Journal of Ubiquitous Systems & Pervasive Networks
Volume 12, No. 2 (2020) pp. 25-30

iKarate: Karate Kata Aiding System

Bassel Emad[a]*, Omar Atef[a], Yehya Shams[a], Ahmed El-Kerdany[a], Nada Shorim[a], Ayman Nabil[a], Ayman Atia[b,c]

[a]*Faculty Of Computer Science, Misr International University (MIU), Cairo, Egypt*
[b]*Faculty Of Computer Science and Artificial Intelligence, Helwan University (HU), Cairo, Egypt*
[c]*Faculty Of Computer Science, October University for Modern Sciences and Arts (MSA), Cairo, Egypt*

Abstract

Karate is a martial art that can be performed using hands and feet to deliver and block strikes. Karate Kata moves must be executed in a unique way, many moves are performed incorrectly during training. In this paper, we offer a system that Karate performers, coaches, judges and sporting clubs could use. The system aids the Karate performers by capturing their moves using Kinect v2 sensor, pre-processing these moves and then analyzing these moves using F-DTW. A report is displayed to the performers that is easily understandable, to learn how the mistakes were made, and how to fix it/learn from them. F-DTW was used for proving the concept, and an average accuracy of 93.65% was achieved. This Paper is mainly concerned about the first Karate Kata (Heian Shodan).

Keywords: *Classification, Kinect, Karate Coaching, Recognition, Analysis, Real-time Feedback, Human Movement Analysis.*

1. Introduction

Karate Kata moves are combination of physical moves. Its methods are qualitative not quantitative which makes it hard for students to perform the accurate version of the motion and harder to evaluate it. The correct order to perform Kata 1 (Heian Shodan) is shown in Fig: 2. Kids nowadays do not get the attention needed from the coach and may find difficulties learning these moves at a young age and since the training may consist of a large number of students so the trainer may be unable to focus on every detail of every student's move, which could result in taking more time to learn and master the move or it may lead to learning the move incorrectly from the beginning.

Despite the importance of the sports field and the fact that skill sports can be learned by video observation [1]. There are no suited viewpoint videos for learning Karate movements. In Addition, not many people were working in the Karate field. Therefore, it was a good opportunity to extend the researches in the Karate field, by developing a system that would make a difference in learning skill-based sports.

The main goal of this paper is to capture the moves of the performers in real-time, pre-processes these moves and then analyze these moves and give them a feedback report to enhance their technique or alert them if they are performing a move or a stance incorrectly. The report includes tips on how to execute the move correctly. As shown in Fig 1, the difference between the correct pose and wrong pose in the two moves is in the position and angle of the joints of the body. One of the challenges we faced while comparing and analyzing the captured motion is that we should take into consideration that the activities might be performed with different speed, body proportions and initial position [2]. Another challenge is that the stream of movements must be segmented to individual moves.

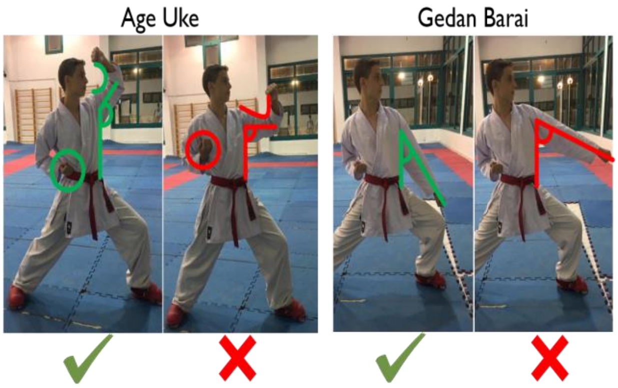


Fig. 1. Age-Uke and Gedan-Barai Correct and Incorrect

* Corresponding author. Tel.: +20-0100-5550-477.
E-mail: Bassel1600604@miuegypt.edu.eg
© 2020 International Association for Sharing Knowledge and Sustainability.
DOI: 10.5383/JUSPN.12.02.004

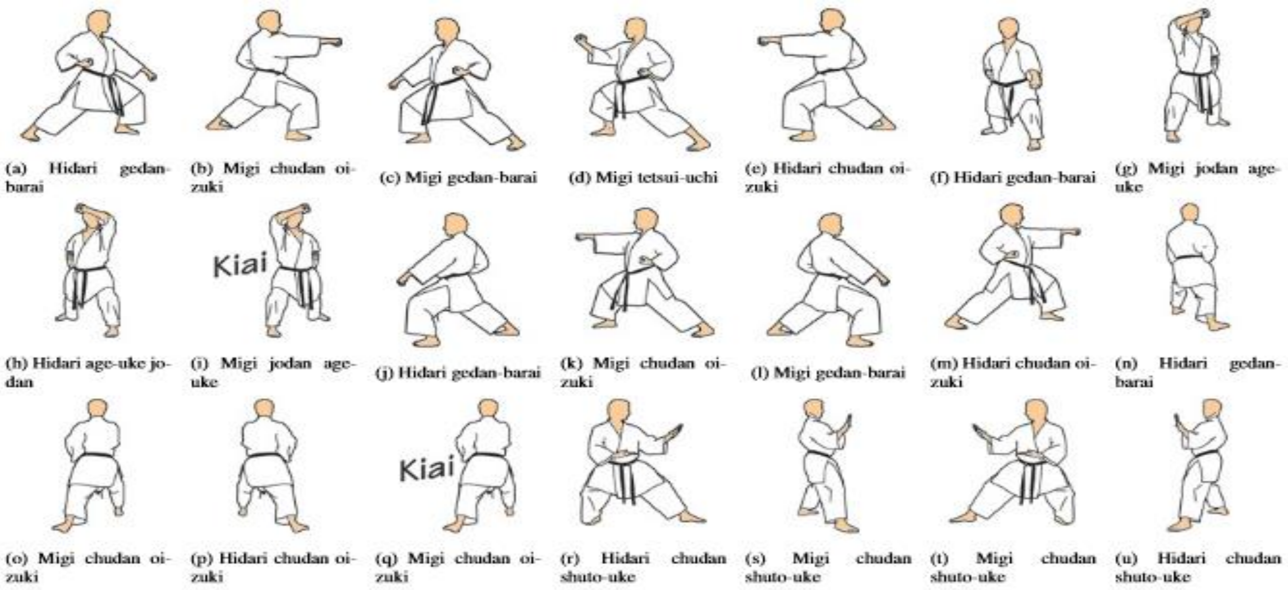


Fig. 2. Karate Kata 1 Moves (Hein Shodan)[2]

The rest of the paper is arranged as follows: Section 2 explains the related work and other approaches. Section 3 explains the proposed approach and the methods that were used. Section 4 shows the experiments and results reached in this paper. Finally, section 5 is the conclusion and the future work.

2. Related Work

T. Hachaj et al. [3] and [4] talked about using Kinect sensors to classify Karate moves and how to achieve the highest results. Initially, T. Hachaj et al.[3] proposed statistical differences between Kinect v1 and Kinect v2 recognition in some of the Karate moves. Their research [3] was done to evaluate the effectiveness of both Kinects in Karate motion recognition and to find the best-suited sensor. They made it clear that Kinect v2 would be more accurate and improve recognition rather than Kinect v1. Secondly, [4] proposed a calibration procedure of three Kinect sensors that integrates the data into one skeleton. The problem was to find a way to increase Karate motion recognition accuracy and effectiveness of fusion of body joints gathered from different sensors. They reached a positioning of the three sensors to improve the non-classified techniques by 48%. Also, A. D. Calin, [5] compared the efficiency of various classifiers tested on six data-sets obtained from both Kinect v1 and v2, to evaluate the accuracy of the two Kinect Sensors and analyze the gesture recognition accuracy of several classifiers. He reached high accuracy but with a time of 65.93 seconds using Multilayer Perceptron algorithm and Kinect v2.

P. Alborno et al. [6] and N. T. Thanh et al. [7] developed a system for evaluating and providing a score based on the quality and performance of each move. In [6] P. Alborno et al. proposed a method to measure the quality of Karate moves. The analysis and evaluation of full human body movements' qualities were a critical problem facing people creating sports systems lately. They discovered a solution by studying how much the limbs are synchronized during relevant motion phases. On the other hand, N. T. Thanh et al. [7] proposed a performance scoring system that utilizes the data of the images from the Kinect. Their aim was to generate a standard model to be used in any system around the world for Vietnamese Traditional martial arts.

T. Hachaj et al. in [2] proposed evaluation and visualization technique for advanced motion analysis. Their paper evaluated the comparison, analyses and visualization method of the similarities and differences between three-dimensional trajectories of human body joints in Karate movements. Their interest was in investigating the differences in human moves, because of the imperfect imitation problem. E. Escobedo-Cardenas and G. Camara-Chave in [8] developed a method for hand gesture recognition using the Kinect. They developed this approach to overcome the problems of hand gesture recognition using sensors and video-based tools. T. Hachaj et al. in [9] proposed actions descriptions with maximally three key-frames. Their aim was to make motion recognition in low-dimensional feature space and the selection of proper features for modelling multiple human actions. Their selection of proper features set to model human actions in low-dimensional space will benefit in prioritizing the feature selection for better recognition. Y. Choubik and A. Mahmoudi in [10] developed a real-time human body poses classification technique. Each Karate move is a sequence of poses, which make poses recognition important in this system. Their problem was applying machine learning algorithms to classify human body poses in real-time. T. Hachaj et al. in [11] developed a video annotation method that enables both numerical and categorical features calculation. Their aim was to design an efficient system for learning Karate.

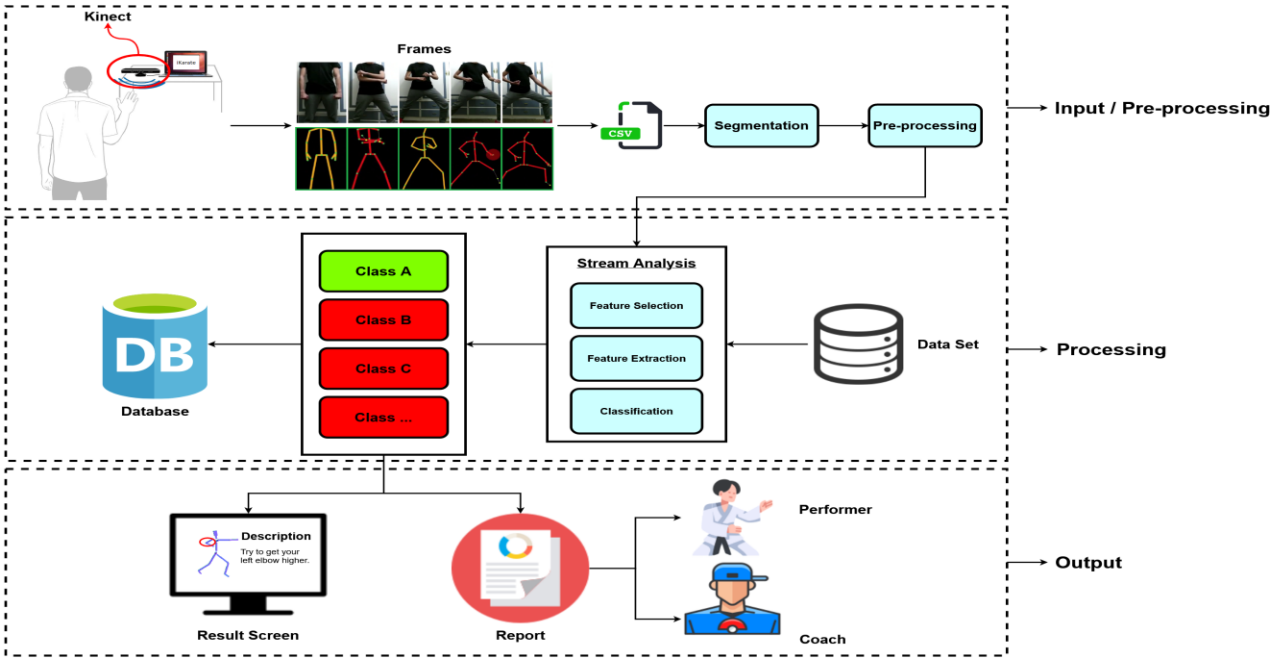


Fig. 3. System Overview

From [3] we may conclude that Kinect v2 is indeed more accurate in terms of precision than Kinect v1, due to its improved measurements of joint positions. In addition, [8] and [9] focused on selecting critical features to enhance the identification of hand and body movements, achieving an accuracy of 88.38% and 88% respectively. In [2] and [10] they concentrated on solving the problem of imitation and different positions that will cause errors during classifications using various algorithms and approaches. Whereas in [5] they made a very detailed comparison of six data-sets taken from Kinect v1 and v2 on various classifiers. Those comparisons will help in selecting the most appropriate hardware and algorithm to achieve the system's highest accuracy. In addition, the quality of karate movements was stated in [7] and [6], they worked on making a measuring technique for the quality of movements, one of which was the measurement of the positioning of the limbs.

3. Methodology

As shown in Fig. 3, the system is composed of a Kinect that would be facing the performer while performing a sequence of moves. Then the frames and the skeleton are extracted from the Kinect and saved in a CSV (Comma-separated values) file. After that, the segmentation and pre-processing will be made simultaneously. Afterwards, the analysis phase will occur, that contains feature selection, feature extraction and the classification. After performing the moves, the performer is presented with the moves' details and whether they were performed correctly or not. The performers are given a score regarding their performance of each move after it has been analyzed. The score evaluation is based on the performers' motion while executing the move and their speed. Dynamic analysis of the movement gives an accurate report to the performer or the coach, making the application more interactive. The report contains the performer's name, age, weight, height, belt color, move name and duration, how well the performer executed the move, how to improve the performer's performance and if any mistakes were made it will be shown in the report.

3.1. Input and Pre-processing

The input of the system is the movements' data. The data is consisted of the 25 joints of the skeleton captured by the Kinect in 3D space (X, Y and Z). Each dimension of each joint is stored in a column in the CSV file. After extracting the data, pre-processing phase takes place. The main stage in the pre-processing is the stream segmentation stage.

3.1.1. Kinect Sensor

The hardware used in this approach is the Kinect camera sensor, which could be utilized by the Microsoft SDK (Software development kit) for Kinect. As mentioned in [12], [13] the Kinect's hardware is composed of an Infrared Emitter to trace the body, presenting the skeleton and the body's joints.

In addition, as stated in Microsoft's book [13] the Kinect is able to provide 30FPS (Frames per second) with a 640 x 480-pixel resolution using its video and depth camera sensors. The Kinect operates by capturing the RGB (Red, green and blue) colors of the person to construct its image. Afterwards, the monochrome sensor and infrared projector begins receiving the rays that were emitted to obtain the third dimension and construct the 3D imagery of the skeleton of the person.

3.1.2. Data-set

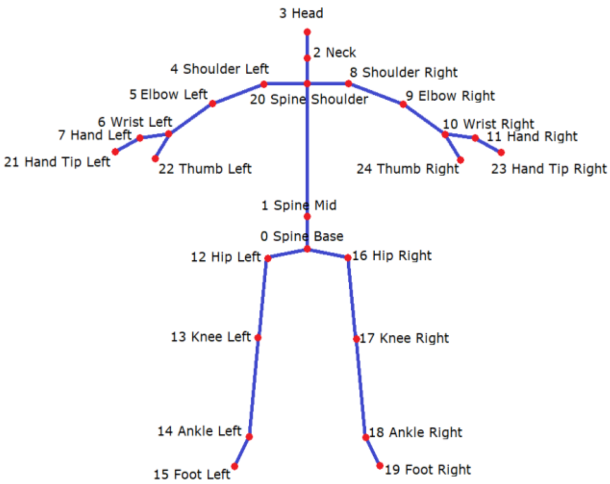


Fig. 4. Skeleton

The data-set used in this approach will be on Kata 1 (Heian Shodan), The data-set contains the 21 moves of Kata 1. As mentioned in section 3.1.1., the data-set has been gathered by the IR camera sensor. Each movement is represented in 25 joints, each joint contains three dimensions (X, Y, and Z). Fig. 4 shows the skeleton's joints that is acquired by the IR camera sensor.

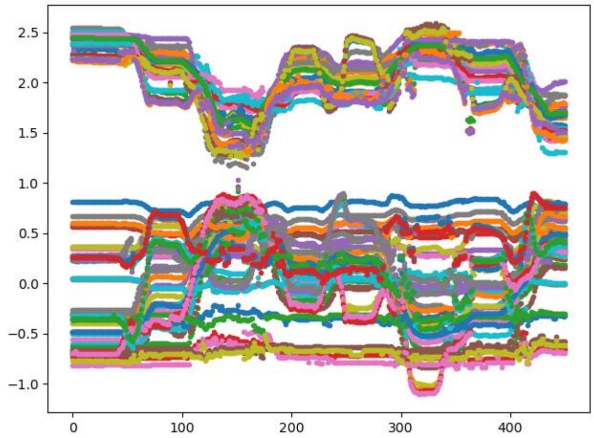
The collected data-set was 70 trials, collected from seven professional performers executing Kata 1. Half of the trials were correct moves and the other half represent an incorrect form of the move that includes a common mistake. These trials were used in training the models. The data were collected in multiple sessions. Before using the data-set, a cross-validation process was made to validate the data and remove any corruption.

3.1.3. Segmentation

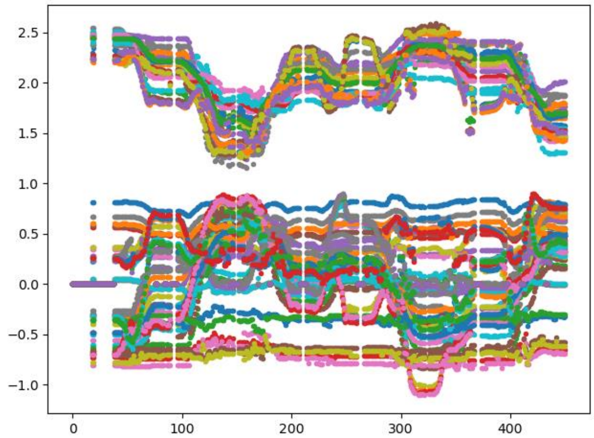
Each move should be analyzed separately from the other moves, so that it could be analyzed and classified without interfering with the other moves. Therefore, a segmentation technique was needed to split each move from the other moves and then send each move to the following stage.

$$Difference = \left| Mean(Window_1) - Mean(Window_2) \right| \quad (1)$$

There is a slight gap before each move in Karate Kata, which was determined after performing multiple tests. Therefore, a technique of segmentation was implemented according to the gathered information. The method used to segment the stream of movements was to build a queue that will serve as a window filled with frames, each frame contains the joints coordinates. A feature selection was used to determine the effective joints. After that, the average of each two consecutive windows is calculated and then subtracted to use their absolute value as shown in Eq: 1. A threshold is calculated based on the value of the difference using an unsupervised clustering technique (ISODATA algorithm). If the difference is below the threshold, then this indicates that there was no motion and the move will be segmented at this point. As shown in Fig: 5, Fig: (5a) represents the data before segmentation, while Fig: (5b) represents the data after segmentation and that the moves was separated.



(a)



(b)

Fig. 5. Before (a) and After (b) Pre-processing

3.2. Processing and Algorithms

3.2.1. F-DTW

After pre-processing the data, F-DTW (Fast Dynamic Time Warping) was used to analyze and handle the different speeds of the performers executing the moves. F-DTW is an algorithm for calculating the similarities between two signals, each signal may have a different speed from the other signals. F-DTW is an alignment algorithm, which is capable of classifying two different time signals. It is also used in [2], [8], [9], [14], [5], [15] and [16]. F-DTW could be applied to various types of data, like videos, audio and graphics data. Therefore, any data that can be converted to a linear sequence could be classified using it. F-DTW was able to recognize the classes of the correct and incorrect moves. F-DTW could be used with different distance equations but the "Euclidean Distance" is the one used in this system.

$$D(P) = \sum_{K=1}^{k=K} D(P_{ki} - P_{kj}) \quad (2)$$

F-DTW calculates the distance and path as shown in Eq.2. Where "D (P)" is the Euclidean Distance of the warp path "P".

Where the wrap path length is “K” and “D (pk_i, pk_j)” represents the distance of two data points (One from the training data and the other is from the testing data) in the “kth” element of the warp path as mentioned by Stan Salvador et al. in [17].

3.2.2. Support Vector Machine

SVM (Support Vector Machine) is a supervised learning classification algorithm. It is effective in recognizing patterns and can be applied to recognize these patterns in a various type of applications. It creates a hyperplane that represents the maximum margin that categorize the data into classes [18], [19].

SVM takes as an input the 3D coordinates of each joint represented in 75 columns, but pre-processing was needed before delivering the data to the SVM. The first pre-processing phase was the interpolation, which was used for adjusting the size of the data before building the SVM model, therefore linear interpolation was used to adjust all the movements to be equal in size without interfering in the signal shape. Then, the data will be flattened to 1D array, so that the SVM can accept it.

Fig. 6. Real-life Experiment

4. Experiments and Results

To prove that the system could run efficiently, 30 trials were collected from three Karate performers to test the system. A data-set has been used for training the model as mentioned in 3.1.2 which contains the first Kata moves to evaluate the system. As shown in Table 1, each trial has been tested with both classifiers after pre-processing. An average accuracy of each classifier was calculated. F-DTW has reached an accuracy of 93.65% while the SVM has reached an accuracy of 79.36%. It was observed that the F-DTW obtains higher accuracy due to its ability to handle signals with different sizes.

Table 1. Results

Algorithm	*Performer 1*	*Performer 2*	*Performer 3*	*AVG %*
F-DTW	95.23%	95.23%	90..47%	**93.65%**
SVM	80.95%	71.42%	85.71%	79.36%

As shown in Fig. 6, the performer is facing the Kinect and performing the moves. The system captures these moves, stores, pre-process them and then classifies the data and presents the results in a report. A Karate coach was present to tell exactly what mistakes were made while performing the moves and to discuss the result of the system. The accuracy has been discussed with the coach to determine whether it is good enough or not. The coach confirms that this accuracy is very good regarding the Karate Kata system.

Fig. 7. Setup

As shown in Fig. 7, the performers need to calibrate their bodies in front of the Kinect to ensure their entire skeleton is detected. Then, they execute the moves and the program will show the outcomes of the entire session when they end.

5. Conclusion and Future Work

In this paper, we introduce iKarate, a Karate Kata classification system. With iKarate, our goal is to enhance the field of Karate Kata smart coaching. Two main challenges were addressed in this approach, which was that (1) we must take into thought that the moves could be executed in different speed and body propositions and that (2) the stream of movements requires segmentation. The proposed approach uses the Kinect to capture the performer's motion and uses F-DTW as a classification method. F-DTW was ideal for this approach, because of its ability to manage different size data. The movements that were used in this experiment as mentioned in 3.1.2 were the Kata 1 (Heian Shodan) moves. F-DTW was able to recognize the movements and classify them correctly with an average accuracy of 93.65%. The system's output is a report containing the information of the moves that the performer had executed. F-DTW accuracy has been discussed with a professional Karate coach and the coach approved of the accuracy.

For future work, we intend to gather bigger data-set, use better algorithms for classification to enhance the accuracy and the response time and remove any noise from the data. Moreover, Karate actions sometimes is invisible to a stationary camera so we might use multiple Kinects.

Acknowledgments

The coaches and performers of Al-Ahly sporting club sponsored this work, we would like to thank them for helping in developing this approach.

References

[1] Tanaka, K. 3d action reconstruction using virtual player to assist karate training. 2017 IEEE Virtual Reality (VR). 2017. https://doi.org/10.1109/VR.2017.7892343

[2] Hachaj, T., Ogiela, M.R., Piekarczyk, M., Koptyra, K. Advanced human motion analysis and visualization: Comparison of mawashi-geri kick of two elite karate athletes. IEEE Symposium Series on Computational Intelligence(SSCI).2017. https://doi.org/10.1109/SSCI.2017.8285269

[3] Hachaj, T., Ogiela, M.R., Koptyra, K. E ectiveness comparison of kinect and kinect 2 for recognition of oyama karate techniques. 2015 18th International Conference on Network-Based Information Systems. 2015.

[4] Hachaj, T., Ogiela, M.R., Piekarczyk, M. Dependence of kinect sensors number and position on gestures recognition with gesture description language semantic classifier. Proceedings of the 2013 Federated Conference on Computer Science and Information Systems. 2013.

[5] Calin, A.D. Variation of pose and gesture recognition accuracy using two kinect versions. 2016 International Symposium on Innovations in Intelligent SysTems and Applications (INISTA). 2016.

[6] Alborno, P., Giorgis, N.D., Camurri, A., Puppo, E. Limbs synchronisation as a measure of movement quality in karate. Proceedings of the 4th International Conference on Movement Computing-MOCO17.2017. https://doi.org/10.1145/3077981.3078027

[7] Thanh, N.T., Tuyen, N.D., Dung, L., Cong, P.T. Implementation of technical data analysis of skeleton extracted from camera kinect in grading movements of Vietnamese martial arts. 2017 International Conference on Advanced Technologies for Communications (ATC). 2017.

[8] Escobedo-Cardenas, E., Camara-Chavez, G. A robust gesture recognition using hand local data and skeleton trajectory. 2015 IEEE International Conference on Image Processing(ICIP).2015. https://doi.org/10.1109/ICIP.2015.7350998

[9] Hachaj, T., Ogiela, M.R., Koptyra, K. Human actions modelling and recognition in low-dimensional feature space. 10th International Conference on Broadband and Wireless Computing, Communication and Applications (BWCCA).2015. https://doi.org/10.1109/BWCCA.2015.15

[10] Choubik, Y., Mahmoudi, A. Machine learning for real time poses classification using kinect skeleton data. 13th International Conference on Computer Graphics, Imaging and Visualization(CGiV).2016. https://doi.org/10.1109/CGiV.2016.66

[11] Hachaj, T., Ogiela, M.R., Koptyra, K. Learning from annotated video: An initial study based on oyama karate tournament recordings. 10th International Conference on P2P, Parallel, Grid, Cloud and Internet Computing (3PGCIC). 2015. https://doi.org/10.1109/3PGCIC.2015.16

[12] Zhang, Z. Microsoft kinect sensor and its effect. IEEE Multimedia.2012. https://doi.org/10.1109/MMUL.2012.24

[13] Miles, R.S. Start here!: learn the Kinect API. Microsoft. 2012.

[14] Hachaj, T., Piekarczyk, M., Ogiela, M.R. How repetitive are karate kicks performed by skilled practitioners? Proceedings of the 2018 10th International Conference on Computer and Automation Engineering – ICCAE. 2018. https://doi.org/10.1145/3192975.3192989

[15] Hachaj, T., Ogiela, M.R., Piekarczyk, M. Real-time recognition of selected karate techniques using gdl approach. Advances in Intelligent Systems and Computing Image Processing and Communications Challenges. 2014.. https://doi.org/10.1007/978-3-319-01622-1_12

[16] Janaki, V.M.S., Babu, S., Sreekanth, S.S. Real time recognition of 3d gestures in mobile devices. IEEE Recent Advances in Intelligent Computational Systems (RAICS). 2013.

[17] Salvador, S., Chan, P. Fastdtw: Toward accurate dynamic time warping in linear time and space. 2004.

[18] Yue Shihong, Li Ping, H.P. Svm classification its contents and challenges. 2001.

[19] SHUJUN HUANG, NIANGUANG CAI, P.P.P.S.N.Y.W.W.X. Applications of support vector machine (svm) learning in cancer genomics. Cancer Genomics Proteomics. 2017.

[20] Shotokan karate do. https://shotokancrsa.com/heianshodan-eng.htm, last accessed on 01/07/20.