FINGERSPELLING INDONESIAN SIGN LANGUAGE USING LEARNING VECTOR QUANTIZATION 3

Dinar Arisandra[1], Galih Hermawan[2]

Informatics Engineering Study Program
[1,2] Faculty of Engineering and Computer Science, Indonesia Computer University
Jl. Dipati Ukur 114 Bandung
E-mail : arisandra54@gmail.com[1], galih.hermawan@email.unikom.ac.id[2]

ABSTRACT

Fingerspelling is used to provide information on names, people and places and other objects that are not yet known or unknown to sign language such as the names of people, companies, brands and so on. Translators in sign language from fingerspelling have been made with various methods, but there are always obstacles or accuracy that are still lacking in translating moving letters, namely letters J and Z. The introduction program that will be created will use vector quantization 3 learning method, where this method is an optimization of the previous method, namely vector quantization learning. Detection of hand signals using a leap motion controller media tool, because this tool has been specified will only detect on the hand. The Leap motion controller will make the detection room 600 mm high, 150 degrees long, and 120 degrees wide. The output generated from the leap motion controller is a vector consisting of palm1 (normal palm), palm2 (palm position), palm3 (palm velocity), hp (x angle), roll (z angle), and yaw (y angle). Based on the results of research, analysis and design, implementation and testing stages, fingerspelling Indonesian signaling system using vector quantization 3 learning method produces an accuracy of 54%.

Keywords: fingerspelling, learning vector quantization 3, leap motion controller

1. INTRODUCTION

Sign language is the primary language of deaf people, the House of tunawicara or the deaf. Indonesia has two sign language that is actively used, i.e. the system Cue Indonesian Language (SIBI) and Indonesia sign language (BISINDO). Although only the official dictionary created SIBI by the Government in order to adapt the sign language with an enhanced spelling (EYD), both of which remain actively used though to BISINDO itself has little difference vocabulary in each region certain. Both languages are not learned by everyone, only a few people interested and who has an interest.

Fingersspelling is commonly used to inform the name of the people, places and other objects that are not yet known or not known by sign language such as names of people, companies, brands and more. In previous research already test method using fingerspelling Viola & Jones that produces the accuracy testing of 55%, [1] and Naive Bayesian method of testing produce accuracy of 69% [2]. Previous research has not been able to detect a moving letter from SIBI, namely the letter J and z. letters with regard to research utilizing technology Leap Motion Controller with base of classification method of Backpropagation can detect moving Letter J and Z as well as generate the accuracy of 90% [3]. While the research is going to do this is using a technology Leap Motion Controller and base klasikasi method using LVQ algorithms development namely LVQ3. In the study of classification of facial identity using LVQ earns a high enough accuracy that is 92% [4]. For LVQ3 himself has compared with LVQ in research classification of nutritional status of the child ", penelitannya earn 88% accuracy results for LVQ and LVQ3 95.2% was for [5].

LVQ was one method of classification of patterns that each unit of output represents a category or a specific group. Processing that occurs on every neuron is looking for the nearest distance between a vector input to the corresponding weights [6]. The advantages of this method are in addition to searching for the nearest distance learning units, so long as the output is positioned with the set up and update the weights through the learning terawasi to estimate the decision classification. As for LVQ3, two vectors (the winner and runner-ups) updated if some conditions are met. LVQ algorithms development idea is if input estimates have the same distance vector with the winners and runners-up, then each such vector should do a study [7].

Based on the above explanation, the study Indonesian Language fingerspelling Gestures System Using Learning Vector Quantizatino 3 is intended to utilize the technology Leap Motion Controller and algorithm development LVQ i.e. LVQ3 on research this time, allow in order for fingerspelling can generate accuracy differs from previous studies in detection of letters that had a great gesture, namely the letter J and Z.

2. THEORETICAL BASIC

2.1 Indonesian Sign Language (SIBI)

Indonesian Sign Language (SIBI) which standardized it is one medium that helps communications fellow House of deaf people in the wider community. His form is a systematic order is about a set of gestures, hands, fingers and a range of motion that symbolizes Indonesia language vocabulary [8].

2.2 Leap Motion Controller

Leap Motion Controller is a useful device to detect the motion of the hand signals that cannot be represented by the mouse and keyboard. Leap Motion Controller is able to detect the finger from a distance of up to 600 millimeters. This tool has two black-and-white cameras and three infrared LED with accuracy up to 1/100 millimetres. Area coverage for the tracking Motion Controller diilustraikan Leap in Figure 1 [9].

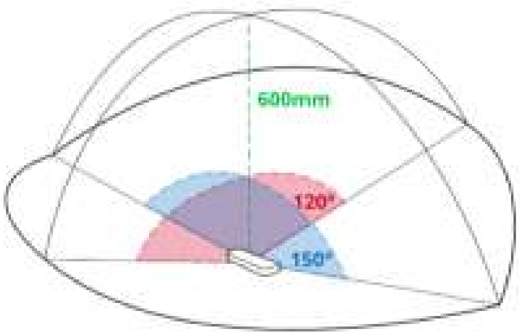

Figure 1. Area Coverage Leap Motion Controller

LMC had a library which provides a value in the form of a vector from every hand gesture at capture using the LMC. The vector taken from LMC consists of 6 vector, i.e.:

1. Palm Position: Center of palm of hand, which was divided into three palm1, i.e. vector palm2, and palm3
2. Hand Pitch: angle of the x-axis
3. the Yaw angle of the y axis:
4. Roll angle of the z-axis:

2.3 Classification Learning Vector Quantization (LVQ)

Pattern classification method of LVQ is that each unit of output represents a particular class or category. LVQ method used for grouping by the number of target or pre-set class [10].

Each class is represented on one set of data used in the training, which was later called the weights. A competitive layer automatically learn to classify the input vectors. Classes are obtained as a result of this competitive layer depends only on the distance between the vectors input vector with weights, which are then weighted from the class winners were revised. If two input vectors is approaching the same competitive layer put both the input vector into the same class. The working principle of the LVQ algorithms is the reduction of nodes that ultimately there is only one output node is selected. Figure 2 shows the architecture of the LVQ.

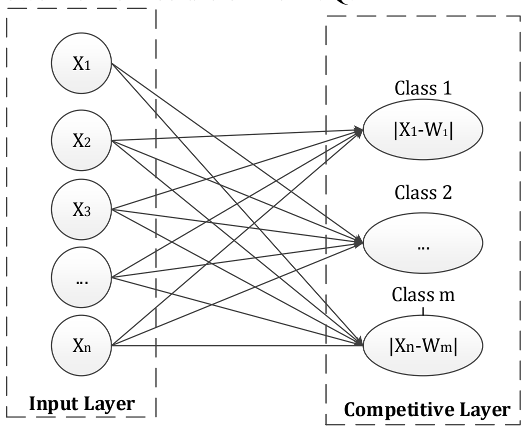

Figure 2. Architecture LVQ

Description:
X 1-Xn is the input vector
W1-Wm is the vector of weights on a competitive layer.

LVQ learning early stage initial weights for each node of the input nodes that represent each class. After determining the weighting, the network was given the number of input dimensions of input neurons. After input of the network, the network received started doing the calculation of distance vector are obtained with the distance between the input vector by a vector of weights.

End of the training obtained final weight (W). This weighting-weighting was later used for the identification of other data. At the moment the only training weights winner with minimum revision done. Stages of LVQ algorithms, namely:

In LVQ3 learning system was developed based on basic provision and LVQ1 algorithm theory LVQ3.

1. Set:
   a. Weighting (w),
   b. Maximum iterations (MaxEpoch)
   c. The learning Rate (α)
   d. Minimum learning rate (eps)
   e. Epoch = 0
2. Do if the iteration has not yet reached the maximum (Epoch < MaxEpoch) or the learning rate is greater than the minimum learning rate (α > eps)
   a. Epoch = epoch + 1
   b. Do for i = 1 to maximum iterations (MaxEpoch)

c. Specify the distance (J) to seek a minimum 1 Equations with w.

$$J = \sqrt{(x_{11} - w_{11})^2 + \ldots + (x_{1n} - w_{1n})^2} \quad (1)$$

d. The value of the window (ε), used as the area that must be met to renew the reference vector winner (Y1) and runner-up (Y2) if it were in a different class. The equation window (ε):

$$\text{Min} = (dc1/dc2, dc2/dc1) > (1\text{-}\varepsilon)\ (1 + \varepsilon) \quad (2)$$

The value of the window (ε) = 0.3 (Fausett: 1994)

e. If it meets the conditions window (ε), then updates using the formula:

$$W1, t + 1 = t, \alpha t\ w1\ (x\text{-}w1t) \quad (3)$$
$$W1, t + 1 = t, \alpha t + w1\ (x\text{-}w1t)$$

f. Whereas if not meet then updates using the formula:

$$W1, (t + 1) = w1\ (t) - \beta\ (t)\ (x\text{-}w1\ (t)) \quad (4)$$
$$W1, (t + 1) = w1\ (t) + \beta\ (t)\ (x\text{-}w1\ (t))$$
$$\beta\ (t) = \varepsilon\alpha\ (t) \quad (5)$$

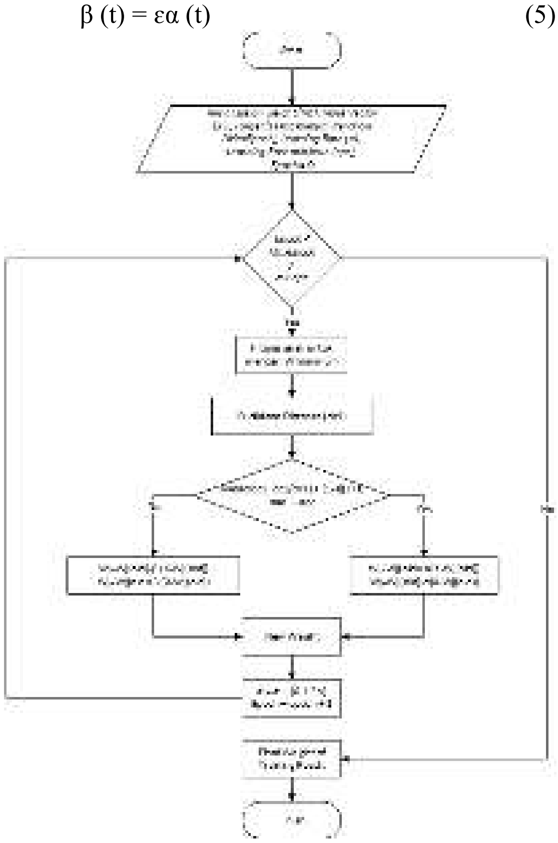

Figure 3. Flow Chart Of Learning (Training) LVQ3

3. RESEARCH METHODS

Research methodology used in this research is quantitative research. Methods used in this research report writing using two methods, i.e., methods of data collection and methods of software development.

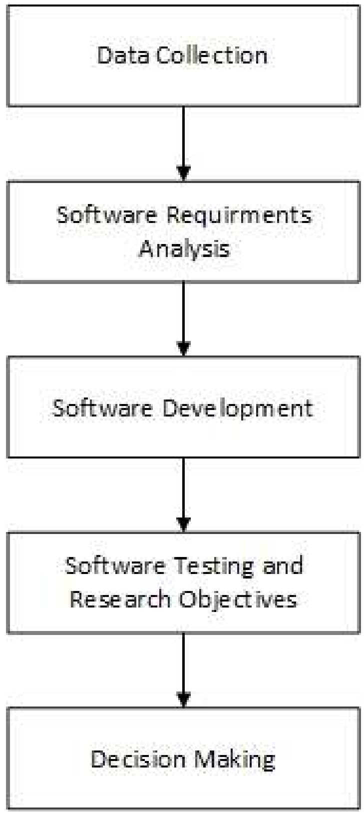

Figure 4. Research Methods

4. RESULTS AND DISCUSSION

4.1 Overview of The System

An overview of the system to be built can be seen in Figure 5.

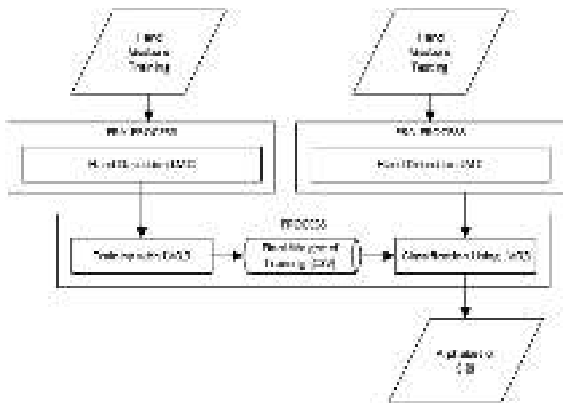

Figure 5. Overview of The System

The construction of the system of fingerspelling was going through several stages. Stages of the making of the new weighting for further classification to be used against test data contained in Figure 6

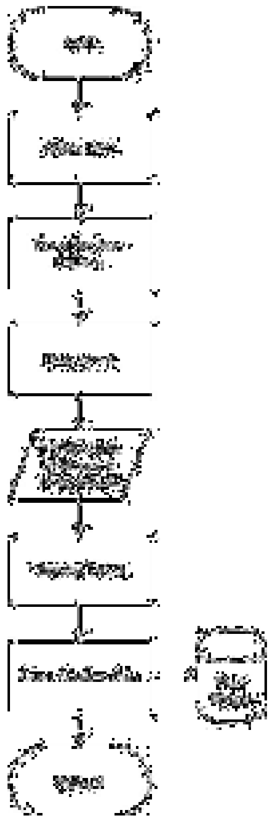

Figure 6. Classification Of Test Data

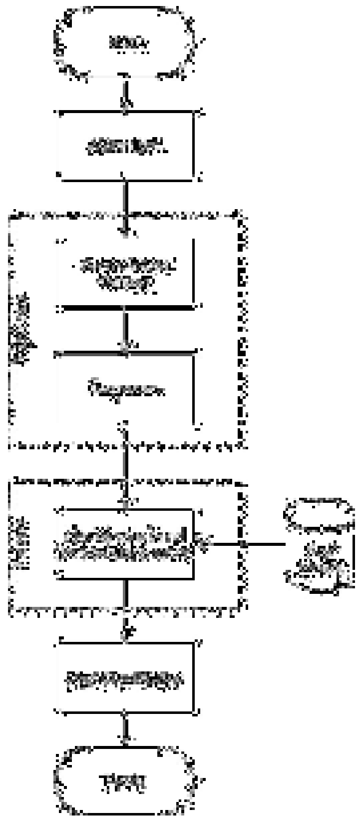

Figure 7. Classification Of Hand Signals

Based on the pictures, making a new weighting system of the system to be built, the main processes, namely the user enters the hand gesture of the LMC and label appropriate gesture taken i.e. the letters of the alphabet the alphabet. Existing processes on the stage of the creation of a new weights are:

a. the user enable LMC
b. Capture a hand gesture
   Retrieval of hand gesture media input into a system using the LMC.
c. Praprocess
   Proceedings before the main proceedings, namely done cropping, grayscalling, thresholding, and conversion matrix on the image input. All this process in a skip because LMC has already done the calculation with praprocess are not shown.
d. hand Gesture taken first made weight. Second hand gesture and subsequently made a data training.
e. the process of training using LVQ3
   The process of training training data against data weighting that will produce a new weighting to be used for the classification against test data.
f. the process of storing data hand-gesture
   The hand gesture is captured using a stored file. CSV.

Stage classification of hand signals in the system there is in Figure 7.

Based on the images on hand signals classification system is divided into stages and praprocess process. The main processes namely users enable LMC later to capture the hand gesture in accordance with iyarat hands.

The first phase of that is praprocess. Existing processes in this stage are:

a. the user enable LMC
b. Capture gesture of the hand.
c. Praprocess.

The next process is the main process. The process of classification of hand signals using LVQ3, the classification data of the previous praprocess results with the results of weighting training data LVQ3 already stored in a file. CSV.

4.2 Pra-process

Pre-process is the process of data capture for all in training by LVQ3 algorithm. Data retrieved by performing data retrieval from LMC. Proper hand position is in front of LMC accompanied not use accessories such as rings, bracelets or other accessories worn hands as in Figure 8.

Figure 8. Capture data from LMC

In data retrieval from LMC calculation are not listed because the data is already taken in the form of a vector. The large number of input vectors is as much as 6, namely palm position (palm1, palm2, palm3) and z (hp) axis, x axis (roll), and the y axis (yaw). All these vector obtained because LMC work system is a form of room sensor as shown in Figure 1.

Table 1. LMC Vector Table

No	Palm1	…	Roll	Class
1	-30.3971	…	-26.8880	I
2	-35.3162	…	-6.0653	J1
3	-37.7638	…	-71.7844	J2
4	-14.7153	…	-49.6420	J3
5	-29.6046	…	-16.0272	Z1
6	-10.0783	…	13.4136	Z2
7	-34.1530	…	-35.6873	Z3
8	-21.4457	…	-26.0776	Z4

In order to characterize the letter moves, this research uses only 4 main frame in its application. All the frames were taken manually, which consists of the frame I, J1, J2, J3 and for the letter J and for the letter Z is Z1, Z2, Z3, and Z4.

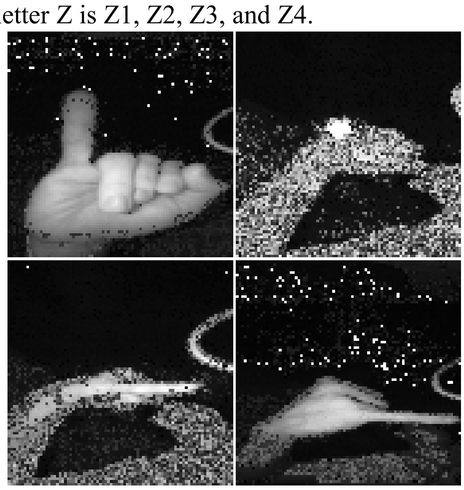

Figure 9. The Combination Of Letter J

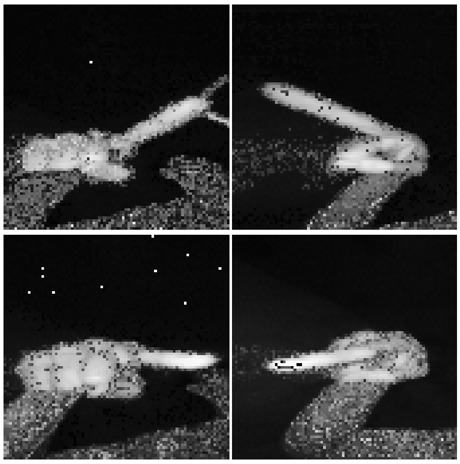

Figure 10. The Combination Of Letter Z

4.3 Training

In this phase of training using LVQ3 against identity letter. The number of neurons has 6 input neurons, which consists of the vector palm1, palm2, palm3, hp, roll, yaw. Meanwhile, the output layer has twenty-six neurons refers to the architecture of the LVQ LVQ Architecture in Figure 11 for the classification system can be seen in Figure 3.5.

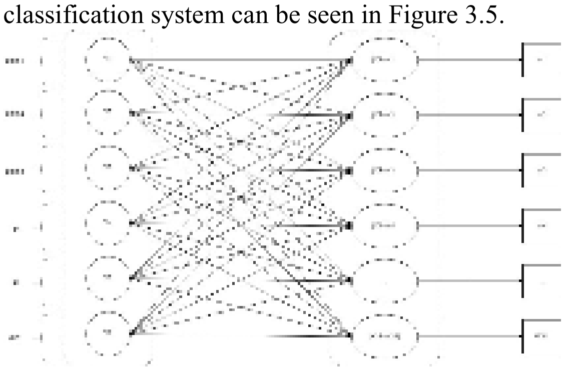

Figure 11. The Architecture Of The LVQ

Based on previous research, set reference rate of learning (α) of 0.01 with a reduction in each iteration 0.1 * α, rate the minimum study (eps) amounted to 0.001, maximum epoch 1000 [4], the value of the window ε 0.3. Note 31 input vectors in twenty-six class accordingly.

Table 2. The Vector Of Weights (W)

No	Palm1	…	Roll	Class
1	-30.3971	…	-26.8880	I
2	-35.3162	…	-6.0653	J1
3	-37.7638	…	-71.7844	J2
4	-14.7153	…	-49.6420	J3
5	-29.6046	…	-16.0272	Z1
6	-10.0783	…	13.4136	Z2
7	-34.1530	…	-35.6873	Z3
8	-21.4457	…	-26.0776	Z4

Table 3. Vector Data Training

No	Palm1	…	Roll	Class
1	-7.0845	…	-31.9043	I
2	-18.0686	…	-31.6101	I
3	-23.6199	…	-29.1312	I
4	-11.2545	…	-30.142	I
5	-4.8634	…	-29.303	I
6	-13.9992	…	-30.5507	I
7	-9.0316	…	-29.4871	I
8	1.6348	…	-26.0745	I
9	11.5798	…	-26.3806	I
10	-17.5413	…	-28.7256	I
11	-37.7202	…	-25.6646	I
12	-43.2258	…	-29.0739	I
13	-25.3187	…	-24.7178	I
14	-1.9447	…	-20.8329	I
15	1.9808	…	-20.1757	I
…	…	…	…	…
170	523.27	…	123.9453	Z4

Table 4. The Vector Of Weights Training Results

Bobot	Palm1	…	Roll
W1	-30.4234	…	-26.9224
W2	-30.3708	…	-26.8536
W3	-37.7638	…	-71.7842
W4	-14.7153	…	-49.6420
W5	-29.5731	…	-16.0195
W6	-10.0728	…	13.3032
W7	-34.1291	…	-35.6737
W8	-21.4514	…	-25.9679

4.4 Classification

At the stage of classification data input in the form of a vector result of pre-process calculation is done using the weighted end LVQ3 learning outcomes.

Known vectors of input from J1 test data: gesture

-14.7233 88.6025 … -52.1370

After calculating the distance between the input vector with weights, the smallest distance can be found on the weights, so that the pattern on the class to the J1 that is class.

4.5 Detection Dynamic Fonts

This stage is the stage of detection of dynamic letter, namely the letter J and z. Groove detection for the letter J and Z is contained in Figure 12.

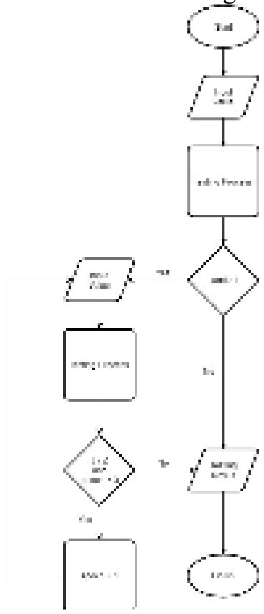

Figure 12. The Dynamic Letters Detection Process

The dynamic letters detection phases in Figure 3.6. IE:

1. Enter the input values or test data
2. Testing using LVQ3
3. is unreadable I or Z1 If yes then go into the next process, if not then displays the test results
4. If Yes, then place the return value or input test data
5. Test process using LVQ3
6. If the J or Z2 and epoch < 3 if yes epoch + 1, when not showing test results
7. when conducting a test to detect the letter J/Z or epoch > 3

4.6 Testing

The testing method is the performance testing phase the methods implemented to build the system introduction cue. The method will be tested is a method of Learning Vector Quantization 3. The method is examined by using 8 data training and test data. These details can be seen in table 5.

Table 5. Testing

No	Types Of Cues	The Amount Of Training Data	The Number Of Test Data
1	I	3	3
2	J1	3	3
3	J2	3	3
4	J3	3	3
5	Z1	3	3
6	Z2	3	3

No	Types Of Cues	The Amount Of Training Data	The Number Of Test Data
7	Z3	3	3
8	Z4	3	3
9	Not The Gesture Cues	0	3

1. Test Scenario 1

On the test 1 this is done manually or not in real time, the data used is the result of data capture gesture manually

Table 6. Confusion Matrix Testing 1

		Predicted Class	
		Positif	Negatif
Actual Class	Positif	7	17
	Negatif	0	2

After performing the classification, calculate the value of accuracy using equation 1.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

$$= \frac{7+2}{7+17+2+0}$$

$$= \frac{9}{26} = 0{,}34$$

Calculate the value of precision and recall by using the equation.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$= \frac{7}{7+17}$$

$$= \frac{7}{24} = 0{,}29$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$= \frac{7}{7+0}$$

$$= \frac{7}{7} = 1$$

Testing accuracy using test data as much as 26, consisting of 8 which includes the sign-language gesture gesture and 2 not including sign language. Classifications generated by the system as much as 7 sign language which are classified correctly, not classified 17 right at sign language and 2 classified is not sign language appropriately, then the proper classification number is 20 sign language .

Based on the testing accuracy, accuracy of results obtained by 34% with a precision of 29% and 100% recall.

1. Test Scenario 2

On testing 2 is done in real time using the LMC, the data examined i.e. 3 data.

Table 7. Confusion Matrix Testing 2

		Predicted Class	
		Positif	Negatif
Actual Class	Positif	12	12
	Negatif	0	2

After performing the classification, calculate the value of accuracy using equation.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

$$= \frac{12+2}{12+12+2+0}$$

$$= \frac{14}{26} = 0{,}54$$

Hitung nilai precision dan recall dengan menggunakan persamaan.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$= \frac{12}{12+12}$$

$$= \frac{12}{24} = 0{,}50$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$= \frac{12}{12+0}$$

$$= \frac{12}{12} = 1$$

Testing accuracy using test data as much as 26, consisting of 8 which includes the sign-language gesture gesture and 2 not including sign language. Classifications generated by the system as much as 7 sign language which are classified correctly, not classified 17 right at sign language and 2 classified is not sign language appropriately, then the proper classification number is 20 sign language .

Based on the testing accuracy, accuracy of results obtained by 54% with a precision of 50% and 100% recall.

2. Test Scenario 3

Testing conducted with the mengujikan some cues for can recognize letters moving, in this case the letter J and Z. Following the test results on the table 8.

Table 8. The test scenario 3

No	Test Parameters	Output	The Test Results	
			Factual Classification	Classification Systems
1	J	I,J2,J2,J2	Positif	Negatif
2	J	I,J1,J2,J3	Positif	Positif
3	J	I,J3,J2,J3	Positif	Negatif
4	J	I,J1,J2,J3	Positif	Positif
5	J	I,J3,J2,J2	Positif	Negatif
6	Z	Z1,Z2,Z3, Z4	Positif	Positif

No	Test Parameters	Output	The Test Results	
			Factual Classification	Classification Systems
7	Z	Z1,Z4,Z3, Z4	Positif	Negatif
8	Z	Z1,Z2,Z3, Z4	Positif	Positif
9	Z	Z1,Z2,Z3, Z4	Positif	Positif
10	Z	Z1,Z2,Z3, Z4	Positif	Positif

Below is a table of confision matrix:

Table 9. Confusion Matrix Testing 3

		Predicted Class	
		Positif	Negatif
Actual Class	Positif	6	4
	Negatif	0	0

After performing the classification, calculate the value of accuracy using equation.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}}$$

$$= \frac{6+0}{6+4+0+0}$$

$$= \frac{6}{10} = 0{,}6$$

Calculate the value of precision and recall by using the equation.

$$\text{Precision} = \frac{\text{TP}}{\text{TP} + \text{FP}}$$

$$= \frac{6}{6+4}$$

$$= \frac{6}{10} = 0{,}6$$

$$\text{Recall} = \frac{\text{TP}}{\text{TP} + \text{FN}}$$

$$= \frac{6}{6+0}$$

$$= \frac{6}{6} = 1$$

Based on the testing accuracy, accuracy of results obtained by 60% with a precision of 60% and 100% recall. Then note that the system less recognize the letter J and Z with accurate.

2. CONCLUSION

Based on the results of the research, analysis and design, implementation as well as to the stages of testing, can be drawn the conclusion that the English system of fingerspelling gestures using Learning Vector Quantization 3 produce accuracy testing of 54%. Suggestions for future research vector characteristics taken from each frame needs to be additions to the number of the lot and number of frames for adding detection of the moving letter in order to be more accurate and accuracy of detection of letters moving into more effective and efficient.

REFERENCES

[1] Wildan Fauzan, Nelly Indriani W, “Finger Spelling untuk kata Khusus dengan Bahasa Isyarat Tangan Menggunakan Metode Viola and Jones,” Teknik Informatika, p. 8, 2017.

[2] N. E. Sihbudi, “Implementasi Metode Naive Bayesian untuk Menerjemahkan Bahasa Isyarat Indonesia,” Teknik Informatika, p. 8, 2017.

[3] Wijayanti Nurul Khotimah, Risal Andika Saputra, Nanik Suciati, Ridho Rahman Hariadi, “Alphabet Sign Language Recognition Using Leap Motion Technology and Rule Based Backpropagation-Genetic Algorithm Neural Network (RBBPGANN),” Jurnal Ilmiah Teknologi Informasi, vol. 15, p. 95, 2017.

[4] Gilang Ramadhan, Esmeralda C Djamal, Tedjo Darmanto, “Klasifikasi Identitas Wajah Untuk Otoritasi Menggunakan Deteksi Tepi dan LVQ,” Seminar Nasional Aplikasi Teknologi Informasi (SNATi), pp. D-37, 2016.

[5] Elvia Budianita, Widodo Prijodiprodjo, “Penerapan Learning Vector Quantization (LVQ) untuk Klasifikasi Status Gizi Anak,” vol. 7, no. 2, p. 155, 2013.

[6] Kusumadewi S, Hartati S, Neuro Fuzzy : Integrasi Sistem Fuzzy dan Jaringan Syaraf, Yogyakarta: Graha Ilmu, 2006.

[7] L. Fausett, Fundamentals of Neural Networks ; Architectures, Algorithms, and Applications, Pretince Hall: Englewood Cliffs, 1994.

[8] Kamus, “Sistem Isyarat Bahasa Indonesia,” Jakarta: Departemen Pendidikan Nasional, 2010.

[9] L. Motion, “Leap Motion Documentation,” Leap Motion, 11 February 2015. [Online]. Available: developer-archive.leapmotion.com. [Diakses 1 January 2019].

[10] M. D. Wuryandari dan I. Afrianto, “Perbandingan Metode Jaringan Syaraf Tiruan Backpropagation dan Learning Vector Quantization pada Pengenalan Wajah,” KOMPUTA, vol. 1, pp. 45-51, 2012.