vector Machine, K-Nearest Neighbors, Pyramidal Histogram of Gradient, Analysis, Sign Language

**History Article:** Submitted 28 April 2021 | Revised 12 May 2021 | Accepted 25 June 2021

**How to Cite:** Imantoko, A. Hermawan, and D. Avianto, "Comparative Analysis Of Support Vector Machine And K-Nearest Neighbors with Pyramidal Histogram of Gradient For Sign Language Detection," *Matrix: Jurnal Manajemen Teknologi dan Informatika*, vol. 11, no. 2, pp.109-120, 2021.

## Introduction

Sign language is one of several communication methods that can be used by deaf and mute persons as well as normal people to communicate. Communication is carried out by mimicking and hand movements that form symbols so that they interpret a letter or word. The method of communication using sign language is very efficient considering that the speed of information delivery is closer to verbal communication (speaking) compared to writing or typing. To make sign language translation easier, a system is needed to translate symbols formed from hand movements (in the form of images) into text or sound.

Image translation to recognize a certain pattern or shape can be done using the Pyramidal Histogram of Oriented Gradient (PHOG) feature extraction method and the Support Vector Machine (SVM) classification method. Research on sign language recognition using these methods has been conducted [1] with an accuracy of up to 86%. Meanwhile, research [2] obtained an accuracy of up to 91.8% using a combination of the Haar Classifier and K-Nearest Neighbors (KNN) with different datasets. Therefore, this study aims to compare performance such as accuracy and computation time of SVM and KNN with PHOG for feature extraction and the same dataset like [1] to find out which one is better at recognizing sign language.

Research on sign language recognition has also been conducted by [3] using the Convolutional Neural Network (CNN) and Recurrent Neural Network (RNN) methods resulting in an average accuracy value of 60.58%. In addition, [4] also conducted research using the Generalized Learning Vector Quantization (GLVQ) method with an alpha value of 0.9 resulting in the

highest accuracy of 71.37%. Another study by [5] used the Hebb Rule to recognize static sign language and resulted in an accuracy of 80.37%.

Research on letter pattern recognition apart from sign language has also been carried out by [6] using Adaline and concluded that Adaline is effective to be applied in the development of letter pattern recognition applications if the reduced alpha value is 0.05, the enlarged tolerance value is 0.1 and uses bipolar activation function. Research by [7] used an artificial neural network (ANN) to recognize the character pattern of the Karo script. The ANN method studied is perceptron and can recognize up to 100%.

The SVM and KNN classifications have been studied by [8] using Partial Least Square (PLS) as dimension reduction, SVM, and KNN as classifications. This study produced the highest accuracy of 98.54% on leukemia data with PLS-KNN, 100% on lung data with KNN, 66.52% on breast data with PLS-KNN, and 85.60% on colon data with PLS-SVM.

Research on the introduction of tajwid law has been carried out by [9] using the KNN with Local Mean which has succeeded in achieving the highest number of 96.43%.

Another research about image recognition was also conducted by [10], but this study focused on license plate recognition using Object Character Recognition (OCR). The success rate for recognizable license plates was 75%.

## Methodology

### Pyramidal Histogram of Gradient (PHOG)

The way the PHOG method works is to do HOG according to the size of the cells and the orientation bin then combine them into 1 feature like the formula [11].

### Dataset

The dataset used is an open-source dataset from [12]. Consists of a collection of near-infrared (near-infrared) images and skeletal information obtained from the Leap Motion sensor. The composition of the dataset includes 16 types of hand movements including palm, L, fist-moved, down, index, ok, palm m, C, heavy, hang, two, three, four, five, palm u, and up from 15 subjects (five male and 10 female). This dataset has a total of 13000 hand motion infrared images. However, based on [1] the types of hand movements used are palm, L, down, index, ok, and C. Out of 13,000 images, [1] states that only 3800 images are used, with the division of 3300 images for training data and 500 test data. However, in the discussion in [1], only 486 images were used, so that the total images used were 3786 images with random sampling. The distribution of the number of images for each type of hand movement can be seen in tables 1 and 2.

**Table 1.** Training data

Class	Subtotal
Palm	550
L	550
Down	550
Index	550
Ok	550
C	550
<b>Total</b>	<b>3300</b>

Table 1 shows the six types of training data hand movements (classes). Palm as many as 550 images, L as many as 550 images, Down as many as 550 images, Index as many as 550 images, Ok as many as 550 images, and C as many as 550 images. The total of images for training is 3300 images.

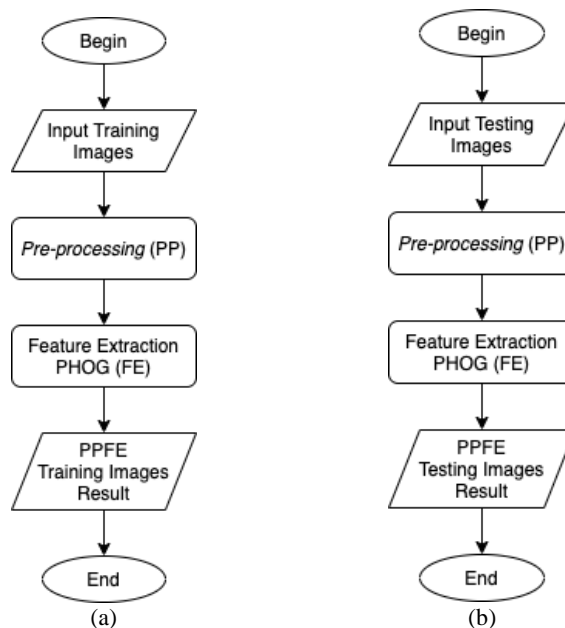
**Table 2.** Testing data

Class	Subtotal
Palm	80
L	80
Down	86
Index	70
Ok	90
C	80
<b>Total</b>	<b>486</b>

Table 2 shows the six types of test data hand movements (classes). Palm as many as 80 images, L as many as 80 images, Down as many as 86 images, Index as many as 70 images, Ok as many as 90 images, and C as many as 80 images. The total number of images for training is 486 images.

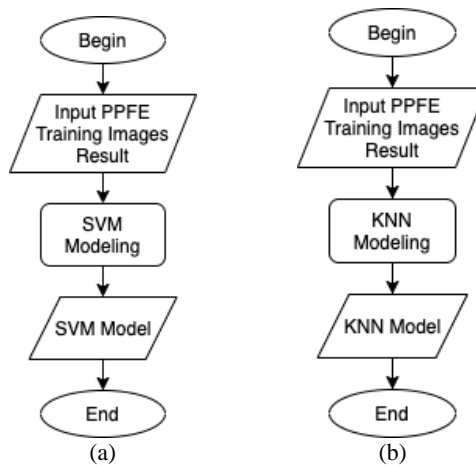
### System Design

This study aims to compare two classification methods, namely SVM and KNN to find out which one is better at recognizing sign language in terms of accuracy and computation time. The feature extraction used for the comparison of the two methods is PHOG levels 1, 2, and 3. The three main stages in this research are Pre-processing & Feature Extraction (PPFE), Training, and Testing.



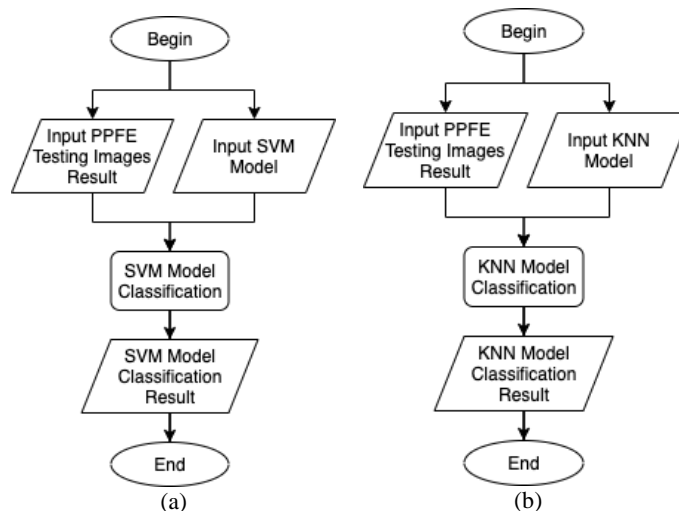
**Figure 1.** (a) PPFE training data stage, (b) PPFE test data stage

Figure 1 shows the PPFE stage for training data (a) and test data (b). Overall, the process for both data is the same, namely through pre-processing and extraction of PHOG features. The output of this stage is training data and test data that have been pre-processed and feature extraction. The output training data is used for the SVM and KNN Training stage, while the test data is used for the SVM and KNN Testing stage.



**Figure 2.** (a) SVM training data stage, (b) KNN training data stage

Figure 2 shows the training stage for SVM (a) and KNN (b). The input of this stage is the PPFE output training data. The output of the PPFE training data processed with SVM and KNN modeling. Each of them produces the SVM model and the KNN model that are used for the Testing stage.



**Figure 3.** (a) SVM testing data stage, (b) KNN testing data stage

Figure 3 shows the testing stage of the SVM (a) and KNN (b) models. At this stage, the input is the PPFE output test data and classification model. The model classifies the PPFE test data output and produces the predictive results of the test data image. The result presented as Confusion Matrix, Precision, Recall, F1-Score, Accuracy, and Computation Time.

### Pre-processing

The image from [12] varies in size from 408 x 264 to 420 x 273, so it is necessary to equalize the image size. All image sizes are converted to 408 x 264 using the crop method. When viewed from the largest size, which is 420 x 273, the reduction that occurs is only 9 to 12 pixels.

### Feature Extraction

After the image pre-processing, the characteristics of the image are extracted using the PHOG method level 1, 2, and 3 for the training image and the test image. Each PHOG level applies HOG with a bin of 9, pixels per cell 8 x 8, and cells per block 2 x 2.

## Training

Training or modeling using the SVM and KNN methods on 3300 training data that has passed the pre-processing and feature extraction processes. SVM uses a polynomial kernel configuration, degree 3, and tolerance of 0.00001 while KNN uses a number of neighbors of 3 ( $k = 3$ ), the configuration of each method applied to all PHOG levels. The result of this training is that there are 3 SVM models and KNN models each following each PHOG level.

## Testing

Testing was carried out with the SVM and KNN models on 486 test data that had passed the pre-processing and feature extraction processes. The test data of the PHOG level 1 feature extraction results were tested with the SVM model and the KNN level 1 model, and so on. The metric used for evaluating the recognition performance is the confusion matrix (CM) presented in graphical form. From CM values, four different measurements have been collected.

Precision (P) is the ratio of true positive predictions compared to the overall positive predicted outcome. The precision formula is as follows:

$$P = \frac{tp}{tp + fp} \quad (1)$$

Recall (R) is the ratio of true positive predictions compared to the total number of true positive data. The recall is calculated through:

$$R = \frac{tp}{tp + fn} \quad (2)$$

F1-score (F) is a weighted comparison of the average precision and recall. F1-score is calculated as:

$$F = 2 \times \frac{P \times R}{P + R} \quad (3)$$

Accuracy (A) is the ratio of Correct predictions (positive and negative) to the overall data. The accuracy formula is as follows:

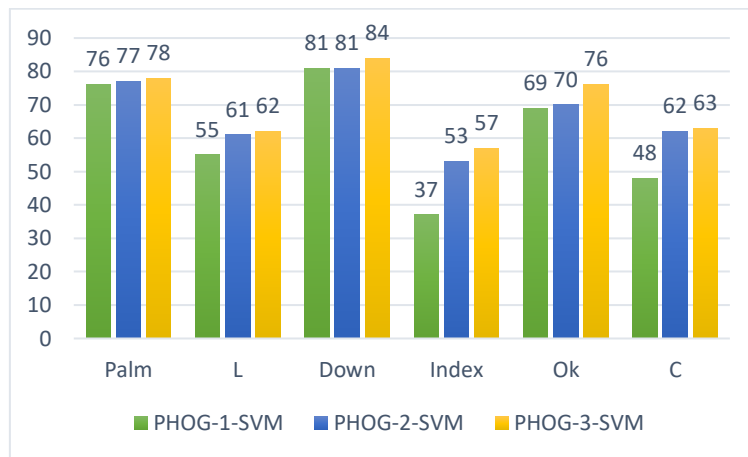
$$A = \frac{tp + tn}{tp + tn + fp + fn} \quad (4)$$

Where  $tp$  represents true positive samples,  $tn$  the true negative samples,  $fp$  the false-positive samples, and  $fn$  the false-negative samples [13].

## Results and Analysis

### Correct Prediction Graph

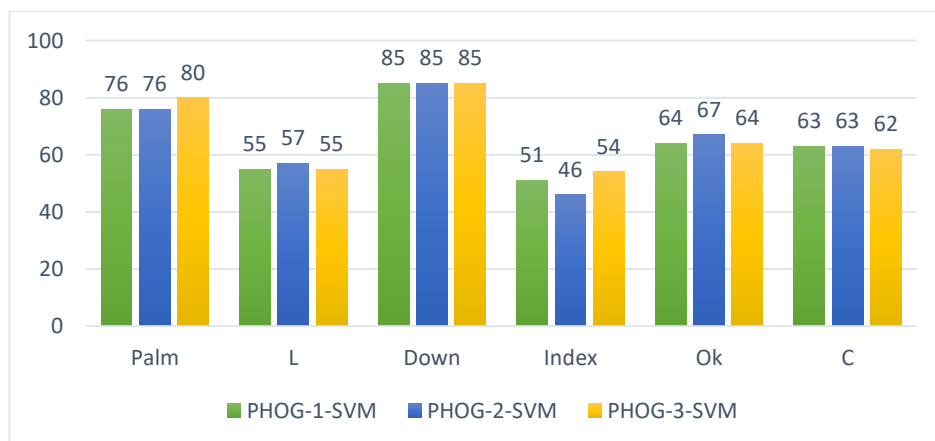
The results of the SVM Confusion Matrix model for each PHOG level in [1] represents in graphical form can be seen in Figure 4.



**Figure 4.** Correct prediction graph PHOG-SVM [1]

Figure 4 shows the number of correct predictions for each class using PHOG Level 1-SVM, PHOG Level 2-SVM, and PHOG Level 3-SVM, respectively in the study [1]. Palm got a correct prediction of 76, 77, and 78. L got a correct prediction of 55, 61, and 62. Down got a correct prediction of 81, 81, and 84. The index got a correct prediction of 37, 53, and 57. Ok got a correct prediction of 69, 70, and 76. C got a correct prediction of 48, 62, and 63.

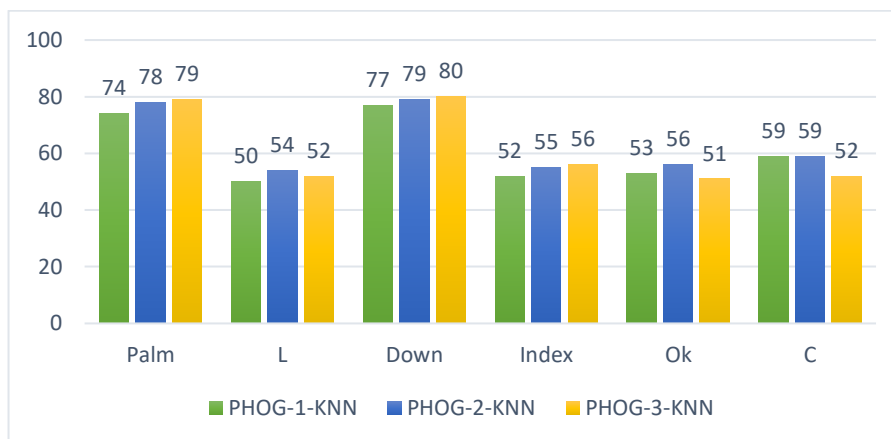
Meanwhile, the results of the SVM Confusion Matrix model for each PHOG level in this research represents in graphical form can be seen in figure 5.



**Figure 5.** Correct prediction graph PHOG-SVM

Figure 5 shows the number of correct predictions for each class using PHOG Level 1-SVM, PHOG Level 2-SVM, and PHOG Level 3-SVM, respectively in this study. Palm got a correct prediction of 76, 76, and 80 out of 80. L got a correct prediction of 55, 57, and 55 out of 80. Down got a correct prediction of 85, 85, and 85 out of 86. Index got a correct prediction of 51, 46, and 54 out of 70. Ok got a correct prediction of 64, 67, and 64 out of 90. C got a correct prediction of 63, 63, and 62 out of 80.

In the other hand, the results of the KNN Confusion Matrix model for each PHOG level in this research represents in graphical form can be seen in figure 6.



**Figure 6.** Correct prediction graph PHOG-KNN

Figure 6 shows the number of correct predictions for each class using PHOG Level 1-KNN, PHOG Level 2-KNN, and PHOG Level 3-KNN, respectively in this study. Palm got a correct prediction of 74, 78, and 79 out of 80. L got a correct prediction of 50, 54, and 52 out of 80. Down got a correct prediction of 77, 79, and 80 out of 86. Index got a correct prediction of 52, 55, and 56 out of 70. Ok got a correct prediction of 53, 56, and 51 out of 90. C got a correct prediction of 59, 59, and 52 out of 80.

### Precision, Recall, F1-Score, and Accuracy

The results of the Precision, Recall, F1-Score, and Accuracy of the SVM model for each PHOG level in [1] can be seen in table 3.

**Table 3.** Precision, recall, F1-score, and accuracy PHOG-SVM [1]

Class	Model	Precision	Recall	F1-Score
Palm	PHOG-1-SVM	0,93	0,95	0,94
	PHOG-2-SVM	0,96	0,97	0,97
	PHOG-3-SVM	0,96	0,97	0,97
L	PHOG-1-SVM	0,62	0,69	0,65
	PHOG-2-SVM	0,82	0,78	0,79
	PHOG-3-SVM	0,82	0,78	0,79
Down	PHOG-1-SVM	0,92	0,94	0,93
	PHOG-2-SVM	0,98	0,98	0,98
	PHOG-3-SVM	0,98	0,98	0,98
Index	PHOG-1-SVM	0,70	0,53	0,60
	PHOG-2-SVM	0,73	0,82	0,77
	PHOG-3-SVM	0,73	0,81	0,77
Ok	PHOG-1-SVM	0,67	0,77	0,72
	PHOG-2-SVM	0,84	0,84	0,84
	PHOG-3-SVM	0,84	0,84	0,84
C	PHOG-1-SVM	0,67	0,60	0,63
	PHOG-2-SVM	0,75	0,78	0,76
	PHOG-3-SVM	0,85	0,79	0,82
<b>Accuracy</b>	PHOG-1-SVM	-	-	0,75
	PHOG-2-SVM	-	-	0,83
	PHOG-3-SVM	-	-	0,86

Table 3 shows the precision, recall, f1-score for each class using PHOG-1-SVM, PHOG-2-SVM, PHOG-3-SVM respectively and accuracy PHOG-1-SVM, PHOG-2-SVM, PHOG-3-SVM respectively in study [1]. Palm obtained a precision of 0.93, 0.96, and 0.96; recall of 0.95, 0.97, and

0.97; f1-score of 0.94, 0.97, and 0.97. L obtained a precision of 0.62, 0.82, and 0.82; recall of 0.69, 0.78, and 0.78; f1-score of 0.65, 0.79, and 0.79. Down obtained a precision of 0.92, 0.98, and 0.98; recall of 0.94, 0.98, and 0.98; f1-score of 0.93, 0.98, and 0.98. Index obtained a precision of 0.70, 0.73, and 0.73; recall of 0.53, 0.82, and 0.81; f1-score of 0.60, 0.77 and 0.77. Ok obtained a precision of 0.67, 0.84, and 0.84; recall of 0.77, 0.84, and 0.84; f1-score of 0.60, 0.77, and 0.77. C obtained a precision of 0.67, 0.75, and 0.85; recall of 0.60, 0.78, and 0.79; f1-score of 0.63, 0.76, and 0.82. The accuracy obtained is 0.75, 0.83, and 0.86.

The results of the Precision, Recall, F1-Score, and Accuracy of the SVM model for each PHOG level in the research conducted can be seen in table 4.

**Table 4.** Precision, recall, F1-score, and accuracy PHOG-SVM

<b>Class</b>	<b>Model</b>	<b>Precision</b>	<b>Recall</b>	<b>F1-Score</b>
Palm	PHOG-1-SVM	0,99	0,95	0,97
	PHOG-2-SVM	0,99	0,95	0,97
	PHOG-3-SVM	0,99	1,00	0,99
L	PHOG-1-SVM	0,79	0,69	0,73
	PHOG-2-SVM	0,77	0,71	0,74
	PHOG-3-SVM	0,86	0,69	0,76
Down	PHOG-1-SVM	0,96	0,99	0,97
	PHOG-2-SVM	0,96	0,99	0,97
	PHOG-3-SVM	1,00	0,99	0,99
Index	PHOG-1-SVM	0,61	0,61	0,67
	PHOG-2-SVM	0,62	0,66	0,64
	PHOG-3-SVM	0,62	0,77	0,69
Ok	PHOG-1-SVM	0,71	0,71	0,71
	PHOG-2-SVM	0,72	0,74	0,73
	PHOG-3-SVM	0,72	0,71	0,72
C	PHOG-1-SVM	0,82	0,79	0,80
	PHOG-2-SVM	0,80	0,79	0,79
	PHOG-3-SVM	0,78	0,78	0,78
<b>Accuracy</b>	PHOG-1-SVM	-	-	0,81
	PHOG-2-SVM	-	-	0,81
	PHOG-3-SVM	-	-	0,82

Table 4 shows the precision, recall, f1-score for each class using PHOG-1-SVM, PHOG-2-SVM, PHOG-3-SVM respectively and accuracy PHOG-1-SVM, PHOG-2-SVM, PHOG-3-SVM respectively in this study. Palm obtained a precision of 0.99, 0.99, and 0.99; recall of 0.95, 0.95, and 1.00; f1-score of 0.97, 0.97, and 0.99. L obtained a precision of 0.79, 0.77, and 0.86; recall of 0.69, 0.71, and 0.69; f1-score of 0.73, 0.74, and 0.76. Down obtained a precision of 0.96, 0.96, and 1.00; recall of 0.99, 0.99, and 0.99; f1-score of 0.97, 0.97, and 0.99. Index obtained a precision of 0.61, 0.62, and 0.62; recall of 0.61, 0.66, and 0.77; f1-score of 0.67, 0.64 and 0.69. Ok obtained a precision of 0.71, 0.72, and 0.72; recall of 0.71, 0.74, and 0.71; f1-score of 0.71, 0.73, and 0.72. C obtained a precision of 0.82, 0.80, and 0.78; recall of 0.79, 0.79, and 0.78; f1-score of 0.80, 0.79, and 0.78. The accuracy obtained is 0.81, 0.81, and 0.82.

The results of the Precision, Recall, F1-Score, and Accuracy of the KNN model for each PHOG level in the research conducted can be seen in table 5.



**Table 5.** Precision, recall, F1-score, and accuracy PHOG-KNN

<b>Class</b>	<b>Model</b>	<b>Precision</b>	<b>Recall</b>	<b>F1-Score</b>
Palm	PHOG-1-KNN	0,85	0,93	0,89
	PHOG-2-KNN	0,87	0,97	0,92
	PHOG-3-KNN	0,91	0,99	0,95
L	PHOG-1-KNN	0,67	0,62	0,65
	PHOG-2-KNN	0,70	0,68	0,69
	PHOG-3-KNN	0,68	0,65	0,66
Down	PHOG-1-KNN	0,88	0,90	0,89
	PHOG-2-KNN	0,95	0,92	0,93
	PHOG-3-KNN	0,95	0,93	0,94
Index	PHOG-1-KNN	0,65	0,74	0,69
	PHOG-2-KNN	0,70	0,79	0,74
	PHOG-3-KNN	0,60	0,80	0,68
Ok	PHOG-1-KNN	0,72	0,59	0,65
	PHOG-2-KNN	0,72	0,62	0,67
	PHOG-3-KNN	0,72	0,57	0,63
C	PHOG-1-KNN	0,72	0,74	0,73
	PHOG-2-KNN	0,75	0,74	0,74
	PHOG-3-KNN	0,71	0,65	0,68
<b>Accuracy</b>	PHOG-1-KNN	-	-	0,75
	PHOG-2-KNN	-	-	0,78
	PHOG-3-KNN	-	-	0,76

Table 5 shows the precision, recall, f1-score for each class using PHOG-1-KNN, PHOG-2-KNN, PHOG-3-KNN respectively and accuracy PHOG-1-KNN, PHOG-2-KNN, PHOG-3-KNN respectively in this study. Palm obtained a precision of 0.85, 0.87, and 0.91; recall of 0.93, 0.97, and 0.99; f1-score of 0.89, 0.92, and 0.95. L obtained a precision of 0.67, 0.70, and 0.68; recall of 0.62, 0.68, and 0.65; f1-score of 0.65, 0.69, and 0.66. Down obtained a precision of 0.88, 0.95, and 0.95; recall of 0.90, 0.92, and 0.93; f1-score of 0.89, 0.93, and 0.94. Index obtained a precision of 0.65, 0.70, and 0.60; recall of 0.74, 0.79, and 0.80; f1-score of 0.69, 0.74 and 0.68. Ok obtained a precision of 0.72, 0.72, and 0.72; recall of 0.59, 0.62, and 0.57; f1-score of 0.65, 0.67, and 0.63. C obtained a precision of 0.72, 0.75, and 0.71; recall of 0.74, 0.74, and 0.65; f1-score of 0.73, 0.74, and 0.68. The accuracy obtained is 0.75, 0.78, and 0.76.

### Computation Time

This study also compared the computation time required by the SVM and KNN methods for training 3300 images and testing 486 images. Computation time is calculated in seconds. The hardware used is the Macbook Pro 2019. The time required for each method to PHOG level n can be seen in Table 6. However, research [1] did not take into account the computation time, so table 6 only shows the computation time carried out in this study only.

**Table 6.** Computation time in seconds

<b>Model</b>	<b>Training</b>	<b>Testing</b>	<b>Total</b>
PHOG-1-SVM	353,26	35,85	389,11
PHOG-2-SVM	280,26	29,86	310,12
PHOG-3-SVM	212,64	23,89	236,53
PHOG-1-KNN	3,41	113,84	117,25
PHOG-2-KNN	1,10	98,53	99,63
PHOG-3-KNN	0,59	77,68	78,27

Table 6 shows the computation time of each model at the training stage (3300 images), testing stage (486 images), and the total time for the two stages in seconds. The PHOG-1-SVM model requires a computational training time of 353.26 seconds, testing 35.85 seconds, and a

total of 389.11 seconds. The PHOG-2-SVM model requires a computational training time of 280.26 seconds, testing 29.86 seconds, and a total of 310.12 seconds. The PHOG-3-SVM model requires a computational training time of 212.64 seconds, testing 23.89 seconds, and a total of 236.53 seconds. The PHOG-1-KNN model requires a computational training time of 3.41 seconds, testing 113.84 seconds, and a total of 117.25 seconds. The PHOG-2-KNN model requires a training computation time of 1.10 seconds, testing 98.53 seconds, and a total of 99.63 seconds. The PHOG-3-KNN model requires a computational training time of 0.59 seconds, testing 77.68 seconds, and a total of 78.27 seconds.

## Analysis

Figure 4 and Figure 5 are expected to have similar results, considering that these figures use the same dataset source, pre-processing method, feature extraction, and classification. However, the results are quite different, in Figure 4 the total correct prediction is 366, 404, and 420, respectively, while Figure 5 are 394, 394, and 400 respectively. Quite a big increase occurs in Figure 4, but only a little in Figure 5.

On the other hand, Figure 4 and Figure 5 have similarities, namely prediction or testing in class L and Index has poor results, followed by classes Ok and C, while the best prediction results are in the Palm and Down classes.

Switch to Precision, Recall, F1-Score, and accuracy results. A clear difference occurs in table 3 as a whole experience an increase in the precision, recall, f1-score, and accuracy results. A clear difference can be seen in the accuracy of Table 3, respectively, namely 75%, 83%, and 86%, while Table 4 hardly experience a significant increase, namely 81%. 81%, and 82%.

The difference in the results of confusion matrix, precision, recall, f1-score, and accuracy is very possible due to two factors. The first factor is the selection of the image dataset from 13000, only 3786 images were selected, it is very possible that the image used in [1] is not the same as the research conducted. The second factor is [1] not explicitly mentioning the SVM configuration used, so in this study, the authors tried several configurations and the best results were obtained on the SVM configuration with a polynomial kernel, degree 3, and tolerance 0.00001.

This difference can still be tolerated considering the results of the two SVMs have not been missed too far and the research conducted has used the source dataset, pre-processing methods, feature extraction, and the same classification as previous studies. For the SVM and KNN comparisons to be valid, comparisons are only made to the SVM and KNN models built in this study.

Returning to the results of the confusion matrix, the SVM model in Figure 5 has better results than the KNN model in Figure 6 where the total correct predictions on PHOG-1-SVM, PHOG-2-SVM, and PHOG-3-SVM respectively were 394, 394, and 400, while the PHOG-1-KNN, PHOG-2-KNN, and PHOG-3-KNN were 365, 381, and 370 respectively. The SVM model experienced an increase in the number of correct predictions although not significant, in the KNN model there was an increase from PHOG-1-KNN to PHOG-2-KNN then get off at PHOG-3-KNN.

There are similarities to the results of PHOG-1-SVM, PHOG-2-SVM, PHOG-3-SVM, PHOG-1-KNN, PHOG-2-KNN, and PHOG-3-KNN, namely predictions in class L and Index have poor results, followed by classes Ok and C, while the results best predictions for the Palm and Down classes. If we review the L and Index images, in plain view there are indeed similarities in several images as in Figures 7 and 8.



**Figure 7.** (a) Sample image L, (b) Sample image index

Figure 7 shows the visible similarities between the L (a) image sample and the index (b) image sample.



**Figure 8.** (a) Sample image Ok, (b) Sample image C

Figure 8 shows the visible similarity between the Ok image sample (a) and the image sample C (b).

Then the results of precision, recall, f1-score, and accuracy from table 4 show that PHOG-1-SVM, PHOG-2-SVM, and PHOG-3-SVM have increased insignificantly, even though there are classes that have low precision, recall, and f1-scores such as L, Index, Ok, and C. Low f1-score results indicate that the model is difficult to distinguish from other classes. Low F1-score due to low precision and recall. The accuracy obtained by the PHOG-1-SVM and PHOG-2-SVM models is 81% and PHOG-3-SVM is 82%. Meanwhile, based on table 5 the PHOG-1-KNN, PHOG-2-KNN, and PHOG-3-KNN models have lower precision, recall, f1-score, and accuracy results than the SVM model. The accuracy obtained by the PHOG-1-KNN model was 75%, PHOG-2-KNN increased to 78%, and there was a decrease in PHOG-3-KNN to 76%. These results indicate that at each PHOG level the SVM model is better than the KNN model.

When viewed from the aspect of computation time based on table 6, the KNN model is superior to the SVM model in the training process. For example, the PHOG-1-KNN model training process takes 3.41 seconds, while the PHOG-1-SVM takes 353.26 seconds. The SVM model is superior to the KNN model during the testing process. It can be seen that the PHOG-1-SVM model takes 35.85 seconds, while the PHOG-1-KNN takes 113.84 seconds. Overall, the KNN model is far superior to the SVM model when viewed from the total computation time required. The PHOG-1-KNN model has a total of 117.25 seconds, while the PHOG-1-SVM Model takes 389.11 seconds. Based on table 6, the total computation time of the KNN model is 78.27 seconds, while the SVM model is 236.53 seconds at PHOG level 3. Looking from another point of view, the SVM and KNN models have the same computation time pattern where the higher the PHOG level, the faster the computation time, so it can be concluded that PHOG level affects the computation time of the two models.

## Conclusion

Experiments in this study with previous research there are differences in the results of confusion matrix, precision, recall, f1-score, and accuracy. The first factor that causes the difference in the results of the two studies is the selection of the image dataset from 13000, only 3786 images were selected, it is very possible that the image used in the previous study was not the same as the research carried out. The second factor is that previous research did not explicitly mention the SVM configuration used, so in this study, the authors tried several configurations and the best results were obtained on the SVM configuration with a polynomial kernel, degree 3, and tolerance 0.00001.

Based on its accuracy, the SVM method is superior to the KNN method at all PHOG levels. PHOG1-SVM produces an accuracy of 81%, while PHOG-1-KNN is only 75%. The PHOG-2-SVM produces 81% accuracy, while the PHOG-2-KNN has an accuracy of 78%. PHOG-3-SVM experienced an increase in accuracy to 82%, while PHOG-3-KNN experienced a decrease in accuracy to 76%. Based on the total computation time required for training and testing, the KNN method is superior to the SVM method at all PHOG levels. The KNN method requires the fastest

total computation time of 78.27 seconds at PHOG level 3, while the SVM method is 236.53 seconds at PHOG level 3.

## References

- [1] R. Novanda, K. N. Ramadhan, and P. E. Yunanto, "Pengenalan bentuk tangan dengan ekstraksi ciri pyramid histogram of oriented gradient ( PHOG ) dan klasifikasi Support Vector Machine ( SVM )," in *e-Proceeding of Engineering*, 2020, vol. 7, no. 2, pp. 8223–8233.
- [2] R. I. Borman, B. Priyopradono, and A. R. Syah, "Klasifikasi objek kode tangan pada pengenalan isyarat alphabet bahasa isyarat indonesia (BISINDO)," in *Prosiding Seminar Nasional Informatika dan Aplikasinya (SNIA)*, 2017.
- [3] D. Yolanda, K. Gunadi, and E. Setyati, "Pengenalan alfabet bahasa isyarat tangan secara real-time dengan menggunakan metode Convolutional Neural Network dan Recurrent Neural Network," *J. Infra*, vol. 8, no. 1, 2020.
- [4] D. Gustiar , S. H. Sitorus , and D. M. Midyanti, "Penerjemahan bahasa isyarat menggunakan metode Generalized Learning Vector Quantization (GLVQ)," *Coding J. Komput. dan Apl.*, vol. 8, no. 3, pp. 1-8, 2020.
- [5] D. Setiawan and A. Z. Falani, "Pemanfaatan Artificial Neural Network dengan metode Hebb Rule untuk pengenalan bahasa isyarat indonesia statis," *j. spirit*, vol. 12, no. 1, pp. 9–15, 2020.
- [6] B. Etikasari and T. D. Puspitasari, "Pengenalan pola huruf dengan menggunakan Algoritma Adaline Bety," *J. Mnemon.*, vol. 2, no. 1, pp. 12–16, 2019.
- [7] D. R. Yani, "Penerapan jaringan syaraf tiruan dalam pengenalan huruf aksara suku karo dengan metode Perceptron," *J. Inf. Sist. Res.*, vol. 1, no. 3, pp. 109–114, 2020.
- [8] S. A. Naufal, A. Adiwijaya, and W. Astuti, "Analisis perbandingan klasifikasi Support Vector Machine (SVM) dan K-Nearest Neighbors (KNN) untuk deteksi kanker dengan data Microarray," *JURIKOM (Jurnal Ris. Komputer)*, vol. 7, no. 1, pp. 162-168, 2014.
- [9] I. Fathurrahman and A. Kurniawardhani, "Pengenalan hukum tajwid pada citra Al-Quran : kajian pustaka," *J. Karya untuk Masy.*, vol. 2, no. 1, 2020.
- [10] W. S. Ismail, P. W. Purnawan, I. Riyanto, and N. Nazori, "Sistem perekaman pelat nomor mobil pada palang pintu parkir menggunakan web kamera dan mikrokontroler," *Matrix J. Manaj. Teknol. dan Inform.*, vol. 10, no. 3, pp. 103–112, 2020.
- [11] F. Sthevanie, P. I. Aristya, and K. N. Ramadhani, "Pengenalan aksara bali menggunakan metode pyramid histogram of oriented gradients," *Indo-Journal Comput.*, vol. 5, no. 1, pp. 73–84, 2020.
- [12] T. Mantecón, "Multi-modal dataset for hand gesture recognition," 2018. <https://www.kaggle.com/gti-upm/multimodhandgestrec> (accessed Jan. 18, 2021).
- [13] B. Raza, Y. J. Kumar, A. K. Malik, A. Anjum, and M. Faheem, "Performance prediction and adaptation for database management system workload using Case-Based Reasoning approach," *Inf. Syst.*, vol. 76, pp. 46–58, 2018.

© 2021 by the author; licensee Matrix: Jurnal Manajemen Teknologi dan Informatika. This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution license (<http://creativecommons.org/licenses/by/4.0/>).