اونيۏرسيتي مليسيا ڤهڠ
UNIVERSITI MALAYSIA PAHANG

ORIGINAL ARTICLE

The Classification of Taekwondo Kicks Via Machine Learning: A feature selection investigation

Muhammad Syafi'i Mass Duki[1], Muhammad Nur Aiman Shapiee[1], Muhammad Amirul Abdullah[1], Ismail Mohd Khairuddin[1] Mohd Azraai Mohd Razman[1] and Anwar P. P. Abdul Majeed[1,2]*

[1]Faculty of Manufacturing and Mechatronics Engineering Technology, Universiti Malaysia Pahang, 26600 Pahang, Malaysia.
[2]IBM Centre of Excellence, Universiti Malaysia Pahang, 26600, Malaysia

ABSTRACT – Martial art strike classification by machine learning has drawn more attention over the past decade. The unique signal from each technique makes it harder to be recognized. Thus, this paper proposed an SVM, Random Forest, k-NN, and Naïve Bayes classification method applied to the time-domain signal to classify the three type of taekwondo technique. Data collected from the total of five participant and statistical features such as mean, median, minimum, maximum, standard deviation, variance, skewness, kurtosis, and standard error mean were extracted from the signal. After that, the data will be trained using selected rank features and hold-out method with k-fold cross-validation applied to the training and testing data. Therefore, with ANOVA test as features selection and 60:40 ratio of a hold-out method, Random Forest classifier score the highest accuracy of 86.7%..

INTRODUCTION

Taekwondo is one of the numerous ancient martial arts in the world [1] that has embraced the future of Olympic games—first introduced in Korea around 1940s and 1950s by Korean martial artists with karate experience, Chinese martial arts, and others traditional practiser. As combat sports between two fighters, the participant's weight category needs to be same [2]. Equipped with the protection is a must before the game starts because of taekwondo popular with its unique kicking leg as a primary attack weapon [3]. In the Olympic Games, scoring was determined using electronic scoring systems embedded in the head and trunk protector [4]. The judges use a manual scoring device to give reward point based on several techniques that hit the opponent. The decision may affect the accuracy in the taekwondo tournament's scoring system. Therefore, several approaches were done to build machine learning classification for classified various type of taekwondo technique.

For example, Samiullah [5] proposed to classified martial art motion from a single wearable sensor. The Metawear C sensor was used as a device to collect dataset consisting of eight activities in Brazilian Jiu-Jitsu martial art. Classifying the activities is more effective when the sensor located around the torso of the practitioner. Equipped with 3-axis accelerometer and a 3-axis gyroscope, approximately 54 datasets of time-series signal were collected. Moreover, standard statistical data features were applied to all 3-axis raw data to extract the features. The author suggests using all total of eight different classifier and 10-fold cross-validation scoring method to balance the accuracy of the calculation. Among all the seven classifiers, Random Forest obtains the highest accuracy of 72% while 31% classification accuracy for Naïve Bayes is the lowest.

In another major study, Zhong [6] introduced the GA-SVM classification method applied to the dynamic evaluation of taekwondo. Genetic algorithm (GA) was used to determine the useful and useless features to reduce the features and improve the accuracy of classification. The method suggests achieved 100% accuracy compared to the C4.5 method only obtain 96.87 classification accuracy. Next, Badawi [7] evaluated the daily activity recognition using wearable sensor via machine learning. The best three statistical features used are standard deviation auto-correlation, mean auto-covariance, skewness, and mean crossing rate were feed to train into four main classifiers. The author also uses k-fold cross-validation method to determine the optimal number of features and select the highest scoring group of features. However, Random Forest attains the highest score of classification accuracy with 96.8% compared to the k-NN obtain only 82.0% accuracy.

In this paper, the study's purpose is to identify the significant time-domain base features by using ANOVA and chi-square method to evaluate the efficiency of different type of machine learning model used as a classifier for a classified distinct type of taekwondo kick. The classification used is Support Vector Machine (SVM), k-Neighbors Nearest (k-NN), Random Forest (RF), and Naïve Bayes (NB). To evaluate the classification performance, confusion matrix will be used to prove the performance of the model.
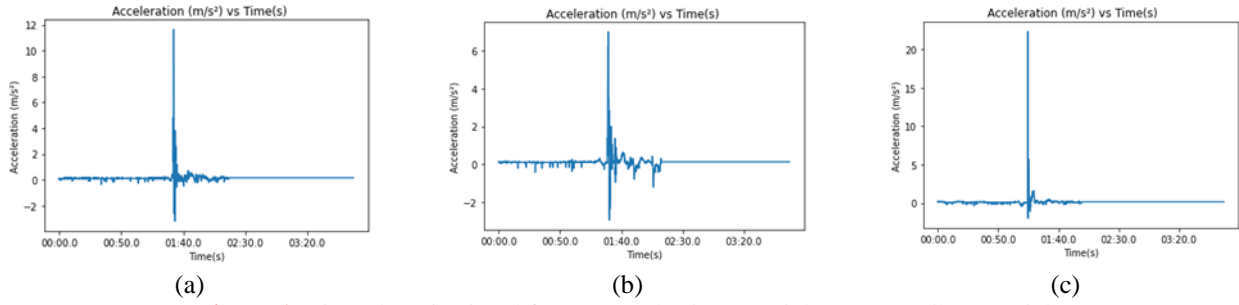
Figure 1. Time-domain signal for (a) Punch, (b) Cut Kick, (c) Roundhouse Kick

METHODOLOGY

Data Collection

In this study, the dataset used from the online Kaggle database was used to classify the taekwondo technique. Five participants were performing three distinct types of taekwondo technique, the roundhouse kick, cut kick, and punch. Time-domain signal obtains from each technique are in the raw format then converted into the acceleration (m/s²). To validate the reliability of the dataset, the assumption of the ADXL335 sensor was used to compare the signal's pattern.

$$Acceleration = \frac{\frac{ADC\ value * Vref}{2048} - Voltage\ Level\ at\ 0g}{Sensitivity\ Scale\ Factor} \tag{1}$$

All the parameter Vref, voltage level at 0g, and sensitivity scale factor will be referred to the datasheet of the ADXL335 sensor. Afterwards converted dataset possibly plotted by using matplotlib library using python programming language at a sampling rate of 20Hz. The example of the converted acceleration signal can be observed in Figure 1 (a), (b), and (c).

Features Extraction

Several statistical time-domain features were extracted in this study by using a python programming language. The extracted features are minimum, maximum, mean, median, standard deviation, variance, skewness, kurtosis, and standard error mean [8].

(1) Mean: Mean represents the segment's amplitude over the sample length of the time-domain signal.

$$mean(\bar{x}) = \frac{1}{N}\sum_{n=1}^{N} x_n \tag{2}$$

(2) Minimum: Minimum is the smallest value that represents of the time domain signal. It would be the first value that indicates of the X_1 for a sample size of n.

(3) Maximum: Maximum is the highest value that represents of the time domain signal. It would be the last value that indicates of the Xn for a sample size of n.

(4) Median: Median is the middle value of the time-domain signal. It commonly used as a position metric where the severe value in dataset presented with less relevant because of the biased distribution or the outlier is inaccurate.

$$Median = \frac{n_{th} + (n+1)_{th}}{2} \tag{3}$$

(5) Standard Deviation: Standard Deviation is a measurement of the distribution of observation within a dataset relative to its mean and square root of variance. It can be denoted as sigma.

$$Standard\ deviation(\sigma) = \sqrt{\frac{1}{N-1}\sum_{n=1}^{N}(x_n - \bar{x})^2} \tag{4}$$

(6) Variance: Variance is a numerical value representing the mean amplitude of the time domain signal and is denoted by sigma-square.

$$Variance\ (\sigma^2) = \frac{1}{N-1}\sum_{n=1}^{N}(x_n - \bar{x})^2 \qquad (5)$$

(7) Skewness: Skewness refers to the mathematical metric used to determine the asymmetry of the probability distribution of random variables by its own mean, and its result can be positive, negative or undetermined.

$$skewness(SKW) = \frac{\sigma^3}{N}\sum_{n=1}^{N}(x_n - \bar{x})^3 \qquad (6)$$

(8) Kurtosis: Kurtosis is the central peak of the time domain signal. The higher value of kurtosis indicates a higher or sharper peak, while the lower value of kurtosis indicates a less distinct peak.

$$Kurtosis(KURT) = \frac{\sigma^4}{N}\sum_{n=1}^{N}(x_n - \bar{x})^4 \qquad (7)$$

(9) Standard Error Mean: Standard error mean also known as the standard deviation of the mean used to approximate the standard deviation of a sampling. If the effect random changes, the higher will be the standard error mean, while if there are no changes in data, the standard error mean is equal to zero.

$$Standard\ Error\ Mean(SEM) = \frac{\sigma}{\sqrt{N}} \qquad (8)$$

Features Selection

In order to decrease the amount of input variable while developing a predictive model, features selection will be used, such as ANOVA and Chi-square. Both methods will rank the features and select the best three based on important features. The result of ANOVA and Chi-square will be compared to identify the features important that give the highest classification accuracy.

(1) ANOVA: ANOVA analyses variance that used variance to determine the statistical differences between the mean of three or more hypotheses [9]. The F-test value identifies the significant variance different between the group and within a group. The sum-of-squares for between the group (SSB) and the sum-of-squares for within the group (SSW) need to be calculated as stated in the formula. The F-test value indicates that the features enough to reject the hypothesis if the value is beyond the confidence level.

$$SSB = \sum (g_i - \bar{x})^2 \qquad (9)$$

$$SSW = \sum (x_i - g_i)^2 \qquad (10)$$

$$F = \frac{SSB/df_B}{SSW/df_w} \qquad (11)$$

For g_i is the group mean, x_i and $\bar{x}$ are i[th] value in the set and mean of all the values, df_B and df_w are degree of freedom for SSB and SSW, respectively.

(2) Chi-Squares: Chi-squares test (χ^2) used to test the hypothesis on the distribution of a different set of features categories [10]. In this method, a dependency between both features and target variable can be determined by observing the count O and expected count E. The value of χ^2 test would be smaller if the observer count closed to the expected count. The higher value of χ^2 indicates that the hypothesis of independent is incorrect.

$$Chi-square(\chi^2) = \sum_{i=1}^{m}\frac{(O_i - E_i)^2}{E_i} \qquad (12)$$

For, O_i count is the observed frequency, E_i count is the estimated frequency.

Classification

Classification is a process to validate selection features and predict the target of input data in supervised learning. In this study, several classifiers were used such as Support Vector Machine (SVM), Random Forest (RF), k-Nearest Neighbors (k-NN) and Naïve Bayes (NB) to compare the classification result.

(1) Support Vector Machine: Support Vector Machine (SVM) is popular among the supervised learning that used method of generating hyperplanes in the number of input features dimensional space which is separate features vectors of different classes. SVM create the margin between the hyperplane to maximize the closest feature vectors on each side.

$$k(x_i \cdot x_j) = (x_i \cdot x_j)^d \tag{13}$$

(2) Random Forest: Random Forest (RF) is one of the most accurate algorithms in the machine learning model. It used a tree as voting in predicting the classes. By combining all decision tree outputs as classifier fusion method, the machine learning model's classification performance will be higher.

(3) k-Nearest Neighbors: k-Nearest Neighbors (k-NN) is one of the most widely used algorithms in classification. In a set of k objects in the training class, the distance between one of the test class samples was calculated. Training sample with the shortest distance represents the test sample. The value of k needs to be minimum to ensure the classification performance better.

$$Euclidean = \sqrt{\sum_{i=1}^{k}(x_i - y_i)^2} \tag{14}$$

(4) Naïve Bayes: Naïve Bayes (NB) work based on Bayes theorem known as a probabilistic classifier used to classify a large amount of data. With the given probability of predator as evidence, the probability of class can be determined. The assumption of predictor is independent and the present of features in a class is not affect the other.

$$P(A|B) = \frac{P(B|A)P(A)}{P(B)} \tag{15}$$

Where P (A | B) is the posterior probability of class or target given predictor or attributes, P(A) is the prior probability of class, P (B | A) is the probability of predictor given class, and P(B) is the prior probability of predictor.

Performance Metrics

Classification accuracy (CA) and confusion matrix were used as a performance matrix to evaluate the different classifier's performance.

(1) Accuracy: An accuracy is a ratio number of correct predictions in the machine learning model to the total number input of samples.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{16}$$

(2) Confusion Matrix: A confusion matrix is a table often used to define a classification model output on a collection of test data for which the true values are known as in Table 1. It is relatively easy to comprehend the confusion matrix itself.

Table 1. Confusion Matrix

		Predicted	
		Positive	Negative
Actual	Positive	TP	FP
	Negative	FN	TN

EXPERIMENTAL RESULTS

Several statistical features were extracted and selected using feature selection method where all feature will be ranked based on the score of features important. Three significant features are selected: skewness, kurtosis, and maximum after that, it will be fed into a classifier to classified taekwondo technique. Figure 2 shows the classification accuracy of different type of feature selection with hold-out ratio 60:40 scoring method.
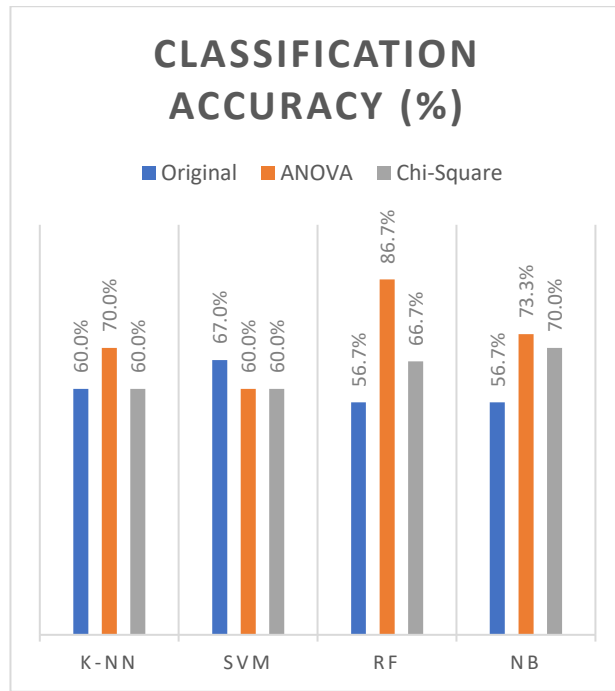
Figure 2. Classification accuracy for different type of feature selection method

From the result shown in Figure 2, Random Forest achieved the highest classification accuracy of 86.7% by using ANOVA feature selection method. However, SVM hit the highest classification accuracy of 67.0% in the original feature compared to the other feature selection. An improvement for Chi-Squares in Random Forest and Naïve Bayes where classification accuracy of 66.7% and 70.0% respectively compared to the original feature. In k-NN classifier, both the original and Chi-Squares method scores have the same accuracy of 60.0% and a rise of 10% for ANOVA feature selection. To conclude, Random Forest classifier is the best model used with ANOVA based on classification performance, and for each feature selection technique, the confusion matrix for the best classifier is shown in Figure 3 to 6.

Figure 3. Confusion matrix for Naïve Bayes on original feature

Figure 4. Confusion matrix for Random Forest on original feature

Figure 5. Confusion matrix for Random Forest on ANOVA feature selection method

Figure 6. Confusion matrix for SVM on Chi-Squares feature selection method

CONCLUSION

In this study, the classification of taekwondo by mean of machine learning was proposed. Time-domain signal obtain from each technique reached the highest classification accuracy of 86.7% used with the ANOVA feature selection method. The evaluation of the efficiency of the machine learning model was achieved. However, to enhance the efficiency of the machine learning model, further suggestion will be stated to increase the number of trials taken for each technique for different practitioners. Moreover, improving the way of data collection to avoid any error and missing data. Hence, hyperparameter tuning to consider the right parameter for classification.

ACKNOWLEDGEMENT

REFERENCES

[1]   M. A. Brudnak, D. Dundero, and F. M. Van Hecke, "Are the 'hard' martial arts, such as the Korean martial art, TaeKwon-Do, of benefit to senior citizens?," *Med. Hypotheses*, vol. 59, no. 4, pp. 485–491, 2002, doi: 10.1016/S0306-9877(02)00203-7.

[2]   A. Gutiérrez-Santiago, R. Pereira-Rodríguez, and I. Prieto-Lage, "Detection of the technical and tactical motion of the scorable movements in taekwondo," *Physiol. Behav.*, vol. 217, no. November 2019, p. 112813, 2020, doi: 10.1016/j.physbeh.2020.112813.

[3]   Y. J. Tsai, C. F. Huang, and G. H. Gu, "The kinematic Analysis of Spin-Whip Kick of Taekwondo in Elite Athletes," *J. Biomech.*, vol. 40, no. July, p. S615, 2007, doi: 10.1016/s0021-9290(07)70603-6.

[4]   C. Menescardi, C. Falco, A. Hernández-Mendo, and V. Morales-Sánchez, "Design, validation, and testing of an observational tool for technical and tactical analysis in the taekwondo competition at the 2016 Olympic games," *Physiol. Behav.*, vol. 224, no. May, p. 112980, 2020, doi: 10.1016/j.physbeh.2020.112980.

[5]   C. Samiullah, S. Devlin, and U. Suess, "Classifying Martial Arts Motion from a Single Wearable Sensor," 2020, [Online]. Available: http://christophergs.github.io/samiullah_devlin_suess_iaai_submission.pdf.

[6]   M. H. Zhong, J. C. Hung, Y. C. Yang, and C. P. Huang, "GA-SVM classifying method applied to dynamic evaluation of taekwondo," *Proc. IEEE Int. Conf. Adv. Mater. Sci. Eng. Innov. Sci. Eng. IEEE-ICAMSE 2016*, pp. 534–537, 2017, doi: 10.1109/ICAMSE.2016.7840191.

[7]   A. A. Badawi, A. Al-Kabbany, and H. Shaban, "Daily Activity Recognition using Wearable Sensors via Machine Learning and Feature Selection," *Proc. - 2018 13th Int. Conf. Comput. Eng. Syst. ICCES 2018*, pp. 75–79, 2019, doi: 10.1109/ICCES.2018.8639309.

[8]   M. Barandas *et al.*, "TSFEL: Time Series Feature Extraction Library," *SoftwareX*, vol. 11, p. 100456, 2020, doi: 10.1016/j.softx.2020.100456.

[9]   B. Sheng, O. M. Moosman, B. Del Pozo-Cruz, J. Del Pozo-Cruz, R. M. Alfonso-Rosa, and Y. Zhang, "A comparison of

different machine learning algorithms, types and placements of activity monitors for physical activity classification," *Meas. J. Int. Meas. Confed.*, vol. 154, p. 107480, 2020, doi: 10.1016/j.measurement.2020.107480.

[10]   S. K. Trivedi, "A study on credit scoring modeling with different feature selection and machine learning approaches," *Technol. Soc.*, vol. 63, no. October 2017, p. 101413, 2020, doi: 10.1016/j.techsoc.2020.101413.