Classification of Walking Gait Features using Markerless-based Approach in ASD Children

Nur Khalidah Zakaria, Nooritawati Md Tahir and R. Jailani

Abstract—This paper discussed the gait classification of autism children versus normal group. The study involved children from 30 typically development and 21 autism children with aged range between 6 to 13 years old. In this study, gait data from both groups were captured using markerless approach namely Kinect sensor. Three types of gait features are extracted namely direct joint feature, reference joint feature and center of mass feature. Additionally, all the features are classified using three different types of classifiers. Further, the effectiveness of the features for classification of walking gait pattern for ASD children is evaluated. Based on the results obtained, artificial neural network (ANN) outperformed the other two classifiers and results showed that the direct joint feature contributed to perfect classification followed by reference joint feature and center of mass feature.

Index Terms—ASD, gait classification, gait features, markerless gait, skeleton joints, walking gait

I. INTRODUCTION

Standard diagnostic criteria for Autism spectrum disorder (ASD) is characterized by severe impairments in social communication and skills, restricted and repetitive behaviors. According to the Diagnostic and statistical manual of mental disorders (DSM-5), existence of motor deficits which include abnormal gait, clumsiness and irregular motor signs are additional characteristics that support the diagnosis of ASD [1]. There are finding that children diagnosed with autism showed some form of different movement as compared to typical control group [2] and tend to be clumsy [3]. This result was supported by another study as reported in [4], that stated significant differences in kinematics and kinetics between autism and control group found on cadence, peak hip and peak ankle at kinematics and kinetics.

Walking gait is referring to the pattern of walking or locomotion. It is shown to be an important indicator of health as it used for diagnosis, monitoring, and rehabilitation [5].

There are several techniques for capturing and extraction of gait features. Recall that most sensor-based techniques for gait monitoring offer portable and wireless approach that is well suited to be used outside the laboratory environment. One example of sensor based techniques was discussed and used in previous studies as reported in [2], [6], [7], [8] & [9]. Marker-based techniques produced 3D data from markers trajectories.

The marker-based technique is an optoelectronic system that synchronizes between electronic devices and systems. Most of these systems are expensive and marker position must be correctly placed on the objects to produce an accurate model. Examples of marker-based techniques used by previous researches are as detailed in [10], [11] & [12]. Conversely, markerless gait technique used camera as the motion sensor detector, for capturing the gait features and parameters that offers low-cost system as compared to the marker-based system. However, the markerless system often requires additional precautions in maintaining its accuracy as it will affect the imperfect object segmentation and reduce the performance of the gait itself [13]. Such study related to ASD, J.A. Vilensky et al. elaborated the features by viewing the lateral films in microreader and chanting on paper for estimation of hip, knee and ankle joint centers locations and measured these parameters using protractor manually [14].

On the other hand, statistical analysis and machine learning method were implemented in gait classification between ASD and control groups too. Some researchers used approaches in statistical analysis methods for instance t-test [6, 15, 16], Pearson correlation [17], Ancova [18] and many [19-21]. Recently, other researchers used machine learning methods such as Linear Discriminant Analysis (LDA) [22], Neural Network [23], SVM [23] and k-nearest neighbor (kNN) [24]. However, the results of the classification groups may differ, depending on variables decision and analysis applications.

Although the importance of variability in gait studies was recognized many years ago, studies related to ASD gait analysis and classification is still going on. Based on previous researches published, walking gait classification focuses on markerless-based techniques are not many and can be further explored. Therefore, this study deems further in developing technique for classifying of walking gait in ASD children based on markerless gait features. This study mainly focuses on the extraction of potential gait features and types of classifiers in

This manuscript is submitted on 7th May 2019 and accepted on 3rd July 2019. Nur Khalidah Zakaria, Nooritawati Md Tahir and R. Jailani are from Faculty of Electrical Engineering, Universiti Teknologi MARA, Shah Alam, 40450, Selangor, Malaysia. (e-mail: rozita@ieee.org)
order to obtain more accurate classification accuracy. All gait features extracted are to be investigated and evaluated along with different classifiers namely artificial neural network (ANN), support vector machine (SVM) and Naïve Bayes classifier (NBC).

II. METHODOLOGY

This section explained in detail the method used in this study. Refer to Figure 1, the overall process methodology in this study is as shown. The process is divided into four sections specifically data acquisition, data preprocessing, data analysis and features extraction, gait classification and followed by performance measures.

A. Data Acquisition

A total of 30 typically development children (TD) and 21 autistic children (ASD) participated in this study. Firstly, parents were given an information sheet that need to be completed in the Consent Form that has been approved by the Universiti Teknologi MARA (UiTM) Shah Alam Ethic Committee for allowing their children to participate in this study. As mentioned earlier, Kinect sensor was used as the motion capture device. Recall that Kinect sensor is able to provide 20 skeleton joints for each frame which consist of xyz-coordinates for each joint in each frame captured [25]. All joints were connected to produce one skeleton stick figure as shown in Figure 2.

For data collection purpose, each subject was requested to perform at least 10 successful walking trials at their comfortable speed in front of the Kinect camera. Additionally, all subjects must be able to walk without any assistant or walking aid during trials. As agreed by the subject and parents or guardians, data collection for ASD children were collected more than ten trials in order to overcome the limitation of number of participant and walking trials from ASD children group.

B. Data Pre-processing

There are four steps involved in data pre-processing. Firstly, an empty frame from skeleton joint data was removed. Empty frames were the result from absence of subject in front of the Kinect while trial was captured. Recall that the skeleton joints occurred once the subject is captured at least 3 meters from the Kinect. Secondly, the skeleton joints were normalized based on the height of the references skeleton at reference frame which is half of the trimmed frame [26]. Next, gait cycle was extended from normalized skeleton joints frames. In this stage, gait cycle was computed based on the distance of the left and right ankle joints at z-coordinates respectively. Lastly, due to different...
number of frames between subject in each trials, the total frames was normalized using spline interpolation and the new total frames upon normalization were set at 30 frames [26]. It is observed that each subject have ten walking trials. For the TD group, the successful walking trials were 300, hence a total of 300 gait cycles were extracted. As for the ASD group, an extra gait cycle was extracted at the same trial so that equal number of gait cycle is achieved for both groups.

C. Data Analysis and Features Extraction

There are three main categories of gait features in this study. All features were extracted and calculated based on skeleton joints obtained from Kinect camera.

1) Direct joint feature

For direct joint feature, there are five features set (set 1-5) that were directly extracted from the skeleton joints as described below [25]:

- Set 1: All skeleton joint in xyz coordinates (point 1 to 20)
- Set 2: Skeleton joint at upper body in xyz coordinates (point 2 to 12)
- Set 3: Skeleton joint at lower body in xyz coordinates (point 1 to 2, 13 to 20)
- Set 4: Skeleton joint at right side of the body in xyz coordinate (point 1 to 4, 9 to 12, 17 to 20)
- Set 5: Skeleton joint at left side of the body in xyz coordinate (point 1 to 8, 13 to 16)

2) Reference joint feature

For reference joint feature, there are nine features set (set 6-14) that were calculated based on the distance between reference joint to other joint points using Euclidean distance in equation (1) [25, 27].

\[ d_i = \sqrt{(ref_x - j_i_x)^2 + (ref_y - j_i_y)^2 + (ref_z - j_i_z)^2} \]  (1)

where \( ref \) is a reference point and \( j_i \) is a second point (other joint point). The description of each set is as detailed below:

- Set 6: Reference joint is 1, to other joint points from 5 to 12
- Set 7: Reference joint is 1, to other joint points from 5 to 12
- Set 8: Reference joint is 1, to other joint points from 5 to 12
- Set 9: Reference joint is 1, to other joint points from 5 to 17 and 20
- Set 10: Reference joint is 1, to other joint points from 5 to 8 and 13 to 16
- Set 11: Reference joint is 5, to other joint points from 6 to 20
- Set 12: Reference joint is 9, to other joint points from 5 to 8 and 10 to 20
- Set 13: Reference joint is 11, to other joint points from 5 to 12 and 14 to 20
- Set 14: Reference joint is 13, to other joint points from 5 to 12, 13 to 16 and 18 to 20

3) Center of mass feature

Center of mass (COM) is defined as the physical point with the volumetric mass distribution is considered as balanced. While body segments are in motion, the center of mass of the whole body continuously changes over time [28]. There are three features set related to COM as computed from the skeleton joints. Recall that skeleton joints from Kinect provide 20 joints along with 19 body segments. Hence, 19 body segments were used to calculate the whole body COM namely both foot, leg, thigh, pelvis, upper arm, forearm, hand, shoulder as well as abdomen, thorax and head. Table 1 showed the defined of body segments with its segment length. Each body segment is calculated based on equation in (2), where proximal and distal are define as closest and furthest to the point of center of body.

\[ COM_{xyz} = (proximal_{xyz} + \left(\frac{\text{length}}{2}\right) (distal_{xyz} - proximal_{xyz})) \]  (2)

<table>
<thead>
<tr>
<th>Body Segment</th>
<th>Segment weight</th>
<th>Segment length (% length)</th>
<th>Proximal</th>
<th>Distal</th>
</tr>
</thead>
<tbody>
<tr>
<td>Head</td>
<td>0.081</td>
<td>1.000</td>
<td>3</td>
<td>4</td>
</tr>
<tr>
<td>Abdomen</td>
<td>0.139</td>
<td>0.440</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>Thorax</td>
<td>0.216</td>
<td>0.820</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>Left shoulder</td>
<td>-</td>
<td>0.712</td>
<td>3</td>
<td>5</td>
</tr>
<tr>
<td>Right shoulder</td>
<td></td>
<td>3.124</td>
<td>3</td>
<td>9</td>
</tr>
<tr>
<td>Left upper arm</td>
<td></td>
<td>0.028</td>
<td>5</td>
<td>6</td>
</tr>
<tr>
<td>Right upper arm</td>
<td></td>
<td>0.436</td>
<td>9</td>
<td>10</td>
</tr>
<tr>
<td>Left forearm</td>
<td>0.016</td>
<td>0.430</td>
<td>6</td>
<td>7</td>
</tr>
<tr>
<td>Right forearm</td>
<td></td>
<td>0.106</td>
<td>10</td>
<td>11</td>
</tr>
<tr>
<td>Left hand</td>
<td>0.006</td>
<td>0.506</td>
<td>7</td>
<td>8</td>
</tr>
<tr>
<td>Right hand</td>
<td>0.112</td>
<td>1.122</td>
<td>11</td>
<td>12</td>
</tr>
<tr>
<td>Left pelvis</td>
<td>0.142</td>
<td>0.105</td>
<td>1</td>
<td>13</td>
</tr>
<tr>
<td>Right pelvis</td>
<td>0.142</td>
<td>0.105</td>
<td>1</td>
<td>17</td>
</tr>
<tr>
<td>Left thigh</td>
<td>0.100</td>
<td>0.433</td>
<td>13</td>
<td>14</td>
</tr>
<tr>
<td>Right thigh</td>
<td>0.176</td>
<td>0.433</td>
<td>17</td>
<td>18</td>
</tr>
<tr>
<td>Left leg</td>
<td>0.0465</td>
<td>0.433</td>
<td>14</td>
<td>15</td>
</tr>
<tr>
<td>Right leg</td>
<td>0.184</td>
<td>0.433</td>
<td>18</td>
<td>19</td>
</tr>
<tr>
<td>Left foot</td>
<td>0.0145</td>
<td>0.500</td>
<td>15</td>
<td>16</td>
</tr>
<tr>
<td>Right foot</td>
<td></td>
<td></td>
<td>19</td>
<td>20</td>
</tr>
</tbody>
</table>

A total of 3 features were calculated in xyz axis. The description of each set is as below:

- Set 15: COM in x-axis
- Set 16: COM in y-axis
- Set 17: COM in z-axis

D. Gait Classification

This study proposed three type of classifiers namely artificial neural network (ANN), support vector machine (SVM) and Naive Bayes classifier (NBC). The classification stage was
performed using matrix-based language in Matlab version R2018a (The MathWorks Inc., USA). All features were fed into the network and the performance accuracy is measured. The inputs of the classifiers were the 17 features obtained from the feature extraction stage. For the output, 0 was set to ASD group and 1 to TD group. The performances of classifier models were evaluated using k-fold cross validation technique with k equal to 10.

1) Artificial Neural Network (ANN)

In this study, the ANN architecture consisted of three layers which are input layer, hidden layer and output layer. The process of training a neural network involves tuning the values of the weights and biases of the network to improve network performance using performance function; mean square error (mse) between the predicted outputs and target outputs. During training, weights are adjusted by employing the Scaled Gradient Conjugate algorithm. The default tan-sigmoid (tansig) transfer function is applied at the hidden and output layer. For the purpose of this study, the hidden neurons in hidden layer it was evaluated between 10 to 100 hidden neurons with 0.1 to 1 for threshold respectively.

2) Support Vector Machine (SVM)

SVM classifies data by finding the best hyperplane that separates all data points of one class from those of the other class. The best hyperplane for an SVM means the one with the largest margin between the two classes. Margin means the maximal width of the slab parallel to the hyperplane that has no interior data points. In this study, three type of kernel functions namely linear, radial basis function (RBF) and polynomial were evaluated with constraint parameter tuning between 0.001 and 1000. This parameter helps to prevent overfitting. On the other hand, kernel functions were used to compute the elements of \( x_j \) and \( x_k \) in \( V(x_j, x_k) \) where \( j \) and \( k \) is the observations in predictor data, \( x \). The mathematical approach using kernels depend on the computational method of hyperplanes. The kernel functions were depicted in Table 2.

<table>
<thead>
<tr>
<th>Kernel function</th>
<th>Mathematical formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>Linear</td>
<td>( V(x_j, x_k) = x_j'x_k )</td>
</tr>
<tr>
<td>RBF</td>
<td>( V(x_j, x_k) = \exp(-|x_j - x_k|^2) )</td>
</tr>
<tr>
<td>Polynomial</td>
<td>( V(x_j, x_k) = (1 + x_j'x_k)^q ) Where ( q ) is the polynomial order.</td>
</tr>
</tbody>
</table>

3) Naive Bayes Classifier (NBC)

Naive Bayes is a classification algorithm that assumes that the presence of a particular feature in a class is unrelated to the presence of any other feature. Its classifies new data based on the highest probability of its belonging to a particular class \( P(X|Y) \), given the densities, \( P \) of the predictors, \( X \) within each class, \( Y \). For classification purposed, two types of probability distribution were used which are normal or Gaussian distribution and kernel density estimation.

In normal distribution, the distribution is suitable for predictors that have normal distributions in each class. The NBC will estimate a separate normal distribution for each class by computing the mean and standard deviation of the training data in that particular class. Meanwhile for kernel distribution, it is suitable to be used for predictors with continuous distribution. For each NBC with kernel distribution, the classifier computes a separate kernel density to estimate each class based on the training data for that class. By default, the classifier will automatically select a suitable width for each class and predictor.

III. RESULTS AND DISCUSSIONS

This section will elaborate the classification results attained based on the three gait features discussed earlier. Table 3 showed the result of the demographic data for TD and ASD groups and presented the mean of height, weight and age for both TD and ASD group respectively. The range of age groups was from age 6 to 13 years old, and the mean aged for both groups were 9 and 8 years old for TD and ASD group respectively. The difference between height for both groups is 3.42 cm where TD group are slightly higher than ASD group. Meanwhile for weight, the mean value of TD group is 33.31 kg whilst ASD group with mean of 33.15 kg.

<table>
<thead>
<tr>
<th>TABLE 3 DEMOGRAPHIC DATA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total participant</td>
</tr>
<tr>
<td>Height, mean (cm)</td>
</tr>
<tr>
<td>Weight, mean (kg)</td>
</tr>
<tr>
<td>Age, mean (years old)</td>
</tr>
</tbody>
</table>

Conversely, kinetic sensor provided 3D skeleton joints that were used as input to data preprocessing. Thus, there would be 60 skeleton joints in one frame where the arrangement of the joints and skeleton points were vertically concatenated via frame by frame respectively. In one trial, at least three complete gait cycle can be extracted. For data processing, only one complete gait cycle was extracted for each trial. Gait cycle was computed based on the distance of the left and right ankle joints at z-coordinates respectively. Additionally, spline interpolation was used to normalize the number of frames and total new frames were set to 30 frames in each trial. Lastly, at the end of data preprocessing, total of 600 successful trials were processed in both groups.

A. Analysis of Features Extraction, Classifiers and Its Performance Measured

In this section, performance measure of ANN, SVM and NB classifiers will be discussed. The gait classification results were tabulated in Table 4 to Table 6 below. The bolded values are the highest accuracy for each feature.
TABLE 4
SUMMARIZED DIRECT JOINT FEATURES AND PERFORMANCE MEASURED

<table>
<thead>
<tr>
<th>Features set</th>
<th>Feature size, m</th>
<th>Performance measure, accuracy (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>ANN</td>
</tr>
<tr>
<td>1</td>
<td>1800</td>
<td>91.1</td>
</tr>
<tr>
<td>2</td>
<td>990</td>
<td>88.9</td>
</tr>
<tr>
<td>3</td>
<td>900</td>
<td>76.7</td>
</tr>
<tr>
<td>4</td>
<td>1080</td>
<td>92.2</td>
</tr>
<tr>
<td>5</td>
<td>1080</td>
<td>91.1</td>
</tr>
</tbody>
</table>

Table 4 showed the direct joint features and its performance measured respectively. These features were directly extracted from Kinect skeleton joints and arranged according to all skeleton points in set 1, upper and lower parts of body in set 2 and 3, right and left side of body in set 4 and 5 respectively. The determination of feature size, m was based on total of skeleton joints multiply by total number of frames which is 30 at xyz coordinate (example for Set 1: m = 20 x 30 x 3 = 1800). For feature set 1 to 3, the total of skeleton joints were 20, 11 and 10 points and the feature set size were 1800, 990 and 900 by 600 correspondingly. Meanwhile for set 4 and 5, total of skeleton joints were 12 points with feature size of 1080 by 600 for both.

Based on direct joint features, ANN outperformed with more than 88% accuracy except for feature set 3. In feature set 3, the highest accuracy was SVM classifier at 79.2% using linear kernel. From the results attained, ANN once again outperformed with direct joint features as input features. Additionally, NBC shows moderate performance for all direct joint features with more than 70% accuracy for each features.

Further, Table 5 tabulated reference joint features and its performance measured respectively. Nine features set that were calculated based on the distance between reference joint to other joint points. For these features set, the determination of feature size, m was based on the total of skeleton joints multiply by total number of frames which is 30 (example for Set 6: m = 16 x 30 = 480). For feature set 6, the total of skeleton joints was 16 points and the feature set size was 480 by 600. Meanwhile for set 7 to 10, the total skeleton joints were 8 points with feature size of 240 by 600 and for feature set of 11 to 14, the total of skeleton joints were 15 points with feature size of 450 by 600 correspondingly.

Next, five out of nine features set in reference joint feature were set 6, 8, 9 and 11 that showed SVM with linear kernel classifier attained highest accuracy at 81.8%. Based on results in Table 5, both ANN and SVM with linear kernel function performed well with reference joint as input features, while NBC shows moderate performance and SVM with RBF and Polynomial kernel shows least performance values.

Subsequently, Table 6 displayed COM features and its performance measures. There are three features set that were
calculated from each segment of the body. For these feature set, m is equal to total number of frames which is 30. All three features were the total of COM in each segment of the body in x, y and z axes correspondingly.

Based on COM features, the highest accuracy achieved was set 17 with 75.6% with ANN as classifier. In addition, set 15 and 16, achieved 73% and 72% accuracy with SVM classifier based on different kernel functions. It can be seen that the accuracy using ANN classifier is lower than 78% using the COM features as inputs due to lesser number of features as compared to the total number of direct and reference joint features. Additionally, with COM features as inputs, SVM with RBF kernel showed highest classification accuracy of 73% for feature set 15 and 72% for feature set 16 using linear SVM for feature set 16.

Table 4, 5 and 6 tabulated the performance measure for each classifier based on experimental analysis conducted. Based on six classifier models specifically ANN, SVM with linear, RBF and polynomial kernel along with NBC using normal distribution and kernel density, results attained showed that ANN outperformed other classifiers especially with direct joints features and reference joint features as inputs to classifiers. Next is SVM classifier for both linear and RBF kernels. This is related to observation made during data collection specifically the ASD children that participated in this study, it was observed that this group of children tend to make upper limbs movement especially hands movement such as waving, clapping and pointing out their finger during data acquisition of walking gait as compared to TD group. Note that this study includes all data of all the hands movement since there are no exclusion criteria for that. These is one of the reason for better classification results for direct joints features as inputs namely set 1, set 2, set 4 and set 5, as well as reference joint features specifically set 6, set 8, set 9, set 11 and set 12 as compared to the COM features. Since there was an alteration in the body segments during walking, the COM of the body was also effect. This is shown from the results attained for COM features with all classification accuracy below 76%.

IV. CONCLUSION

As a conclusion, this study has reported and presented the classification of walking gait features using markerless based approach for TD and ASD children. Based on performance measured of all three features namely direct joint features, reference joint features and COM features, it is observed that ANN outperformed all other classifiers. Next is SVM with linear kernel that achieved promising good accuracy as well. Conversely, SVM with polynomial showed least accuracies for almost all feature sets. In addition, the highest performance measured obtained was direct joint features and reference joint features were likely due to hand movements of the ASD subjects during data collection since the ASD children has less attention during experimental and data acquisition stage as compared to TD children. Hence direct and reference joint features can be used to determine the differences between walking gait pattern between TD and ASD children. Future work will focus on the implementation of different feature extractions algorithms with other possible classifiers and further investigation on possible intervention procedures.

V. ACKNOWLEDGMENT

The authors would like to acknowledge Institute of Research Management and Innovation (IRMI), Universiti Teknologi MARA (UiTM), Shah Alam, Selangor for the grant awarded No: 600-IRMI 5/3/GIP (084/2018), Human Motion Gait Analysis Laboratory, IRMI Premier Laboratory, for the instrumentation and experimental facilities provided as well as Faculty of Electrical Engineering (FKE), UiTM Shah Alam for the support given in this research. The authors also wish to acknowledge the National Autism Society of Malaysia (NASOM), all participants and their families, and colleagues for their contributions to the study.

REFERENCES


Nur Khalidah Zakaria received her B.Eng in Electronic Engineering (Instrumentation) and MSc in Electrical Engineering from Universiti Teknologi MARA, Malaysia. She is currently a postgraduate researcher at the Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia. Her research interests include gait analysis and pattern recognition using intelligent signal processing technique.

Nooritawati Md Tahir received her PhD in Electrical, Electronic & System Engineering from Universiti Kebangsaan Malaysia. She is currently a Professor at the Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia. Her research interests include pattern recognition and artificial intelligence.

Rozita Jailani received her PhD in Automatic Control and System Engineering from Sheffield University, UK. She is currently an Associate Professor at the Faculty of Electrical Engineering, Universiti Teknologi MARA, Malaysia. Her research interests include intelligent control system, rehabilitation engineering, instrumentation and artificial intelligence.