Sign Language A-Z Alphabet Introduction American Sign Language using Support Vector Machine

Muhammad Rasuandi [a,1], Muhammad Fachrurrozi [b,2*], Anggina Primanita [b,3]

[a] Informatics Engineering, Faculty of Computer Science, Universitas Sriwijaya, Indonesia
[1] andremrasuandi@gmail.com*; [2] mfachrz@unsri.ac.id; [3] anggina.primanita@ilkom.unsri.ac.id
* corresponding author

A R T I C L E I N F O

Article history
Received 2023-05-11
Revised 2023-05-31
Accepted 2023-08-23

Keywords
Sign Language
American Sign Language
Support Vector Machine

ABSTRACT

Deafness is a condition where a person's hearing cannot function normally. As a result, these conditions affect ongoing interactions, making it difficult to understand and convey information. Communication problems for the deaf are handled through the introduction of various forms of sign language, one of which is American Sign Language. Computer Vision-based sign language recognition often takes a long time to develop, is less accurate, and cannot be done directly or in real-time. As a result, a solution is needed to overcome this problem. In the system training process, using the Support Vector Machine method to classify data and testing is carried out using the RBF kernel function with C parameters, namely 10, 50, and 100. The results show that the Support Vector Machine method with a C parameter value of 100 has better performance. This is evidenced by the increased accuracy of the RBF C=100 kernel, which is 99%.

1. Introduction

Deafness is a condition where a person's hearing cannot function normally. As a result, these conditions affect ongoing interactions, making it difficult to understand and convey information. Existing research indicates that 360 million people (328 million adults and 32 million children) are deaf and dumb, or more than 5% of the global population. [1]

The communication problems of the deaf are addressed through the introduction of various forms of sign language. However, most deaf people have difficulty adapting the hand code rules in sign language from the expressions of direct body movements at first

American Sign Language (ASL) is one of the most widely used sign languages in the world. In fact, American Sign Language is used in collaboration with the Indonesian Sign System. ASL is expressed with one hand movement, this is what makes ASL often chosen to facilitate the introduction of the alphabet in sign language communication.

Computer Vision-based sign language recognition often takes a long time to develop, is less accurate, and cannot be done directly or in real-time. As a result, a solution is needed to overcome this problem. In the system training process, the Support Vector Machine (SVM) method produces high accuracy. Because it can be done in Real-Time. The SVM method itself has two sessions for the introduction process, namely training and introduction. [2]

The purpose of this research is to identify the American Sign Language (ASL) alphabet from A to Z. With 26 ASL alphabet labels used. Computer Vision approach was used in conducting this research.

2. Literature Study

A. American Sign Language

Deaf people can communicate with each other through sign language. Because this system can be used as a communication tool to understand the intent of the interlocutor by recognizing alphabetic patterns formed by the fingers as signs in word order, Sign Language is useful in everyday life. An example of sign language is American Sign Language. All of the natural languages that make up American Sign Language were spoken by hand and share many of the same linguistic features as spoken languages. You can demonstrate this type of ASL with only one hand, both left and right. [3]

Fig. 1. American Sign Langage

B. Support Vector Machine

Support Vector Machine is an algorithm classification that can classify data non-linearly. [4] SVM works by defining the limit between classes through the creation of a hyperplane with a maximum margin. [5] Margin is the distance between the hyperplane and the nearest data of each class. The general hyperplane function is shown in Equation 1.

$$w_i . x_i + b = 0 \qquad (1)$$

If $\boldsymbol{w_i . x_i + b = +1}$ is a supporting hyperplane of class +1 (positive) and the equation $\boldsymbol{w_i . x_i + b = -1}$ is a supporting hyperplane of class -1 (negative) then the margin can be calculated by finding the distance of the two hyperplanes supporting the two classes. The margin value is defined as follows.

$$\frac{2}{||w||} \qquad (2)$$

To determine the best hyperplane, the Quadratic Programming (QP) problem formula is used, which is a form of optimization using the inverse of the following equation.

$$\min \frac{1}{2} ||w||^2 \qquad (3)$$

Where $(y_i(wx_i) + b - 1 \geq 0$ and the limitations functions can be written as follows.

$$\sum_{i=1}^{n} a_i \, [y_i \, (wx_i) + b - 1] \qquad (4)$$

Optimization of Quadratic Programming problems can be solved with the Lagrange Multiplier method defined through the following function.

$$L(\vec{w}, b, a) = \frac{1}{2}\|w\|^2 - \sum_{i=1}^{n} a_i (y_i (\vec{x}_i . \vec{w} + b) - 1)) \quad (5)$$

With a_i is the value of Lagrange multipliers with a value of zero or positive ($a_i \geq 0$). The optimum value of Lagrange multipliers can be calculated by minimizing the variables $\vec{w}$ and b and maximizing the a_i variables. This the equations can be written as follows.

$$\sum_{i=1}^{n} (a_i y_i x_i) = 0 \quad (6)$$

The Lagrange function in (5) is converted into (7) by substituting (6) into (5)

$$L(a) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j \vec{x}_i . \vec{x}_j \quad (7)$$

The determination of the best separation of fields is written as follows.

$$\max (L(a)) = \sum_{i=1}^{n} a_i - \frac{1}{2} \sum_{i,j=1}^{n} a_i a_j y_i y_j \vec{x}_i . \vec{x}_j \quad (8)$$

Based on the function of the obstacle, equation (8) becomes,

$$w = \sum_{i=1}^{n} (a_i y_i x_i) \quad (9)$$

So the value of b is obtained as follows,

$$b = \frac{1}{2} \hat{w}. [x_r + x_s] \quad (10)$$

Where x_r and x_s are support vectors that comply with each class. The decision function obtained based on (9) and (10) is written as follows.

$$f(x) = sgn (\hat{w}. x_i + b) \quad (11)$$

C. Evaluation

Evaluation is a process to measure the performance of the method being tested. In measuring the performance of the classification process, four standard calculations can be used, namely accuracy, precision, recall, and f-measure.[6]

1. Accuracy is a comparison between the data that is clarified correctly and the overall data

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

2. Precision is the ratio of the amount of data of the correctly classified positive category to the total data classified as positive

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

3. Recall, the success rate of the software in separating similar documents into the same classification

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

4. F-measure, the average ratio of Precision and Recall values

$$F - measure = 2 \times \frac{Precision \times Recall}{Precision + Recall}$$

The parameters in these calculations are obtained from the Confusion Matrix table. A confusion Matrix is a method used to review classification results obtained by calculating the frequency of truth. [7]

Table 1. Confusion Matrix

		Prediction	
		Negative	Positive
Actual Class	Negative	TN	FP
	Positive	FN	TP

Where TN (*True Negatives*) shows the value obtained when the model can classify negative data into negative, TP (*True Positives*) shows the value obtained when positive data is classified as positive, FP (*False Positives*) shows the value obtained when the model classifies negative data into positive, and FN (*False Negatives)* shows the value that obtained when positive data is classified to negative. [8]

3. Research Method

A. Data Collection

26 pictures of sign language alphabets taken via webcam from the researcher's own hands which serve as primary data with as many as 2,600 data (100 data for each sign). The data is divided into 2080 training data and 520 test data.

Data is taken by hand pointing towards the front of the webcam and without any obstruction. The image format is (.png).

B. Framework

The research phase consisted of acquiring sign language such as motion picture data, developing a research framework, setting test criteria, developing test data formats, selecting tools to use, conducting research tests, evaluating test results, and formulating research conclusions. [9] We can see the stages of developing research that have down are shown in Figure 2.

Based on Figure 2, starting from text pre-processing, In the process of converting video to frames, the video that has been recorded in the program will be converted into frames per second in .jpeg format. The test data is used to place the model being trained during the training process through its steps. Each test measures the precision, recall, accuracy and F1-Score of each label. Then splitting the data which divides into training data and testing data with a split of 80% and 20%.

After performing the preprocessing, the TF-IDF weighting procedure is performed, which is useful for converting words to numbers. Next, the TF-IDF weighted data is classified using SVM, thereby creating a training model. After that, the final results are evaluated in the form of a confusion matrix including accuracy, precision, recall, and f1-score [10].

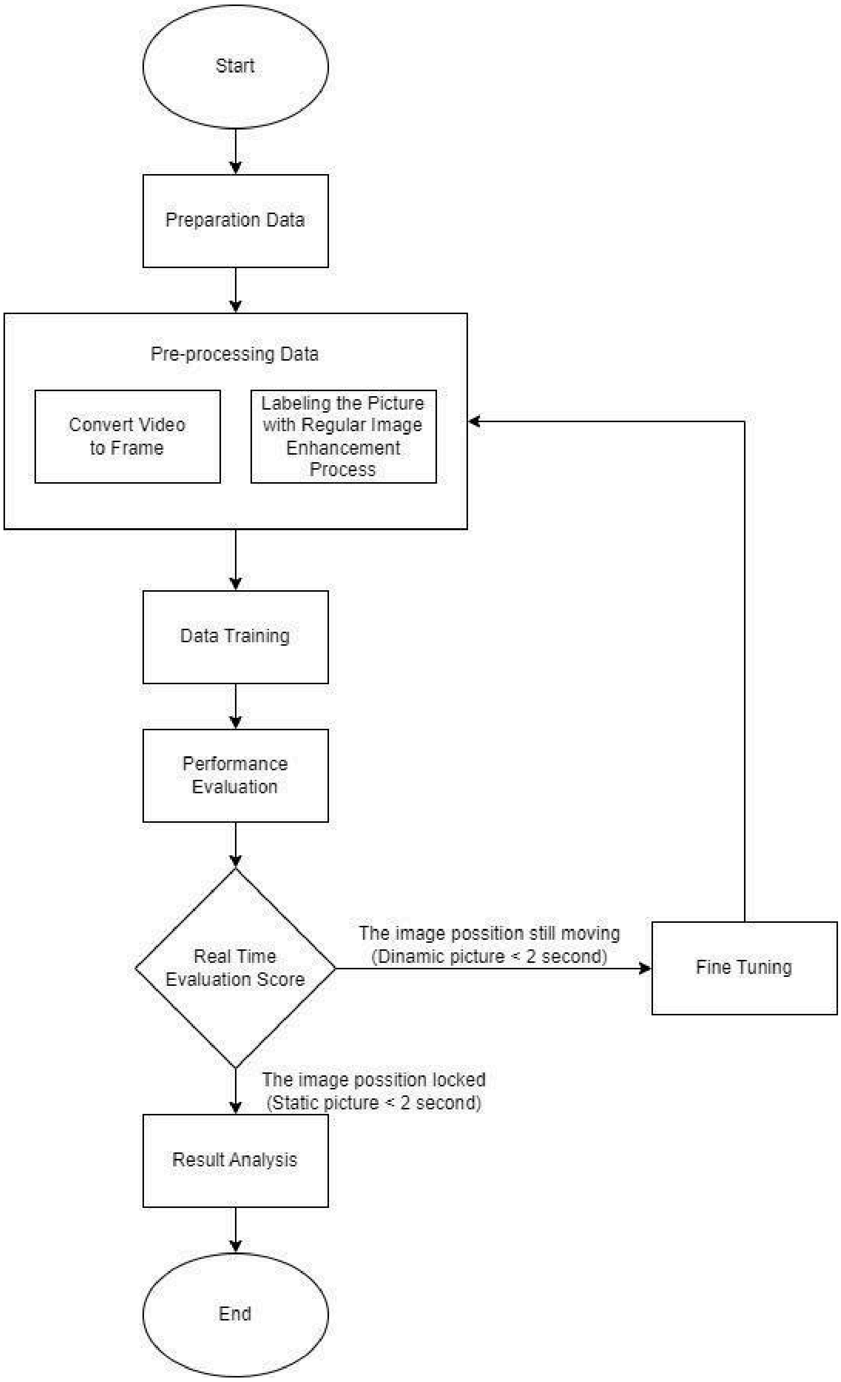

Fig. 2. Framework

4. Result and Discussion

Training on the primary dataset is carried out by dividing the dataset into two parts, namely test data and train data randomly using the sklearn library with the train_test_split function with a ratio of 80: 20. The training process is carried out using the Support Vector Machine method by testing hyperplane line parameters / constant values C is 10, 50, and 100, the kernel approach is "rbf', and the gamma value is 10%. After the training process is complete, the results can be seen in the array that holds it.

A. The result of testing

The result of testing based on the process that has been done, obtained a confusion matrix as shown in Figure 3 - 5.

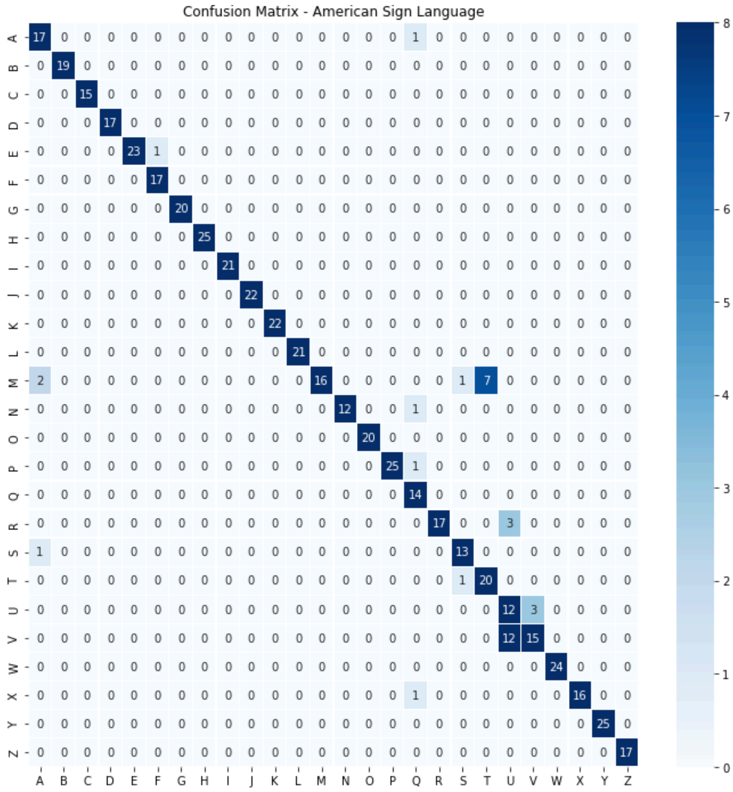

Fig 3. Result of Confusion Matrix (C = 10)

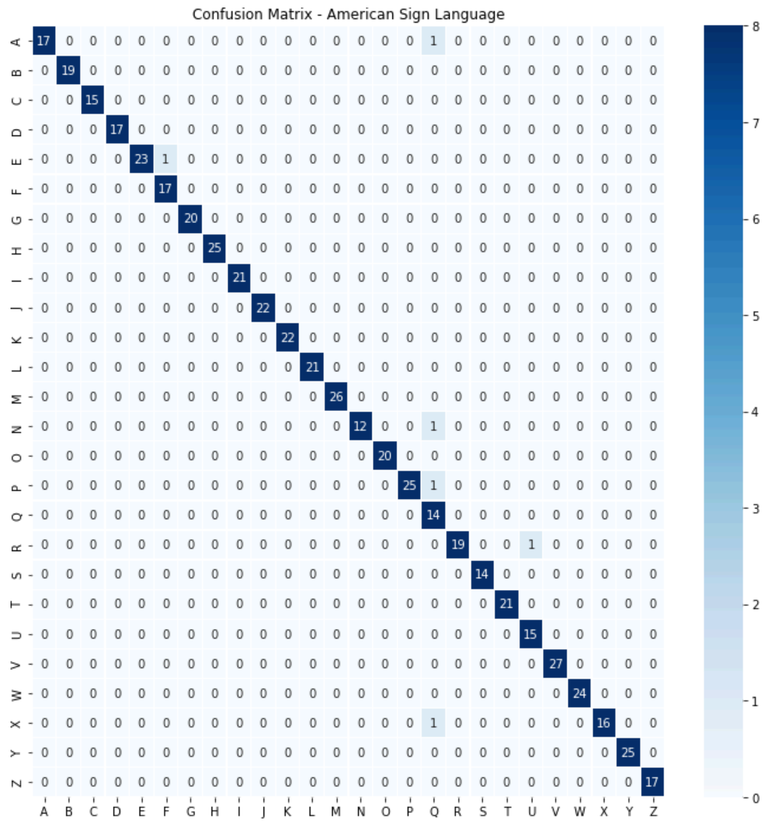

Fig 4. Result of Confusion Matrix (C = 50)

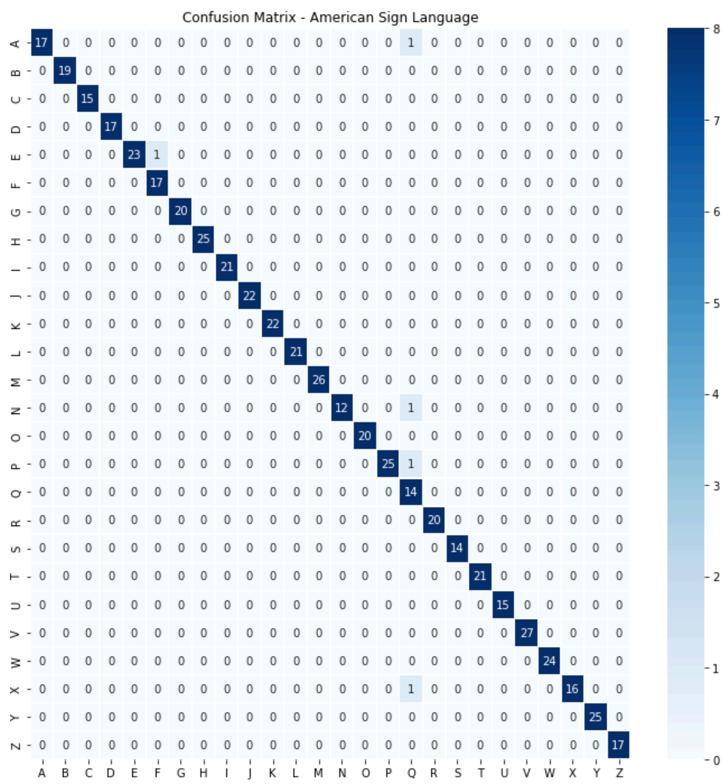

Fig 5. Result of Confusion Matrix (C = 100)

Table 2-4 displays the results of the Confusion Matrix calculations for the best model of the SVM classification procedures for each C varying parameters. Furthermore, the Confusion Matrix results from the best fold will be used to calculate the evaluation value. The results are compared to the comparison table on each C parameter score.

Table 2. Evaluation Data from C = 10

No.	Label	TP	TN/FP/FN	Precision	Recall	F1-Score	Accuracy
1.	A	17	1	0.85	0.94	0.89	0.89
2.	B	19	0	1.00	1.00	1.00	1.00
3.	C	15	0	1.00	1.00	1.00	1.00
4.	D	17	0	1.00	1.00	1.00	1.00
5.	E	23	1	1.00	0.96	0.98	0.98
6.	F	17	0	0.94	1.00	0.97	0.97
7.	G	20	0	1.00	1.00	1.00	1.00

Table 3. Evaluation Data from C = 50

No.	Label	TP	TN/FP/FN	Precision	Recall	F1-Score	Accuracy
1.	A	17	1	1.00	0.94	0.97	0.97
2.	B	19	0	1.00	1.00	1.00	1.00
3.	C	15	0	1.00	1.00	1.00	1.00
4.	D	17	0	1.00	1.00	1.00	1.00
5.	E	23	1	1.00	0.96	0.98	0.98
6.	F	17	0	0.94	1.00	0.97	0.97
7.	G	20	0	1.00	1.00	1.00	1.00

Table 4. Evaluation Data from C = 100

No.	Label	TP	TN/FP/FN	Precision	Recall	F1-Score	Accuracy
1.	A	17	1	1.00	0.94	0.97	0.97
2.	B	19	0	1.00	1.00	1.00	1.00
3.	C	15	0	1.00	1.00	1.00	1.00
4.	D	17	0	1.00	1.00	1.00	1.00
5.	E	23	1	1.00	0.96	0.98	0.98
6.	F	17	0	0.94	1.00	0.97	0.97
7.	G	20	0	1.00	1.00	1.00	1.00

The table above is a comparison of the results of the evaluation of the SVM method classification which is divided into 3 different score of C parameters.

5. Conclusion

The highest accuracy value of 99% is obtained by using the constant value C = 100, compared to C = 10 or C = 50. The precision values obtained are different because there are similarities between several sign language alphabets such as between the alphabet A & M, the alphabet U & V, and the fewest errors are found at a constant value of C = 100.

This is the reference for researchers to choose the constant value to get the highest accuracy value compared to the comparison of other constant values.

Real-time input is used to run test scenarios. Precision, recall, accuracy, and F1-score were all found to have an average value of 0.99, or 99%, in trials conducted for research.

References

[1] G. R. Mauludi and A. Setiyadi, "Pembangunan Alat Penerjemah Huruf dan Angka Bahasa Indonesia bagi Tunarungu dan Tunawicara menggunakan Arduino," 2019.

[2] I. Nadia, "Jurnal Mahasiswa Universitas Muhammadiyah Ponorogo," *Jurnal Mahasiswa Universitas Muhammadiyah Ponorogo*, 2018.

[3] M. E. Al Rivan, H. Irsyad, K. Kevin, and A. T. Narta, "Pengenalan Alfabet American Sign Language Menggunakan K-Nearest Neighbors Dengan Ekstraksi Fitur Histogram Of Oriented Gradients," *Jurnal Teknik Informatika dan Sistem Informasi*, vol. 5, no. 3, Jan. 2020, doi: 10.28932/jutisi.v5i3.1936.

[4] A. Rozani, "PENERAPAN METODE JARINGAN SYARAF TIRUAN PADA APLIKASI PENGENALAN BAHASA ISYARAT ABJAD JARI," 2017.

[5] P. A. Octaviani, Y. Wilandari, and D. Ispriyanti, "PENERAPAN METODE KLASIFIKASI SUPPORT VECTOR MACHINE (SVM) PADA DATA AKREDITASI SEKOLAH DASAR (SD) DI KABUPATEN MAGELANG," vol. 3, no. 4, pp. 811–820, 2014, [Online]. Available: http://ejournal-s1.undip.ac.id/index.php/gaussian

[6] T. M. Kadarina et al, "Pengenalan Bahasa Pemrograman Python menggunakan Aplikasi Games untuk Siswa/i di Wilayah Kembangan Utara," 2019. [Online]. Available: https://codecombat.com/.

[7] I. Monika Parapat and M. Tanzil Furqon, "Penerapan Metode Support Vector Machine (SVM) Pada Klasifikasi Penyimpangan Tumbuh Kembang Anak," 2018. [Online]. Available: http://j-ptiik.ub.ac.id

[8] A. R. Shetty, F. B. Ahmed, and V. M. Naik, "CKD Prediction Using Data Mining Technique As SVM And KNN With Pycharm," *International Research Journal of Engineering and Technology*, p. 4399, 2008, [Online]. Available: www.irjet.net

[9] S. Wulan Purnami and W. wibowo, "PARAMETER OPTIMIZATION OF SUPPORT VECTOR MACHINE USING TAGUCHI APPROACH FOR HIGH-DIMENSIONAL DATA," 2017.

[10] R. I. Borman, B. Priopradono, and A. R. Syah, *Klasifikasi Objek Kode Tangan pada Pengenalan Isyarat Alphabet Bahasa Isyarat Indonesia (Bisindo)*. 2017.