International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

Concealed Firearm Detection in Male and Female on Video using Machine Learning Classification: A Comparative Study

Henry Muchiri[1], Ismail Ateya[2], Gregory Wanyembi[3]
[1, 2]Faculty of Information Technology, Strathmore University, Nairobi, Kenya
[3]School of Computing and Informatics, Mount Kenya University, Thika, Kenya

Abstract— *An abnormal human gait has been identified as one of the key behavioral indicators used by CCTV operators to detect concealed firearm carrying. Gait has however been found to be significantly different in male and female more so that it has been proposed as a gender determinant. Using skeletal tracking and machine learning, this study sought to compare the performance of various machine learning algorithms to classify armed and unarmed features obtained from male and female depth datasets. The findings indicate that features selected from the datasets are significantly different. However, the classification ability of machine learning algorithms was almost similar for both datasets with K-nearest neighbor algorithm outperforming other algorithms for both datasets. These findings suggest that concealed firearm carrying affects male and females differently therefore a concealed firearm detector based on human motion should factor in these differences for a more holistic system.*

Keywords— *Concealed Firearm Detection; Depth Data; Human gait; Human Skeletal Tracking; Machine Learning.*

I. INTRODUCTION

The detection of persons carrying concealed firearms has become crucial in maintaining the safety and security of public places [1], [2]. This has been necessitated by the increase in crimes involving illegal firearms [3]. According to [4], [5] these firearms are normally carried in a concealed manner on a persons' body to the place where they will be used to commit a crime. The most common place to conceal a firearm has been found to be on the right hip to aid in easier access [4], [5]. According to [6], [4] a firearm concealed on the right hip results in the disruption in the gait of a person. This disruption according to [7] is used as one of the major indicators of concealed firearm carrying by trained CCTV operators. Gait is defined as the repetitious sequence of limb (Arms and Legs) motion that move the body forward [8].

Studies by [9], [10] have found that the use of CCTV operators to monitor video surveillance to be inefficient due to human weakness such as the inability to remain alert over long periods of time, task interruptions, and visual overload. Automation of this process has since been proposed by [11], [12]. In concealed firearm detection, automation using infrared and Passive millimetre wave imagers have been proposed by [11], [12]. However, these techniques have been found to be effective only when persons being screened are stationary and aware of the screening process [13], [14]. A desirable practical solution would be one where people are screened while they are in motion doing their businesses and are unaware of the screening process.

This study will train machine learning classifiers using RGB-D gait datasets to distinguish between armed and unarmed scenarios in male and female's and evaluate the performance. This is guided by studies by [15], [16], who have found that the gait of male and females is significantly different such that it has been used to distinguish between the genders with great accuracy and success.

The success of this approach will provide a concealed firearm detector that can be integrated into existing video surveillance networks and detect firearms when people are in motion and unaware of the screening process. In addition, a comparison of the classifier performance for male and female datasets will inform whether a generalized or specialized solution would be ideal.

II. MATERIALS AND METHODS

The Study did not find any existing RGB-D datasets of persons carrying concealed firearms hence created the data sets from scratch. The data creation procedure is described in the subsequent sub sections.

A. Participants

A total of 26 participants consisting of 19 male and 7 females with a mean age of 20.0 ± 1.3 years, were included in the study. General study group characteristics are as shown in table 1. The participants were selected using non-probabilistic convenience sampling technique. participants who reported any injury or deformity that affected their gait or were not predominantly right legged were excluded from the study.

TABLE 1. Participants Demographic Characteristics.

	Mean ± Standard Deviation
Age (Years)	20.0 ±1.3
Height (Meters)	1.7 ± 0.1
Weight (Kg)	65.8 ± 13.0

The study was approved by the Strathmore University Institutional Ethics Review Committee (SU-IERC) Certificate number SU-IRB 0246/18, dated 05.07.2018. In addition, a written consent was obtained from all participants prior to study procedures.

B. Materials

Ceska- ČZ 75 firearm-A ceska pistol loaded with 11 rounds of ammunition was used in the simulation. The pistol had a

Henry Muchiri, Ismail Ateya, and Gregory Wanyembi, "Concealed firearm detection in male and female on video using machine learning classification: A comparative study," *International Research Journal of Advanced Engineering and Science*, Volume 4, Issue 1, pp. 187-194, 2019.

International Research Journal of Advanced Engineering and Science
ISSN (Online): 2455-9024

weight of 2.77kg and a length of 0.2 meters. This firearm was used in the study since it's the most illegally used firearm in Kenya. Authorization to use the firearm was granted by the office of the Inspector General of Police in Kenya.

Microsoft Kinect RGB-D Sensor-The Microsoft Kinect Xbox one RGB-D sensor was used to record the videos. The sensor also referred to as Microsoft Kinect v2 was introduced by Microsoft in 2012 [17]. Originally the sensor was designed to be used as an Xbox accessory enabling players to interact with the Xbox through gestures or voice commands instead of controllers [17]. Reference [18] notes that the sensor has recently become popular for modelling and analyzing human motion with various authors [19], [20], [21] confirming its validity for motion capture. The sensor was used in this study due to the below benefits;

a) Lower price compared to other motion capture sensors [22].
b) High Accuracy and resolution as compared to its predecessor Kinect Xbox 360 [23].
c) Outdoor applicability as compared to Kinect Xbox 360 which was limited to indoors [24].
d) Improved skeletal tracking.

Microsoft Kinect v2 integrates four sensors into a single device. These sensors are (a) RGB color camera (b) A depth sensor (c) infrared emitters (d) Multi-array Microphones as shown in fig. *1* [17].

The four sensors generate four streams of data as illustrated by fig. 2; (a) Infrared data provided by the infrared sensor (b) depth and skeletal data provided by the depth sensor, (c) RGB data from the color camera and (d) voice provided by the multiarray microphones as illustrated in figure *2*. The sensor acquires a 512 × 424 depth map using the depth and infrared emitters and a 1920 × 1080 color image from the color camera at 15 to 30 frames per second (fps) depending on the lighting condition [24]. Each frame contains an array containing all the extracted points at the moment of capture. The infrared Camera can recognize and track up to six users in detail.

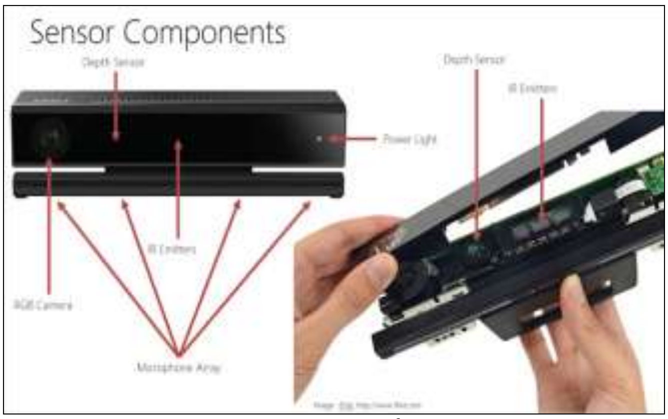

Fig. 1. Sensors in Microsoft Kinect v2.

Kinect for windows software Development Kit (SDK) v2.0-The Kinect for windows SDK is an opensource software that provides tools, device interfaces, API's and code examples for windows that allow programmers to develop applications based on the data streams (Video, depth and sound) generated by the Kinect v2 sensor [22].

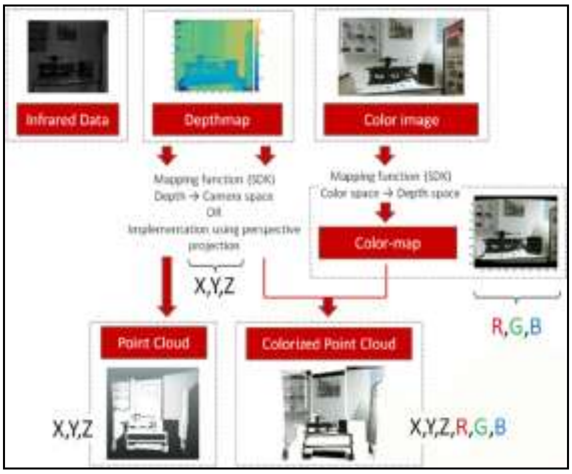

Fig. 2. Data Streams Generated by Kinect V2.

This study used the Kinect studio tool provided by the SDK for skeletal tracking. The recorded RGB-D videos were stored in *.xef* file format. Kinect Studio uses the skeleton tracking module in the SDK to provide detailed information about the position and orientation of 25 tracked joints on an individual located in front of the sensor [22]. Each tracked joint is identified by its name as illustrated by fig. 3. This position information is provided as a set of 25 three-dimensional skeleton depth point clouds [25]. The joint positions are provided as Euclidean distance in meters from the sensor. The origin (x=0, y=0, z=0) is located at the center of the IR sensor. X coordinate grows to the sensor's left (Subjects right), Y coordinate grows up and Z coordinate grows out in the direction the sensor is facing as shown in fig. 4. This point cloud depth data was used since it's immune to color variations, view-point variations, human appearance and lighting conditions [18], [26].

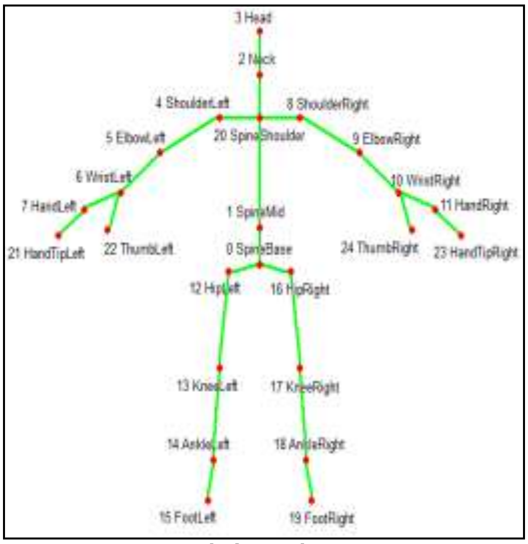

Fig. 3. 25 tracked Joints by Kinect SDK.

Fig. 4. 3D Coordinate placement on the Kinect Sensor.

Background Questionnaire- Participants responded to a background questionnaire regarding their gender, data of birth, height, weight and whether they had an injury or deformity that affected their gait.

C. Study Design

The experiment adopted a repeated measures design whereby all participants were recorded when armed and unarmed. The independent variables were armed and unarmed scenarios while the dependent variable was the 3D point cloud positions of the 25 tracked joints.

D. Procedure

Dataset Creation- The Study did not find any existing RGB-D datasets of persons' carrying concealed firearms hence created the data set from scratch. The participants motion was recorded using the Kinect v2 as they walked towards the sensor in two scenarios; (a) Armed (b) Unarmed. In the armed scenario, the ceska firearm was tucked and concealed on the right hip without a holster for both male and female participants. All Participants were dressed in trousers and a long jacket or sweater to conceal the firearm.

This video recordings were done in a brightly lit room with 6 meters long by 1 meter wide clearly marked footpath. The sensor was connected using a Kinect for windows adapter via USB 3.0 to a laptop with Intel(R) Core (TM)i7-7700HQ CPU @2.80GHZ processor, 16gb RAM and running Windows 10 operating system and Microsoft SDK v2.0. The sensor was elevated 1.2 meters above the ground and 1 meter away from one end of the walking path to ensure that the subjects were within the range and field of view of the sensor [27]. A physical marker was placed near the end of the footpath, so that the subjects were aware of where they needed to stop without having to look down which [28] found to skew the results. The video recordings for all 26 participants were stored and clearly labelled and distinguished as female and male and further as *armed* and *unarmed*. Each recording had an average of 80 frames.

To extract the skeletal joint depth information from the recorded RGB-D data, Kinect2 toolbox master application adopted from [29] was used. The depth information was dumped into text files which contained the joint number, point cloud 3D skeletal joint positions coordinates, and the skeleton tracking state which could either be Tracked-2, Not Tracked-0, or inferred-1. A sample text file is represented by fig 5.

Data Normalization- The skeletal joint position depth data obtained was based on the Euclidean distance between the sensor and the test subjects and was also dependent on the dimensions of the test subject such as height and limb length [30]. This data could not be used as provided and required normalization. The pair wise relative position-spatial displacement human feature representation technique adopted from the works of [31], [32], [30] was used for the normalization since it has been shown to yield data that is more independent with respect to the position between the sensor and the user and the person's specific size and speed of movement. This technique involved two stages.

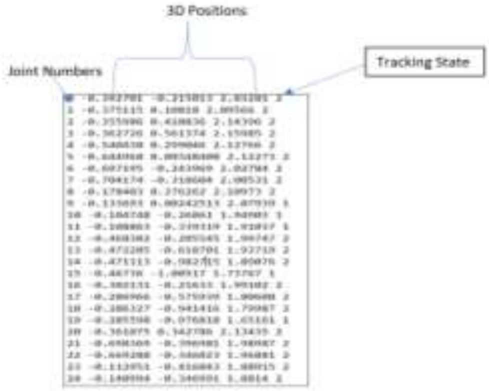

Fig. 5. Skeletal Joint Position Data.

In stage 1, the data was normalized to make it independent of the distance to the sensor. According to this normalization technique, the reference point was moved from the camera to the spine mid/torso joint by translation. The spine mid joint is represented as joint number 1. It was selected as the reference joint since it contains limited movement [31].

A normalized joint J_i was therefore obtained using 1.

$$J_i = j_i - J_1 \tag{1}$$

Where: J_1 is the spine mid joint location and j_i is the raw joint locations obtained from the Kinect sensor.

In stage 2, the invariance of the obtained data and the test subjects' dimension (height and limb length) was obtained by scaling each of the 25 joints with respect to the Euclidean distance between the neck joint and torso joint as show in 2. The obtained features were a set of distance vectors which connect each joint to the joint of the torso.

$$k_i = \frac{J_i}{\|d\|},\ 0 \geq \text{i} \leq 24 \tag{2}$$

Where: J_i is the partially normalized joint in 1

$\|d\|$ is the Euclidean distance between the neck and spine mid joint.

Ki are the fully normalized joint positions

These computations were performed using MATLAB (R2018a) numerical computing environment.

Feature Engineering- After data normalization, a feature vector f was obtained for each skeleton frame as illustrated by 3. A set of N feature vectors was computed, having a recording constituted by N frames.

$$f = [k_0, k_1, k_2, \ldots\ldots\ldots k_{24}] \tag{3}$$

Where: f represents the feature vector per frame and k are the fully normalized joint positions obtained from the data normalization step.

Henry Muchiri, Ismail Ateya, and Gregory Wanyembi, "Concealed firearm detection in male and female on video using machine learning classification: A comparative study," *International Research Journal of Advanced Engineering and Science*, Volume 4, Issue 1, pp. 187-194, 2019.

Further, all the features vectors f from all the N frames were combined into a row feature vector (FV) for each scenario (armed or unarmed) as shown by 4. All 25 joints were included in the feature vector resulting to 75 position features representing the 3-Dimension (X, Y, Z) for the 25 tracked joints.

$$FV = \left\{ \begin{array}{c} \left\{ \begin{array}{c} (J0_{X0}, J2_{X0}, \dots\dots\dots\dots J24_{X,0}) \\ (J0_{Y,0}, J1_{Y,0}, \dots\dots\dots . J24_{Y,0}), \\ (J0_{Z,0}, J1_{Z,0}, \dots\dots . J24_{Z,0}) \end{array} \right\} \\ \left\{ \begin{array}{c} (J0_{X,1}, J1_{X,1}, \dots . J24_{X,1}), \\ (J0_{Y,1}, J1_{Y,1}, \dots\dots\dots . J24_{Y,1}), \\ (J0_{Z,1}, J1_{Z,1}, \dots\dots . J24_{Z,1}) \end{array} \right\} \\ \left\{ \begin{array}{c} \vdots \end{array} \right\} \\ \left\{ \begin{array}{c} (J0_{X,n}, J1_{X,n}, \dots . J24_{X,n}), \\ (J0_{Y,n}, J1_{Y,n}, \dots\dots\dots . J24_{Y,n}), \\ (J0_{Z,n}, J1_{Z,n}, \dots\dots . J24_{Z,n}) \end{array} \right\} \end{array} \right\} \quad (4)$$

Where: $J1\ x,0$ represents the x component of the position of joint $J1$ in frame number 0, $J1\ y,0$ and $J1z,0$ represent the position of joint J1 in y and z respectively.

Lastly, both armed and unarmed features per gender dataset were combined to form the machine learning dataset. A new column was included to classify the features as either armed or unnamed classes. Each row/instance was considered individually. This resulted in a row vector with 75*N elements for each dataset. Where N is the number of frames. A detailed description of the datasets is provided by table 2.

TABLE 2. Description of Dataset.

		Instances		
Dataset	**No. of features**	**Armed**	**Unarmed**	**Total**
Female	75	511	581	1092
Male	75	1662	1771	3433

Feature Selection- The generated feature vectors described had high dimensionality as a result of having 75 features. Some of these features were noise and would reduce computational efficiency and result in overfitting [33]. To prevent this, the study conducted feature selection on both data sets using wrapper technique and best fit search approach. The aim of feature selection is to remove the irrelevant and redundant features [33]. Wrapper technique uses a predefined machine learning model to score feature subsets. Each new subset is used to train a model, which is then tested on a hold-out set. Since our dataset was small, this technique was specifically employed since it has been found to produce the best performing feature set in small datasets [33].

The classification algorithm employed with the wrapper technique was the C4.5-decision tree classifier. It was employed since it is one of the most prevalent and effective algorithms for supervised machine learning and is easy to train [34]. The classification algorithm used is not necessarily the algorithm to be used in modelling the problem. WEKA data mining software Version 3.8.2 was used to run the feature selection experiments. The selected features were then used to train various machine learning models.

Machine Learning Classification Algorithms- In order to develop a model to predict the presence of firearm carrying the study employed classification machine learning technique. Classification is a supervised machine learning technique whose aim is to predict the class/label of a given data points. The classification problem in this case was binary in nature for all datasets with armed and unarmed as the classes.

Naïve Bayes, K-Nearest Neighbor (K-NN), C4.5 Decision Tree, Random forest, Support vector machine, and Zero-Rule classification algorithms were employed. Their descriptions are provided below. Their performance was compared, and the algorithm that gave the best performance for each dataset was selected and compared.

The Naïve Bayes (NB) classifier is an algorithm in the family of probabilistic classifiers that calculates a set of probabilities by counting the frequency and combinations of values in a given data set [35]. The algorithm applies the Bayes theorem and theorem of total probability and assumes all attributes are independent given the value of the class variable [35]. That is, the presence of one particular feature does not affect the other (naïve) [36].

The K-NN/Instance based learning algorithm (IBK) is a non-parametric method, instance-based learning algorithm [37]. Instance-based algorithms are lazy-learning algorithms since they delay the generalization process until classification is performed. The algorithm is based on the principle that the instances within a dataset will generally exist near other instances that have similar properties [38]. If the instances are classified, then the class of an unclassified instance can be determined by observing the class of its nearest neighbors [39]. Instances are considered as points within an n-dimensional instance space. The absolute position of this instances within this space is not as significant as the relative distance between instances. This relative distance is determined by using a distance metric such as Euclidean, Manhatta or Minkowsky. Ideally, the distance metric must minimize the distance between two similarly classified instances, while maximizing the distance between instances of different classes [39].

The C4.5 is an algorithm used to generate decision trees [40]. A decision tree is a tree where each node represents a feature(attribute), each link(branch) represents a decision(rule) and each leaf represents the class. The decision trees are developed using the information entropy concept. The training data set consists of already classified samples. Each sample consists of a dimensional vector which represents the features of the sample, as well as the class it belongs to. At each node of the tree, the algorithm chooses the attribute of the data that most effectively splits its set of samples into subsets enriched in one class or the other. The splitting criterion is the difference in entropy. The attribute with the highest normalized information gain is chosen to make the decision. The algorithm then recurses on the partitioned sub lists. In Weka Version 3.8.2 data mining software, the C4.5 algorithm is implemented using the J48-open source java implementation.

Random forests or Random decision forests are an ensemble learning method that constructs a multitude of decision trees at training time [41]. The algorithm combines bagging sampling approach and the random selection of features to construct a collection of decision trees with controlled variation [42]. To classify a new object based on new attributes each tree gives a classification. This is referred to as a vote for a particular class. The forest chooses the classifications having the most votes from all the trees in the forest to classify the instance [42], [41].

A Support Vector Machine (SVM) is a discriminative classifier formally defined by a separating hyperplane [43]. In two-dimensional space this hyperplane is a line dividing a plane in two parts where in each class lays in either side. The SVM is motivated by the statistical learning theory developed by [44]. An SVM training algorithm builds a model of data points in space so that the data points of the separate categories are divided by a hyperplane. New examples are then mapped into that same space and predicted to belong to a category based on which side of the plane they fall on [45].

The Zero Rule/Zero-R algorithm provides the baseline results in a classification problem [46]. The baseline results are the class value that has the most observations in the training dataset and provides a point of reference/benchmark from which to compare other machine learning algorithms [46].

To run this classification algorithms, WEKA data mining software Version 3.8.2 was used.

Performance Evaluation Metrics- Performance evaluation metrics play a critical role in achieving the optimal classifier during classifier training. Thus, the selection of a suitable evaluation metric for a specific problem is key for discriminating and obtaining the optimal classifier [47]. Cross-validation was applied to split the data set into training and testing sets. The 10-fold cross-validation was specifically utilized since it has been found to be a good and reasonable compromise between providing robust performance estimates and being computationally feasible and avoids overfitting [48]. The value of 10 was chosen since it has been found through experimentation to generally result in a model with low bias. To apply the 10-fold cross-validation the dataset was split into 10 subsets of equal size. Then, over a total of 10 iterations, one of the subsets was held back as a test set, whilst the remaining nine were used to train the classification model. The model was then evaluated using the test set. The process was repeated 10 times, so that each subset was used once as a testing set. This yielded 10 sets of performance values, and their average represented the cross-validated performance estimates for the classification model.

To evaluate the classifier performance, this study employed the confusion matrix. For a binary class problem, the matrix is a square of 2×2 as shown in table 3. The columns represent the classifier prediction, while the rows represent the class labels.

TABLE 3. Confusion matrix for Binary classifier.

	Predicted Outcome		
Actual Value	**Class A**	**Class B**	<-Classified as
	TP	FN	**Class A**
	FP	TN	**Class B**

The acronyms TP, FN, FP, and TN refer to the following: TP = True positive, the number of positive cases that are correctly identified as positive, FN = False negative, the number of negatives cases that are misclassified as positive cases, FP = False positive, the number of positive cases that are incorrectly identified as Negative cases, TN = True negative, the number of negative cases that are correctly identified as negative cases

A series of performance measures were defined based on the confusion matrix. These measures are usually dependent on the balanced or imbalanced nature of the dataset. As presented in table 2, the male dataset had a total of 3433 instances with 1662 instances (48.41%) representing armed class and 1771 instances (51.59%) representing unarmed class. The female dataset had a total of 1090 instances with 510 instances (46.79%) representing armed class and 580 instances (53.21%) representing unarmed class. The datasets were therefore balanced, and the following balanced measures were employed for the classification problem a) Accuracy b) Precision c) Recall /sensitivity d) Specificity e) F-Measure f) Cohen Kappa Statistic. A description of this metrics is provided below.

a) Classification Accuracy/Correctly Classified Instances: The accuracy of a classifier is the percentage of test set instances that are correctly classified [49]. The computation is as shown in 4.

$$\frac{TP+T_n}{TP+FP+FN+T_n} \tag{4}$$

b) Precision (p): It is the proportion of predicted positive which are actual positives. A high precision means that an algorithm returned substantially more relevant results than irrelevant ones. The formula is defined as shown in 5 [50].

$$P=\frac{TP}{TP+FP} \tag{5}$$

c) Recall/Sensitivity(R): It is the proportion of actual positives which are predicted positive. It measures how good a test is at detecting the positives by avoiding of false negatives. This may sometimes result in many false positives/alarms. The formula is defined as shown in 6.

$$R=\frac{TP}{TP+FN} \tag{6}$$

d) Specificity: This measure approximates the probability of the negative label being true [50]. It is a measure of how good a test is at reducing false alarms.

$$\frac{TN}{TN+FP} \tag{7}$$

e) F-Measure: F-measure is a measure that combines precision and recall and is considered as the harmonic mean of precision and recall [45].

$$\frac{2(PR)}{P+R} \tag{8}$$

f) Cohen's Kappa Statistic: This is a metric that compares an observed accuracy with an expected accuracy. The expected accuracy is described as random accuracy generated by a random classifier. The score can range from 0 to +1, where 0 represents the amount of agreement that can be expected from random chance, and 1 represents perfect agreement. According to [51] the Kappa result can be interpreted as follows: values ≤ 0 as indicating no agreement and 0.01–0.20 as none to slight, 0.21–0.39 as minimal, 0.40– 0.59 as weak, 0.60–0.79 as moderate, and 0.80–0.90 as strong and above 90% to be almost perfect agreement. 80% agreement is considered as the minimum acceptable interrater agreement.

The desirable machine learning model for concealed firearm detection should report a high F-Measure score which means the model will report a high recall or sensitivity (ability to detect a good proportion of armed persons (True Positive rate)) and as well have a high specificity meaning it should report unarmed persons correctly (True Negative rate)

III. RESULTS

This section reports on the results from the Feature selection and Machine learning experiments

A. Feature Selection Experiment Results

18 features were selected from the male dataset while 14 features were selected from the female dataset as presented in table 4. The features represent the coordinate axis and the joint number. For instance, X-7 refers to the joint seven (7) X-axis feature.

TABLE 4. Selected features for the Male and Female Datasets.

Dataset	Features Selected
Male	X-7, X-9, X-11 X-12, X-14, X-17, X-20, X-21, X-22, X-24, Y-0, Y-8, Y-13, Y-19, Z-0, Z-2, Z-6, Z-12
Female	X-0, X-3, X-4, X-11, X-12, X-22, X-24, Y-0, Y-2, Y-3, Y-8, Y-11, Y-23, Z-3

Fig. 6a shows the positions of the selected joints in the male dataset and fig. 6b represents the positions of the selected joints in the female dataset on the human skeleton for better visualization.

This selected feature datasets can be accessed from fig share.com with the file name concealed firearm detection male and female datasets.

B. Machine Learning Experiment Results

The confusion matrix of the machine learning algorithms is presented in table 5 for the male dataset and in table 6 for the female dataset.

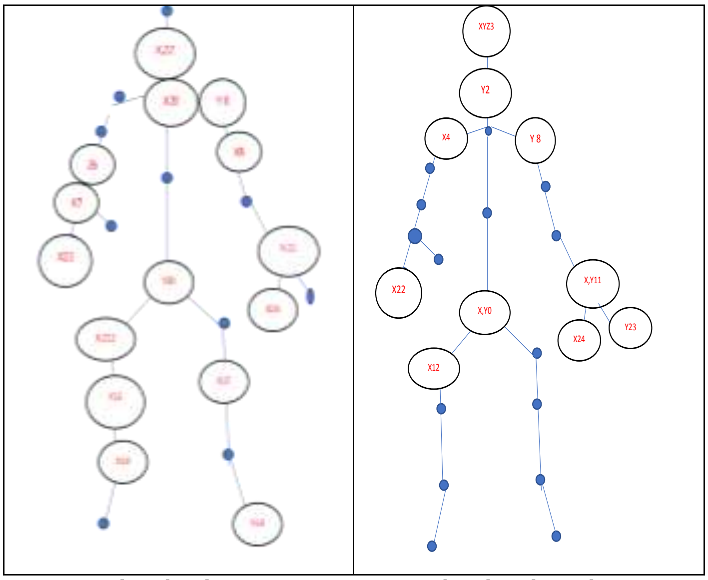

Fig. 6a. Selected Male Features. Fig. 6b. Selected Female Features.

TABLE 5. Confusion Matrix for the Male Dataset.

Algorithm	TP	FP	TN	FN
Zero-R Baseline	0	0	1771	1662
Naïve Bayes	373	235	1536	1289
kNN/IBK	1627	50	1721	35
C4.5 Decision Tree	1354	298	1473	308
Random Forest	1573	88	1683	89
Support Vector Machine (SVM)	872	530	1241	790

TABLE 6. Confusion Matrix for the Female Dataset.

Algorithm	TP	FP	TN	FN
Zero-R Baseline	0	0	580	510
Naïve Bayes	370	272	308	140
kNN/IBK	503	8	572	7
C4.5 Decision Tree	438	64	516	72
Random Forest	492	16	564	18
Support Vector Machine (SVM)	292	105	475	216

The classification performance of the considered machine learning algorithms is presented in table 7 for the male dataset and in table 8 for the female dataset.

TABLE 7. Classification Performance on the Male Dataset.

Algorithm	Accuracy	Precision	Recall/Sensitivity	Specificity	F-Measure	Cohen-Kappa Statistic
Zero-R Baseline	0.516	0.000	0.516	0.516	0.000	0.000
Naïve Bayes	0.541	0.481	0.541	0.533	0.473	0.620
kNN/IBK	0.978	0.978	0978	0.979	0.978	0.956
C4.5 Decision Tree	0.824	0.824	0.824	0.827	0.824	0.648
Random Forest	0.946	0946	0.946	0.945	0.946	0.891
Support Vector Machine (SVM)	0.616	0.617	0.616	0.611	0.613	0.228

TABLE 8. Classification Performance on the Female Dataset.

Algorithm	Accuracy	Precision	Recall/Sensitivity	Specificity	F-Measure	Cohen-Kappa Statistic
Zero-R Baseline	0.532	0.000	0.532	0.532	0.000	0.000
Naïve Bayes	0.622	0.635	0.622	0.688	0.619	0.253
kNN/IBK	0.986	0.986	0.986	0.988	0.986	0.972
C4.5 Decision Tree	0.875	0.875	0.875	0.878	0.875	0.749
Random Forest	0.967	0.967	0.967	0.967	0.967	0.934
Support Vector Machine (SVM)	0.706	0.711	0.706	0.687	0.700	0.401

IV. DISCUSSION OF RESULTS

The findings demonstrate that carrying a concealed firearm affects the motion of male and female's differently. This is demonstrated by figure 6a and figure 6b where different features were selected from the male and female datasets during the feature selection experiment. This agrees with previous studies by [15], [16] that found a difference in motion for men and women.

The findings of the feature selection experiment show that a majority of features selected in both datasets were mostly those representing the X-axis. This is an indication that carrying a firearm concealed on the right hip has a greater impact on the left and right movement along this axis.

The findings show that the all algorithms outperformed the baseline algorithm in all metrics. This is an indication that the algorithm performance was not by chance. The K-Nearest Neighbor algorithm performed overall best in all metrics for both male and female datasets. However, the algorithm performed better in classifying the female dataset as compared to the male dataset. The K-Nearest Neighbor algorithm reported a high F-Measure performance for all datasets meaning that the algorithm had a good balance between specificity and sensitivity.

Naïve Bayes algorithm performed poorly in all metrics for the female dataset. However, in the male dataset, the algorithm performed poorly but had a higher Cohen kappa statistic as compared to support vector machine (SVM) algorithm. This is an indication that the performance reported by SVM algorithm largely occurred by chance. Therefore, the algorithm had the poorest performance in the male dataset.

The empirical results reported herein should be considered in the light of the limitations that the data collection was performed in a controlled lab environment which simulated participants carrying a concealed firearm.

V. CONCLUSION

The detection of persons carrying concealed firearms is crucial in maintaining the safety and security in public places. Simulated videos of male and female participants carrying a concealed firearm tucked on the right hip were recorded using a depth sensor. The high dimensionality of the collected data was reduced using wrapper feature selection technique. The features selected from the male and female datasets were significantly different signifying that the concealed firearm had a different effect on the motion of male and female. The development therefore of machine learning models for concealed firearm detection based on motion characteristics must include data from both men and women for a more holistic model.

Various machine learning algorithms were employed to classify the selected features from the male and female datasets and their performance was compared. K-Nearest Neighbor algorithm outperformed all other algorithms. Performance of the algorithm in the female dataset was better with 98.6% F-Measure and accuracy scores while in the male data set a 97.8% F-Measure and accuracy score was reported. This high F-Measure score signifies the ability of the algorithm to have a high firearm detection rate while maintaining a low false alarm rate. This demonstrates the viability of the algorithm to develop concealed firearm detector models that perform well regardless of gender.

REFERENCES

[1] S. K. Bandyopadhyay, B. Datta and S. Roy, "Identifications of concealed weapon in a human," *International Journal of Scientific & Engineering Research*, pp. 1-7, 2012.
[2] J. Yuan and C. Guo, "A deep learning method for detection of dangerous equipment," in *8th International Conference on Information Science and Technology*, Seville, 2018.
[3] Gun-Violence-Archive, "past-tolls," Gun Violence Archive, 14 September 2018. [Online]. Available: http://www.gunviolencearchive.org/past-tolls. [Accessed 14 September 2018].
[4] N. C. Meehan and C. Strange, "Behavioral indicators of legal and illegal gun carrying," Naval Research Laboratory, Washington, DC, 2015.
[5] K. Porter, "Characteristics of the armed individual," U.S. Secret Service Publication, Washington, 2010.
[6] I. T. Darker, A. G. Gale and A. Blechko, "CCTV as an automated sensor for firearms detection: Human-derived performance as a precursor to automatic recognision," *Society of Photo-Optical Instrumentation Engineers*, vol. 7112, pp. 73-80, 2008.
[7] T. Darker, Blechko and A. G. Gale, "The role of emotion recognition from nonverbal behaviour in detection of concealed firearm carrying," in *Human Factors and Ergonomics Society 53rd Annual Conference*, 2009.
[8] A. Kharb, V. Saini, Y. K. Jain and S. Dhiman, "A review of gait cycle and its parameters," *IJCEM International Journal of Computational Engineering & Management*, pp. 78-84, 2011.
[9] R. Arroyo, J. Yebes, L. M. Bergasa, I. G. Daza and J. Almazán, "Expert video-surveillance system for real-time detection of suspicious behaviors in shopping malls," *Expert Systems with Applications*, pp. 7991-8005, 2015.
[10] T. SenthilKimar and G. Narmatha, "Suspicious human activity detection in classroom examiniation," Springer Science+Business Media B.V, pp. 99-105, 2016.
[11] V. M. Patel, J. N. Mait, D. W. Prather and A. Hedden, "Computational millimeter wave imaging: Problems, progress, and prospects," *IEEE Signal Processing Magazine*, pp. 109-118, 2016.
[12] N. Salmon, "Passive millimeter wave imagers for safety and security," *SPIE*, pp. 1-4, 2008.
[13] S. M. Hegde, K. Shivaprasad and N. Neh, "Imaging for concealed weapon detection," in *2015 National Conference "Electronics, Signals, Communication and Optimization*", 2015.
[14] N. M. Sirakov, "Firearms identification through partonomy," in *SPIE Conference*, Maryland, 2015.
[15] D. A. Bruening, R. E. Frimenko, C. D. Goodyear, D. R. Bowden and A. M. Fullenkamp, "Sex differences in whole body gait kinematics at preferred speeds," *Pubmed*, pp. 540-545, 2015.
[16] O. M. Hassan, A. M. Abdulazeez and M. V. TİRYAKİ, "Gait-Based human gender classification using lifting 5/3 wavelet and principal component analysis," in *International Conference on Advanced Science and Engineering (ICOASE)*, Duhok, 2018.
[17] N. Khamsemanan, C. Nattee and N. Jianwattanapaisarn, "Human identification from freestyle walks using posture-based gait feature," *IEEE Transactions on Information Forensics and Security*, vol. 13, no. 1, pp. 119-128, 2018.
[18] L. Tao, P. Adeline, D. Damen, M. Mirmehdi, S. Hannuna, M. Camplani, T. Burghardt and I. Craddock, "A comparative study of pose representation and dynamics modelling for online motion quality assessment," *Elsevier*, pp. 136-152, 2016.
[19] K. Otte, B. Kayser, S. Mansow-Model, J. Verrel, F. Paul, A. U. Brandt and T. Schmitz-Hübsch, "Accuracy and reliability of the Kinect version 2 for clinical measurement of motor function," *Plos One*, pp. 1-17, 2016.
[20] B. Müller, W. Ilg, M. A. Giese and N. Ludolph, "Validation of enhanced Kinect sensor-based motion capturing for gait assessment," *Plos One*, pp. 1-18, 2017.

[21] J. Mahvash, B. Alireza and A. Ahmadrez, "A study on validating KinectV2 in comparison of Vicon system as a motion capture system for using in health engineering in industry," *Nonlinear Engineering*, pp. 95-99, 2017.
[22] S. K. Jarraya, "Computer vision-based fall detection methods using the Kinect camera: A survey," *International Journal of Computer Science & Information Technology*, pp. 73-92, 2015.
[23] G. McFassel, S.-J. Hsieh and B. Peng, "Prototyping and evaluation of interactive and customized interface and control algorithms for roboticassistive devices using Kinect and infrared sensor," *International Journal of Advanced Robotic Systems*, pp. 1-9, 2018.
[24] S. Zennaro, M. Munaro, S. Milani, P. Zanuttigh, A. Bernardi, S. Ghidoni and E. Menegatti, "Performance Evaluation of the 1st And 2nd Generation Kinect For Multimedia Applications," in *IEEE International Conference on Multimedia and Expo (ICME)*, Turin, 2015.
[25] E. Lachat, H. Macher, T. Landes and P. Grussenmeyer, "Assessment and Calibration of a RGB-D Camera (Kinect v2 Sensor) Towards a Potential Use for Close-Range 3D Modeling," *Remote Sensing*, pp. 13070-13097, 2015.
[26] F. Han, B. Reily, W. Hoff and H. Zhang, "Space-Time representation of people based on 3D skeletal data: A review," *Computer Vision and Image Understanding*, pp. 85-105, 2015.
[27] C. Tucker, I. Behoora, H. B. Nembhard, M. Lewis, N. W. Sterling and X. Huang, "Machine learning classification of medication adherence in patients," *Elsevier*, pp. 1-15, 2015.
[28] S. Liu, X. Meng and C. Tam, "Building information modeling-based building design optimization for sustainability," *Energy and Buildings* 105, pp. 139-153, 2015.
[29] Z. Yixin, Z. Yibiao and Z. Song-Chun, "Understanding tools: Task-oriented object modeling, learning and recognition," in *IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, Boston, Massachusetts, 2015.
[30] A. Manzi, L. Fiorini, R. Limosani, P. Dario and F. Cavallo, "Two-person activity recognition using skeleton data," *IET Computer Vision*, pp. 27-35, 2018.
[31] E. Cippitelli, S. Gasparrini, E. Gambi and S. Spinsante, "A human activity recognition system using skeleton data from RGBD sensors," *Computational Intelligence and Neuroscience*, pp. 1-15, 2016.
[32] S. Gaglio, G. LoRe and M. Morana, "Human activity recognition process using 3-D posture data," *IEEE Transactions on Human-Machine Systems*, pp. 586-597, 2015.
[33] K. Sutha and J. Tamilselvi, "A review of feature selection algorithms for data mining techniques," *International Journal on Computer Science and Engineering (IJCSE)*, pp. 63-67, 2015.
[34] B. N. Lakshmi, T. S. Indumathi and N. Ravi, "A study on C.5 decision tree classification algorithm for risk predictions during pregnancy," in *International Conference on Emerging Trends in Engineering, Science and Technology (ICETEST- 2015)*, Kerala, 2015.
[35] T.R. Patil and S. S. Sherekar, "Performance analysis of naive Bayes and J48 classification algorithm for data classification," *International Journal of Computer Science and Appli*, pp. 256-261, 2013.
[36] P. Domingo and M. Pazzani, "On the optimality of the simple Bayesian classifier under zero-one loss," *Machine Learning*, pp. 103-130, 1997.
[37] N. Altman, "An introduction to kernel and nearest-neighbor nonparametric regression," *The American Statistician*, pp. 175-185, 1992.
[38] T. Cover and P. Hart, "Nearest neighbor pattern classification," *IEEE Transactions on Information Theory*, pp. 21-27, 1967.
[39] S. B. Kotsiantis, "Supervised machine learning: A review of classification techniques," in *Proceedings of the 2007 conference on Emerging Artificial Intelligence Applications in Computer Engineering: Real Word AI Systems with Applications in eHealth, HCI, Information Retrieval and Pervasive Technologies*, The Netherlands, 2007.
[40] R. Quinlan, *C4.5: Programs for Machine Learning*, Burlington: Morgan Kaufmann Publishers, 1992.
[41] T. K. Ho, "The random subspace method for constructing decision forests," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, p. 832–844, 1998.
[42] K. Fawagreh, M. M. Gaber and E. Elyan, "Random forests: From early developments to recent advancements," *Systems Science & Control Engineering*, pp. 602-609, 2014.
[43] C. Cortes and V. Vladimir, "Support-vector networks," *Machine Learning*, pp. 273-297, 1995.
[44] V. Vapnik, *The Nature of Statistical Learning Theory*, New York: Springer, 1998.
[45] S. Amarappa and. S. V. Sathyanarayana, "Data classification using Support vector Machine (SVM), a simplified approach," *International Journal of Electronics and Computer Science Engineering*, pp. 435-445, 2014.
[46] C. Nasa and Suman, "Evaluation of Different Classification Techniques for WEB Data," *International Journal of Computer Application*, pp. 34-41, 2012.
[47] M. Hossin and M. N. Sulaiman, "A review one valuation metrics for data classification evaluations," *International Journal of Data Mining & Knowledge Management Process*, pp. 1-11, 2015.
[48] N. Mansbridge, J. Mitsch, N. Bollard, K. Ellis, G. G. Miguel-Pacheco, T. Dottorini and J. Kaler, "Feature selection and comparison of machine learning algorithms in classification of grazing and rumination behaviour in sheep," *Sensors*, pp. 1-16, 2018.
[49] O. Miho, W. peng, M. Kenji, K. Shigeru, W. Hideyuki and R. Anca, "Confusion-Matrix-Based kernel logistic regression for imbalanced data classification," *IEEE Transactions on Knowledge and Data Engineering*, pp. 1806-1819, 2017.
[50] M. Bekkar, H. K. Djemaa and T. A. Alitouche, "Evaluation measures for models assessment over imbalanced data sets," *Journal of Information Engineering and Application*, pp. 27-40, 2013.
[51] M. L. McHugh, "Interrater reliability: The kappa statistic," *Biochem Med (Zagreb)*, p. 276–282, 2012.