Indonesian Sign Language (BISINDO) Recognition Using Accurate and Fast Dynamic Time Warping Learning Model

Tri Handhika, Dewi Putrie Lestari, Ilmiyati Sari, Revaldo Ilfestra Metzi Zen, and Murni

Abstract—**Sign language recognition problem should be represented as a time series classification model with high accuracy. In the previous studies, Indonesian sign language (BISINDO) had been modeled with one of stochastic time series classification model, i.e. Hidden-Markov Model (HMM), but has low accuracy. In other studies, BISINDO had been recognized with high accuracy but using an unrepresentative model (non-time series classification model), i.e. the modified Generalized Linear Vector Quantization (mGLVQ) model with mode function. In this paper, we tried to use a deterministic time series classification model, named Accurate and Fast Dynamic Time Warping (AF-DTW) model. AF-DTW model is a modified form of Dynamic Time Warping (DTW) model. It is not only to improve the accuracy of DTW but also to accelerate the finding of optimal warping path. The output results showed that AF-DTW has a much higher accuracy than HMM, although it is not as accurate as mGLVQ.**

Index Terms—**BISINDO, dynamic time warping, sign language recognition, time series classification.**

I. INTRODUCTION

Humans communicate with each other using language. Different with ordinary people who use spoken language to communicate, the deaf use sign language because of their limitations in speaking and listening. Sign language uses hand gestures and facial expressions.

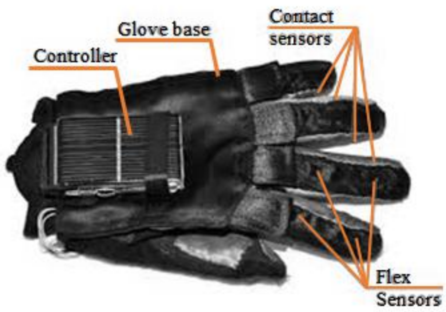
Sign language recognition is one of more popular research topics in computer vision. It is used to solve the communication problems between the deaf and ordinary people who does not know about sign language by creating a software application that can translate sign language into a spoken/written language [1]. Several studies have been conducted to build an automatic sign language translator through computer vision technology based on gesture recognition such as Microsoft Glove and Kinect XBox (see Fig. 1).

Manuscript received October 31, 2019; revised January 19, 2020. This work was supported in part by the Ministry of Research, Technology, and Higher Education, Indonesia to the Gunadarma University under PUPT 2018 Research Grant No. 023/KM/PNT/2018.

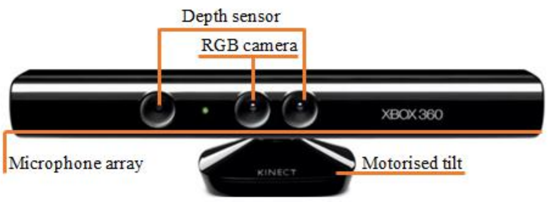
Tri Handhika, Dewi Putrie Lestari, Ilmiyati Sari, and Murni are with the Center for Computational Mathematics Studies, Gunadarma University, 100 Margonda Raya St., Pondok Cina, Depok, West Java, 16424, Indonesia (e-mail:trihandika@staff.gunadarma.ac.id, dewi_putrie@staff.gunadarma.ac.id, ilmiyati@staff.gunadarma.ac.id, murnipskm@staff.gunadarma.ac.id).

Revaldo Ilfestra Metzi Zen is with the Big Data Division, Metra Digital Media, Wisma Aldiron Dirgantara 2nd Floor Suite 202-209, South Jakarta 12780, Indonesia (e-mail: revaldo.ilfestra@mdmedia.co.id).

In this paper, we used Microsoft Kinect Xbox [2]–[5] as a data acquisition tool because it is equipped with various sensor features that can receive multi-modal inputs such as gestures of shoulders, upper arms, forearms, hands, fingers, and face [6] not only fingers gestures like those acquired by gloves.



(a)



(b)

Fig. 1. (a) Glove (b) Kinect.

After the sign language gestures data has been recorded, it is recognized by some of classification models. However, the chosen classification model needs to represent the sign language that expressed as a sequence of gestural patterns to convey a meaning of words at uniform time of intervals [7]. The sign language gestures data in the form of sequence of frames is classified as time series data such that it needs to be represented by time series classification model [8]. In recent decades, large number of studies have focused on the development of time series classification models where it can be roughly divided into two approaches, i.e. the deterministic and stochastic approaches that represent conditions under certainty and uncertainty, respectively.

There are two kinds of sign languages applied in Indonesia, i.e. Sistem Isyarat Bahasa Indonesia (SIBI) [9], [10] and Bahasa Isyarat Indonesia (BISINDO) [1], [6]. We tried to continue our research by setting BISINDO recognition as our research objectives because of its gestures are more natural and practical for the deaf or hard-hearing people [6]. In the previous studies, BISINDO has been modeled with one of stochastic time series classification model, i.e. Hidden-Markov Model (HMM), but has low accuracy [6]. In other studies, BISINDO has been recognized with high

doi: 10.18178/ijmlc.2020.10.2.946

accuracy but using an unrepresentative model (non-time series classification model), i.e. the modified Generalized Linear Vector Quantization (mGLVQ) model with mode function [1]. Based on these results, we tried to find a representative time series classification model with high accuracy. And finally, we decided to use one of deterministic time series classification model, named Dynamic Time Warping (DTW) model [11]–[14].

DTW model has recently been widely used to measure similarity between two time series by finding an optimal warping path in a cumulated matrix. DTW model suffers from high computational costs and space complexity when it works for large time series datasets. However, it is still potential to be a feasible answer to find an optimal warping path by providing a time series similarity measurement which is flexible and easy to interpret [13]. To improve the computational efficiency and the accuracy of similarity measurement, Li and Yang modified the DTW model into an Accurate and Fast Dynamic Time Warping (AF-DTW) model [15].

AF-DTW model is divided into three parts i.e. backward strategy, reduced scope process, and threshold value determination. Different with DTW model which found the optimal warping path by using forward strategy, AF-DTW model use backward strategy instead. Moreover, to narrow the warping path scope, it reduces the scope and choose a method to determine the threshold value. In the implementation, backward strategy and reduced scope process are run concurrently.

AF-DTW model is known as an optimization model for aligning two time series data which may vary in time or speed. In the AF-DTW model, sign language gestures data are compared one-by-one for each sign language gestures record in the database. The time series classification result is a word of BISINDO with warping cost below a certain threshold within the test sequence [12]. In other words, AF-DTW model is used to optimize the similarity between sign language gestures records from a sign language gestures database [11]. The detail how to implement AF-DTW model on the Indonesian sign language (BISINDO) recognition is explained comprehensively in Section II. The results of BISINDO recognition using AF-DTW model and its comparison with previous studies are given in Section III. Finally, the conclusions and future works of this paper are given in Section IV.

II. Research Methodology

In this paper, the sign language gestures data consist of some variables obtained from various Kinect sensors input so that classified as multivariate time series data. Each variable represents the shoulder-center joint angle for each *X*-axis and *Z*-axis (see [6] for the detail). We used sign language gestures data from previous studies [1], [6] for comparative purposes. It is a recording of the gestures demonstrated by two deaf people (male and female) from the Indonesian Sign Language Interpreter Service Center, Jakarta.

Dynamic Time Warping (DTW) is one of popular models to measure similarity between two time series [15]. DTW model finds an optimal path which align two sets of time series data without requirement that they have the same length. DTW model measures the similarity or distance between these two sets of time series data by warping them non-linearly in time where the distance metric used does not necessarily hold triangle equality [16]. Fig. 2 illustrates alignment of two time series data.

DTW model allows to compute the distance and alignment between two time series data in the following procedure. Assuming there are two records of sign language gestures data:

$$x = \{x(i) \in \mathfrak{R}: i = 1, 2, \ldots, m\}$$

and

$$y = \{y(i) \in \mathfrak{R}: i = 1, 2, \ldots, n\}.$$

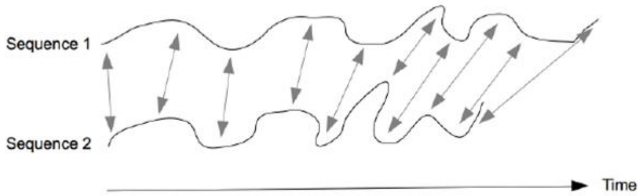


Fig. 2. Alignment of two time-dependent sequences [16].

To align these two time series data using DTW model, we need to define the distance matrix D containing Euclidian distances between all pairs of points $(x(i), y(j))$, $i \in [1,m]$ and $j \in [1,n]$ as

$$D(i,j) = d(x(i), y(j)), \tag{1}$$

where

$$d(x(i), y(j)) = |x(i) - y(j)|.$$

The matrix element $D(i,j)$ corresponds to the alignment between values $x(i)$ and $y(j)$ of the sign language gestures data *x* and *y*.

We then construct a warping path $W = \{w_1, w_2, \ldots, w_K\}$ of matrix elements $D(i,j)$ in equation (1). At the same time, the warping path *W* much satisfy at least three constraints such as boundary conditions, continuity, and monotonicity:

1. Boundary conditions: $w_1 = D(1,1), w_K = D(m,n)$;
2. Continuity: for $w_K = D(i_K, j_K)$ and $w_{K+1} = D(i_{K+1}, j_{K+1})$, $i_{K+1} - i_K \le 1$ and $j_{K+1} - j_K \le 1$; and
3. Monotonicity: $i_{K+1} - i_K \ge 0$ and $j_{K+1} - j_K \ge 0$.

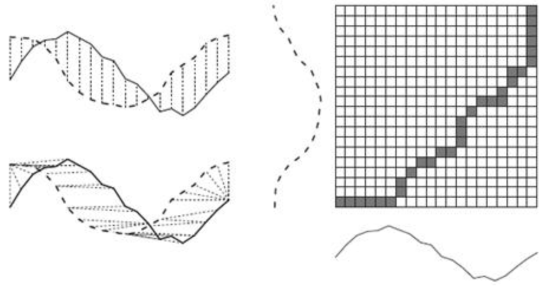


Fig. 3. Top left: two similar time series data but out of phase produce a large Euclidean distance; Bottom left: Non-linear alignment of DTW model; Right: warping matrix [17].

To form a warping path, we start at element $D(1, 1)$ and then move at most one index right or up until ending at $D(m,n)$ (See Fig. 3). DTW creates a matrix D depicted as a black and white grid in the Fig. 3 on the right side. The DTW model then runs through D from the first index (bottom left) up to the last index (top right). Enumerate all paths w and find an optimal warping path which is shown as the darkened line. The alignment between these two time series data is shown in the Fig. 3 on the left side [13]. The path which minimizes the warping cost gives us the distance value of DTW model [17]:

$$DTW\left(x, y\right) = \min_{W} \left\{ \sum_{k=1}^{K} w_k \right\}. \tag{2}$$

In general, dynamic programming is used to get the optimal warping path of DTW model by adding the distance $d(i, j)$ with the minimum of the cumulative distance of the three adjacent elements which defines as a cost matrix [15], i.e.

$$R(i,j) = d(i,j) + min \begin{cases} R(i,j-1) \\ R(i-1,j-1) \\ R(i-1,j) \end{cases}, \tag{3}$$

where $R(0,0) = 0$, $R(i,0) = R(0,j) = \infty$.

In this paper, we used the improved version of DTW model called Accurate and Fast Dynamic Time Warping (AF-DTW) model. It consists of three parts: (i) constructing the main idea of AF-DTW model by using a backward strategy; (ii) reducing the scope of cumulative distance in the cost matrix; and (iii) determining a threshold value used to reduce the scope in part (ii) [15].

The algorithm of AF-DTW model is constructed by using a backward strategy. This strategy starts from (m,n) to $(1, 1)$ where each of elements in the cost matrix is less than the three top right adjacent ones. There is a significant difference between DTW and AF-DTW models in the cost matrix construction. In equation (3), DTW model used the minimum of the cumulative distance of three bottom left adjacent elements for calculating a current element $R(i, j)$ while AF-DTW model used the maximum of the cumulative distance of three top right adjacent elements ones. This component is then subtracted by the distance $d(i, j)$ to obtain the current element $R'(i, j)$ of AF-DTW model, i.e.,

$$R'(i,j) = max \begin{cases} R'(i,j+1) \\ R'(i+1,j+1) - d(i,j), \\ R'(i+1,j) \end{cases} \tag{4}$$

where $i,j = n,n-1,\ldots,1; R'(n+1,n+1) = 0;$ and $R'(i,n+1) = R'(i,n+1) = -\infty$.

Similar with DTW model, a new warping path of AF-DTW model, $W' = \{w'_1, w'_2, \cdots, w'_K\}$, contains some of distance matrix elements with subject to the boundary conditions, continuity, and monotonicity. The optimal warping path of AF-DTW model with maximum warping cost can be obtained by also using the dynamic programming in this backward strategy as follow,

$$DTW_{BS}(x,y) = \max_{W'} \{\textstyle\sum_{k=1}^{K} w'_k\}. \tag{5}$$

However, $DTW_{BS}(x,y)$ formula in equation (5) has the same time and space complexity with $DTW(x,y)$ formula in equation (2). This complexity can be decreased by reducing the scope of cumulative distance R' in the cost matrix. We do not need to work with all cells in the original cost matrix because the optimal warping path always exist in the reduced scope (e.g. only the positive cells) with regards to a large enough threshold value θ. $DTW_{AF}(x,y;\theta)$ in equation (6) reduce the time and space complexity of $DTW_{BS}(x,y)$ for improving the performance of AF-DTW model [15],

$$DTW_{AF}(x,y;\theta) = \theta - DTW_{BS}(x,y). \tag{6}$$

Note that the reduced scope does not necessarily reduce the accuracy of the result.

First, choose $\theta \geq DTW(x,y)$ such that the optimal warping path is surrounded by the positive cells in the original cost matrix. There are some rules to construct the reduced scope by using backward strategy such as the current cell is set to be zero when it has negative value or the three top right adjacent cells are zero so that the current iteration is broken and AF-DTW model continued by calculating the positive cells until it gets the reduced scope. The minimum distance between time series x and y is retrieved from that reduced scope.

The selection of θ affects the computational time of AF-DTW model implementation because the larger value of θ the increasing number of positive cells in the cost matrix. However, two different large values of θ does not have a significant difference in terms of number of positive cells. This dilemma can be solved by determining an optimal threshold value such that the initial purpose of the reduced scope to reduce the time and space complexity of AF-DTW model can be achieved.

We know that the increasing number of positive cells in the cost matrix the more time and space are needed. In other words, $\theta = DTW(x.y)$ produces the smallest scope of cumulative distance in the cost matrix even though its minimum distance is unknown and needs to be calculated first. To find the θ which is closed to $DTW(x.y)$, we need to determine the relatively small initial value of θ as a sum of elements in the initial warping path (e.g. the opposite diagonal warping path), i.e. $\theta = \sum_{k=1}^{K} w_k$. Li and Yang showed that the reduced scope obtained by this technique is equal with the optimal reduced scope obtained by $\theta = DTW(x.y)$ [15]. There are many techniques that can be developed to obtain small θ which is larger than $DTW(x.y)$ but produces the same reduced scope to reduce the time and space complexity of AF-DTW model.

III. Results and Discussion

The Indonesian sign language (BISINDO) gestures which captured by Microsoft Kinect Xbox as an extracted skeleton data is divided into three groups based on their gender, i.e (i) male, (ii) female, and (iii) mixed. The extracted skeleton data consists of various features related to the gestures of shoulders, upper arms, forearms, hands, fingers, and face. However, in this study we only used some of them such as shoulders, upper arms, forearms, and hands by first transforming all gestures into angles between each of them and the shoulder-center. At the same time, we also label our recordings according to the BISINDO words that were exhibited. Furthermore, we used supervised learning approach to recognize BISINDO by using AF-DTW model.

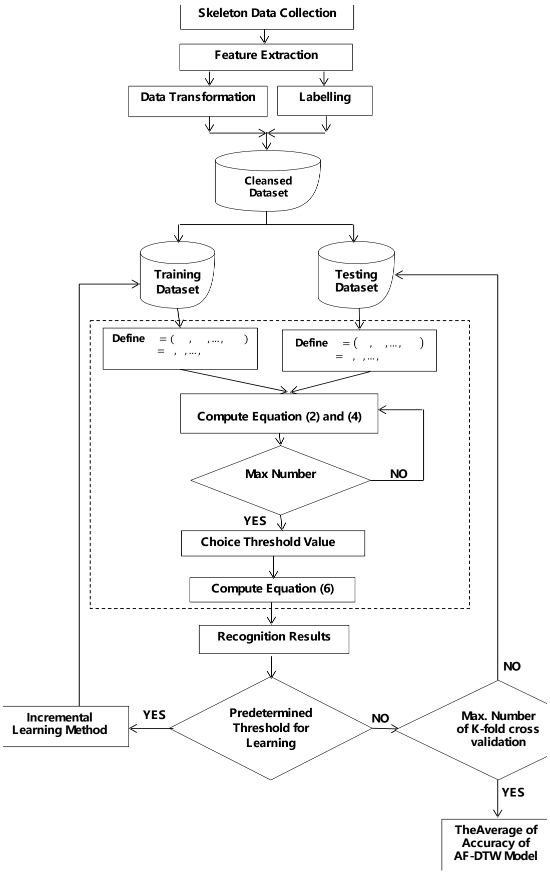


Fig. 4. Flowchart of BISINDO recognition using AF-DTW model.

Fig. 4 shows the flowchart the AF-DTW model to recognize BISINDO. We worked on the training and testing dataset which are obtained by dividing the cleansed dataset into two groups. The cross-validation technique is used to see the performance of Accurate and Fast Dynamic Time Warping (AF-DTW) model in BISINDO recognition. The cross-validation procedure is carried out by dividing the data into 5 groups, then 2 groups are used as training dataset while the rest as testing dataset. AF-DTW model as explained on the previous section provided with a dotted line boundary in Fig. 4 is worked on the testing dataset by using training dataset as referenced templates [18]. The recognition results are then evaluated by using a predetermined threshold to decide whether further learning is needed or not. If necessary, we use incremental learning method as recommended by Ding, *et al*. [18]. We added the worst series of frames in testing dataset recognized by AF-DTW into a new referenced template in the training dataset such that the accuracy of AF-DTW model can be improved. But if not, the cross-validation procedure can be continued until it is ready to calculate the average of accuracy of AF-DTW model to recognize BISINDO.

TABLE I: COMPARISON OF THE AVERAGE OF ACCURACY FOR BISINDO RECOGNITION

Sex	HMM [8]	mGLVQ [1]	AF-DTW
Male	63.94%	94.375%	90.95%
Female	61.41%	93.975%	92.875%
Mixed	70.31%	91.8125%	91.625%

In this paper, AF-DTW algorithm in Fig. 4 is coded by Python programming language that produces output as shown in Table I.

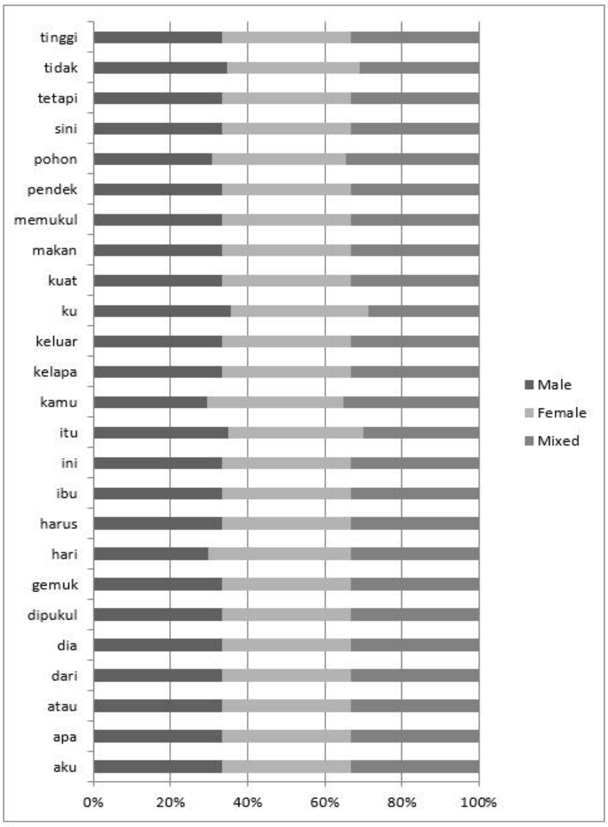


Fig. 5. The accuracy of AF-DTW model to recognize each word of BISINDO for each experiment.

The average of accuracy of AF-DTW model for each of experiments is then compared with the average of accuracy of the other models that have been applied in the previous studies, i.e. Hidden-Markov Model (HMM) [8] and modified Generalized Linear Vector Quantization (mGLVQ) model with mode function [1]. It shows that HMM has the lowest average of accuracy when compared to the other two models. Meanwhile, the average of accuracy of mGLVQ and AF-DTW models are not much different. Although the average of accuracy of AF-DTW model is less than mGLVQ model, it represents the sign language recognition problem which is time series classification. This is contrast to the Generalized Linear Vector Quantization (GLVQ) model which classifies sequence of frames in each frame. Therefore, we need to modify GLVQ into mGLVQ by involving some functions such as mode so that the real problem can be approached more objectively [1]. Objectivity is important to make our model becomes explainable.

The objectivity of AF-DTW model is also reflected in this result where the average of accuracy for mixed group is in between the other two groups, i.e. male and female groups. This is different with the results obtained in the previous studies where mixed group has the highest average of accuracy for HMM. In contrast, mGLVQ model produced the lowest average of accuracy in the mixed group. In fact, a model has its own specifications. Objectively, when the dataset is mixed, it will certainly reduce the average of accuracy of the best specification, and most likely be above the worst specification. AF-DTW model has better specifications in representing the female than the male groups so that the mixed group has the average of accuracy in between the other two groups. This characteristic did not occur in the other two models so that the average of accuracy of HMM and mGLVQ models cannot be explained further.

The accuracy of AF-DTW model to recognize each word of BISINDO for each experiment is shown in Fig. 5. Most of the words that are tested can be well recognized, only a little number of words which are slightly different performance for each experiment. However, the accuracy of AF-DTW model is relatively more balanced for each word when compared to mGLVQ (see [1] for the detail).

IV. Conclusion

Accurate and Fast Dynamic Time Warping (AF-DTW) model as one of deterministic time series classification model, is an appropriate model to recognize the Indonesian sign language (BISINDO). It represents the time series classification problems with high accuracy, i.e. around 90% – 92%. Based on the results obtained in this paper and previous studies related to BISINDO recognition, we decided to use AF-DTW model for our next research in recognizing sentence of BISINDO. However, some deep learning algorithms can be tried to improve the accuracy of BISINDO recognition, as in [19], [20]. In fact, the sign language is used by the deaf or hard-hearing people to communicate in the form of sentence, not a word. Some of previous studies forced to recognize a sentence of sign language with word-by-word techniques which is unnatural because we need to pause for a moment to show each word in the sentence. There is a problem for detecting the change point of multivariate time series in this new topic. We need to split the sentence of BISINDO into words and then applied AF-DTW to recognize these words. This, of course, has a direct impact on the significant increasing of computational time so that needs to be solved by parallel computing technique. Furthermore, we need to consider several things in linguistic perspectives such as morphology, syntax, or semantics.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Tri Handhika conducted the research design and analysis of this paper and wrote the manuscript by compiling all of the results and discussion during this research; Dewi Putrie Lestari organized the state-of-the-art of this paper; Ilmiyati Sari concerned with the methodology used in this research; Revaldo Ilfestra Metzi Zen was responsible for coding the program; and Murni collected the data; all authors had approved the final version.

Acknowledgment

We would like to thank Pusat Layanan Juru Bahasa Isyarat Indonesia (PLJ BISINDO) for their support and encouragement during this research. We sincerely appreciate the efforts of each member to introduce the Indonesian sign language (BISINDO), to provide the BISINDO dictionary, and to demonstrate the BISINDO for data collection requirements.

References

[1] T. Handhika, D. P. Lestari, I. Sari, R. I. M. Zen, and Murni, "The generalized learning vector quantization model to recognize Indonesian sign language (BISINDO)," in *Proc. the 3rd International Conference on Informatics and Computing*, 2018.

[2] H. Takimoto, J. Lee, and A. Kanagawa, "A robust gesture recognition using depth data," *International Journal of Machine Learning and Computing*, vol. 3, no. 2, pp. 245-249, Apr. 2013.

[3] M. Fraiwan, N. Khasawneh, H. Ershedat, I. A.-Alali, and H. A.-Kofahi, "A Kinect-based system for Arabic sign language to speech translation," *International Journal of Computer Applications in Technology*, vol. 52, no. 2/3, pp. 117-126, 2015.

[4] Z. Halim and G. Abbas, "A Kinect-based sign language hand gesture recognition system for hearing- and speech-impaired: A pilot study of Pakistani sign language," *Assistive Technology*, vol. 27, no. 1, pp. 34-43, Mar 2015.

[5] R. Lun and W. Zhao, "A survey of applications and human motion recognition with Microsoft Kinect," *International Journal of Pattern Recognition and Artificial Intelligence*, vol. 29, no. 5, pp. 1555008-1-1555008-48, Aug. 2015.

[6] T. Handhika, R. I. M. Zen, Murni, D. P. Lestari, and I. Sari, "Gesture recognition for Indonesian sign language (BISINDO)," *Journal of Physics: Conf. Series*, vol. 1028, pp. 1-8, June 2018.

[7] Q. Ma, L. Shen, W. Chen, J. Wang, J. Wei, and Z. Yu, "Functional echo state network for time series classification," *Information Sciences*, vol. 373, pp. 1-20, Dec. 2016.

[8] T. Handhika, Murni. D. P. Lestari, and I. Sari, "Multivariate time series classification analysis: State-of-the-art and future challenges," in *Proc. IOP Conference Series: Materials Science and Engineering*, vol. 536, no. 1, article number 012003, 2019.

[9] E. Rakun, M. Andriani, I. W. Wiprayoga, K. Danniswara, and A. Tjandra, "Combining depth image and skeleton data from Kinect for recognizing words in the sign system for Indonesian language (SIBI [Sistem Isyarat Baha Indonesia])," in *Proc. IEEE Advanced Computer Science and Information Systems*, pp. 387-392, 2013.

[10] W. N. Khotimah, Y. A. Susanto, and N. Suciati, "Combining decision tree and back propagation genetic algorithm neural network for recognizing word gestures in Indonesian sign language using Kinect," *Journal of Theoretical and Applied Information Technology*, vol. 95, no. 2, pp. 292-298, Jan. 2017.

[11] K. Barczewska and A. Drozd, "Comparison of methods for hand gesture recognition based on dynamic time warping algorithm," in *Proc. the 2013 Federated Conference on Computer Science and Information Systems*, pp. 207-210, 2013.

[12] A. H. -Vela, M. A. Bautista, X. P.-Sala, V. P.-Lopez, S. Escalera, X. Baro, O. Pujol, and C. Angulo, "Probability-based dynamic time warping and bag-of-visual-and-depth-words for human gesture recognition in RGB-D," *Pattern Recognition Letters*, vol. 50, pp. 112-121, Dec. 2014.

[13] S. Seto, W. Zhang, and Y. Zhou, "Multivariate time series classification using dynamic time warping template selection for human activity recognition," in *Proc. 2015 IEEE Symposium Series on Computational Intelligence*, pp. 1399-1406., 2015.

[14] A. Switonski, H. Josinski, and K. Wojciechowski, "Dynamic time warping in classification and selection of motion capture data," *Multidimensional Systems and Signal Processing*, vol. 30, no. 3, pp. 1437-1468, July 2019.

[15] H. Li and L. Yang, "Accurate and fast dynamic time warping," *Lecture Notes in Computer Science: Advanced Data Mining and Applications*, vol. 8346, pp. 133-144, 2013.

[16] E. Isikligil, "A method for isolated sign recognition with Kinect," M. S. Thesis, Science in Computer Engineering Department, Middle East Technical University, Turkey, 2014.

[17] T. Gorecki and M. Luczak, "Multivariate time series classification with parametric derivative dynamic time warping," *Expert Systems with Applications*, vol. 42, pp. 2305-2312, Apr. 2015.

[18] I.-Jr. Ding, C.-T. Yen, and Y.-M. Hsu, "Developments of machine learning schemes for dynamic time-wrapping-based speech recognition," *Mathematical Problems in Engineering*, vol. 2013, article ID 542680, pp. 1-10, 2013.

[19] H.-J. Kim, S.-J. Park, and S.-K. Lee, "Sign language recognition using motion history volume and hybrid neural networks," *International Journal of Machine Learning and Computing*, vol. 2, no. 6, pp. 750-753, Dec. 2012.

[20] A. H. Vo, V.-H. Pham, and B. T. Nguyen, "Deep learning for Vietnamese sign language recognition in video sequence," *International Journal of Machine Learning and Computing*, vol. 9, no. 4, pp. 440-445, Aug. 2019.

Tri Handhika was born in Sanga-Sanga, the East Borneo, Indonesia, on April 14, 1987. He received the B.S. degree in mathematics with concentration in actuarial statistics from University of Indonesia, Depok, Indonesia in 2008; the M.S. degree in mathematics with concentration in computational finance from University of Indonesia, Depok, Indonesia in 2011; and the Ph.D. degree in computer science with concentration in high-performance computing from Gunadarma University, Depok, Indonesia in 2015.

Currently, he is a lecturer at the Faculty of Industrial Technology as well as a researcher at the Center for Computational Mathematics Studies, Gunadarma University, Indonesia. He is also a senior data science consultant for several technology companies in Indonesia in building an automated recommendation system to support data-driven decision management. His research interests include stochastic modeling, actuarial mathematics, numerical simulation, artificial intelligence, and high-performance computing.

Dr. Handhika is a member of Indonesian Data Science Association, Indonesian Statistician Group and an editorial member of Applied and Computational Mathematics Journal. He has also served as an invited reviewer in some scientific publications. He received the Best Paper Award from the technical program committee of the 3rd IEEE International Conference on Informatics and Computing in 2018 and the 2nd International Conference of Computer and Informatics Engineering 2019, and the Best Presenter Award in the 4th International Conference on Computer Science and Computational Intelligence 2019.

Dewi Putrie Lestari was born in Serang, Indonesia, on December 26, 1985. She received the B.S. and M.S. degrees in mathematics from University of Indonesia, Depok, Indonesia in 2008 and 2012, respectively; and the Ph.D. degree in Information Technology from Gunadarma University, Depok, Indonesia in 2015.

Currently, she is a lecturer at the Faculty of Industrial Technology as well as a researcher at the Center for Computational Mathematics Studies, Gunadarma University. She has published extensively in the area of medical imaging. Her research interests include dynamic models, statistics, image processing, and medical imaging.

Ilmiyati Sari was born in Jakarta, Indonesia, on October 25, 1987. She received the B.S. and M.S. degrees in mathematics from University of Indonesia, Depok, Indonesia in 2009 and 2012, respectively; and the Ph.D. degree in Information Technology from Gunadarma University, Depok, Indonesia in 2018.

Currently, she is a lecturer at the Faculty of Industrial Technology as well as a researcher at the Center for Computational Mathematics Studies, Gunadarma University. She has published extensively in the area of gesture recognition. Her research interests include video processing, graph, and artificial intelligence.

Revaldo Ilfestra Metzi Zen was born in Jakarta, Indonesia, on October 5, 1986. He received the B.S. and M.S. degrees in mathematics and computer science from University of Indonesia, Depok, Indonesia in 2008 and 2012, respectively.

Currently, he is a data scientist at Big Data Division, PT. Metra Digital Media (Indonesia. His research interests include computer vision, big data, and machine learning.

Murni was born in Tangerang, Indonesia, on September 24, 1986. She received the B.S. and M.S. degrees in mathematics with concentration in actuarial statistics and computational finance from University of Indonesia, Depok, Indonesia in 2008 and 2011, respectively; and the Ph.D. degree in computer science with concentration in high-performance computing from Gunadarma University, Depok, Indonesia in 2015.

Currently, she is a lecturer at the Faculty of Industrial Technology as well as a researcher at the Center for Computational Mathematics Studies, Gunadarma University, Indonesia. Her research interests include stochastic analysis, graph theory, and high-performance computing.

Dr. Murni is a member of Indonesian Data Science Association.