

## Human gait recognition using preprocessing and classification techniques

Samer Kais Jameel, Jihad Anwar Qadir, Mohammed Hussein Ahmed

Department of Computer Science, University of Raparin, Iraq

---

### Article Info

#### Article history:

Received Jul 6, 2019

Revised Nov 26, 2019

Accepted Dec 10, 2019

---

#### Keywords:

Classification

Gait recognition

Kinect sensor

Preprocessing

---

### ABSTRACT

Biometric recognition systems have been attracted numerous researchers since they attempt to overcome the problems and factors weakening these systems including problems of obtaining images indeed not appearing the resolution or the object completely. In this work, the object movement reliance was considered to distinguish the human through his/her gait. Some losing features probably weaken the system's capability in recognizing the people, hence, we propose using all data recorded by the Kinect sensor with no employing the feature extraction methods based on the literature. In these studies, coordinates of 20 points are recorded for each person in various genders and ages, walking with various directions and speeds, creating 8404 constraints. Moreover, pre-processing methods are utilized to measure its influences on the system efficiency through testing on six types of classifiers. Within the proposed approach, a noteworthy recognition rate was obtained reaching 91% without examining the descriptors.

*Copyright © 2020 Institute of Advanced Engineering and Science.  
All rights reserved.*

---

### Corresponding Author:

Samer Kais Jameel,  
Department of Computer Science,  
University of Raparin,  
Rania, Iraq.  
Email: samer.kais@uor.edu.krd

---

## 1. INTRODUCTION

The automated identifying system has improved barely in recent decades, in particular in excessive safety areas e.g. Airports and banks Biometric authentication makes use of organic or behavioral traits to verify the identification of a person [1, 2]. In the midst of the violations that occur constantly in most parts of the globe a multiplied attention has been taken to the prevention of terrorist attacks, throughout build sophisticated and swift systems to identify the humans. Many biometric technologies have emerged for identifying and verifying persons through analyzing face, fingerprint, palm print, iris, gait, or a mixture of these traits [3-6].

Biometrics is the automatic use of physiological or behavioral traits to determine or confirm the identification of a person, the physiological biometrics examines physiological characteristics like, iris, faces, fingerprints, DNA, and hand geometry; the behavioral biometrics examines behavioral issues, such as voice, signature, and gait [1]. The significance of computerized identification of people has accelerated throughout the previous decades, particularly in excessive security areas such as airports and banks [5, 7]. Each character has distinguishable unique traits (biometrics) that can be used by identification systems to verify and pick out the person's identity [8]. To distinguish unique humans in the manner they stroll is a horrible project humans performed each day. Psychological studies [4, 8, 9] have shown that gait signatures bought from video can be used as a reliable cue to become aware of individuals. These findings stimulated researchers in computer imaginative and prescient to extract potential gait signatures from pictures to pick out people. It is challenging, however, to locate idiosyncratic gait aspects in marker-less motion sequences, where the use of markers is prevented due to the fact it is intrusive and not suitable in commonplace gait

cognizance settings. Ideally, the focus features extracted from photos be invariant to elements different than gait, such as color, texture, or kind of clothing [9]. Two Compared to different biometric methods, gait consciousness offers several unique characteristics. The most appealing characteristic is its unobtrusiveness, which does no longer require observed subjects' attention and cooperation. In addition, human gait can be captured at some distance besides requiring physical information from subjects. This favorable characteristic has brilliant advantages, mainly when man or woman facts such as face image are confidential [5, 8]. Moreover, gait recognition presents superb achievable for consciousness of low-resolution videos, where other biometrics technologies may additionally be invalid because of insufficient pixels to identify the human topics [9]. The common framework of computerized gait focus consists of challenge detection, silhouette extraction, function extraction, function selection, and classification. Once transferring topics are captured, humans will be detected and separated from the image background [5].

This paper was aimed to develop an integrated and sophisticated system for recognizing humans in terms of a set of points found by using a Kinect sensor, particularly since not requiring a camera directly focused on the human face or on any human biometric like other systems as a major challenge in some cases. The filters (Resample, Discretize, and Spread sub-sample) were used as a step pre-processing to treat the data and to minimize long time courses while increasing system performance, which is measured via pass the data into six types of classifiers (Sequential Minimal Optimization, Decision Tree, Naïve Bayes, Random Tree, Rule, and Bayes Net). The influence of the filters on system efficiency was tested to understand the appropriateness of filters for such classifiers and their positive effects in identifying systems.

## 2. REVIEW OF LITERATURE

Different ways of recognizing people by their gaits are wide mentioned within the literature for several years. The first add psychologists administrated this space in 1971, once Johansson connected light-weight points to the joints of people's bodies during a dark space. Participants were then asked to run, run, or ride a bicycle [10, 11]. The results prompt that individuals will acknowledge one another by their individual walking designs. The biomechanics studies of Perry et al. [12], Murray [13] and Winter [14] semiconductor diode to the idea that gait may be a characteristic and probably individual attribute of an individual. Gait recognition may be a pattern recognition drawback. Most of the prevailing gait recognition approaches depend upon AN analysis of the binary silhouette of walking persons for identification [15, 16]. Cutting, et al. [17] studied human perception of gait exploitation moving light-weight displays (MLD) the same as that employed by Johansson and showed human person identification results [18] and gender classification results [19]. They showed that human observers may determine gender with more or less seventieth accuracy exploitation solely the visual cues from MLD. Bobick and Johnson [20] calculate four distances of human bodies, particularly the gap between the pinnacle and foot, the gap between the pinnacle and pelvis, the gap between the foot and pelvis, and therefore the distance between the left foot and right foot. They use the four distances to make 2 teams of static body parameters and reveal that the second set of parameters are additional view-invariant comparison to the primary set of body parameters. Given the flexibility of humans to spot persons and classify gender by the joint angles of a walking subject, Robert Hutchings Goddard [21] developed a connectionist formula for gait recognition exploitation joint locations obtained from MLD. However, computing joint angles from video sequence remains a tough drawback, although many tries are created thereon [22-24]. There is a variety of looks based mostly on algorithms for gait and activity recognition. Dealer and Davis [25] used self-correlation of moving foreground objects to tell apart walking humans from different moving objects like cars. Polana and Nelson [26] detected regularity in optical flow and used these to acknowledge activities like frogs jumping and human walking. Very little and Boyd [27] used moment options and regularity of foreground silhouettes and optical flow to spot walkers. Nixon, et al. [28] used principal part analysis of pictures of a walking person to spot the walker by gait. Shutler, et al. [29] used higher-order moments summed over sequent pictures of a walking sequence as options within the task of distinguishing persons by their gait. The work delineated during this paper is closely associated with that of very little and Boyd [30]. However, rather than exploitation moment descriptions and regularity of the whole silhouette and optical flow of a walker, we tend to divide the silhouettes into regions and cipher statistics on these regions. We tend to add more study the capability of our options in tasks on the far side person identification, like gender classification.

## 3. DATASET

There are a few gait sensors databases, and the data used in this paper has been collected using a Kinect sensors device. The gait was recorded for a group of volunteers of 49 persons, 9 of whom were women and the rest were men. Each person was tested for walking five times for both left and right

directions, in front of the sensor at a 90-degree angle, and height of the 0.6 meters from the ground. The Kinect sensor presents the human skeleton see Figure 1, which provides values of X, and Y coordinates for 20 points from head to toe, where each person was registered between 90 and 190 records, which results in 8404 records for all persons.

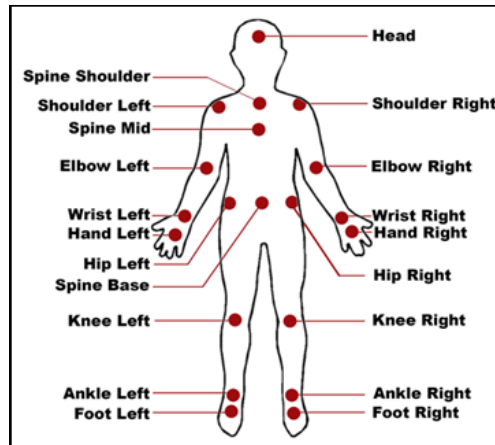


Figure 1. A human skeleton by Kinect sensor

## 4. BACKGROUND

### 4.1. Bayesian networks

Bayesian networks are a kind of probabilistic graphical mannequin that uses Bayesian inference for chance computations. Bayesian networks goal to mannequin conditional dependence, and therefore causation, via representing conditional dependence by using edges in a directed graph. Through these relationships, one can correctly habit inference on the random variables in the graph through the use of factors [31, 32]. Two it is also regarded as "belief networks" or "causal networks" are graphical fashions for representing multivariate chance distributions. Each variable  $X_i$  is represented as a vertex in a directed acyclic graph ("dag"); the probability distribution  $P(x_1, x_2, x_3, \dots, x_n)$  is represented in factorized form as follows:

$$P(X_1, X_2, X_3, \dots, X_n) = \prod_{i=1}^N p(X_i | \Pi_{X_i})$$

where  $\Pi_{X_i}$  is the set of vertices that are  $X_i$ 's parents in the graph. A Bayesian network is fully specified by the combination of:

- The graph structure, i.e., what directed arcs exist in the graph.
- The probability table  $p(X_i | \Pi_{X_i})$  for each variable  $X_i$ .

It can be used for a huge range of tasks together with prediction, anomaly detection, diagnostics, automatic insight, reasoning, time series prediction and decision making under uncertainty [29, 31, 33].

### 4.2. Decision tree algorithm J48

J48 classifier is an easy C4.5 selection tree for classification. It creates a binary tree. The selection tree strategy is most beneficial in the classification problem [34]. Decision Trees embody a supervised classification approach. A decision tree is a simple structure the place non-terminal nodes signify checks on one or extra attributes and terminal nodes replicate selection outcomes. The concept got here from the normal tree shape which is made up of a root and nodes (the positions where places branches divide), branches and leaves. In the same way, the choice tree consists of nodes which stand for circles, the branches stand for segments connecting the nodes. A Decision Tree begins from the root, strikes downward and normally are drawn from left to right, so it is less difficult to draw it. The node from where the tree begins is referred to as a root node. The node then place the chain ends is recognized as the "leaf" node. From every interior node (i.e. no longer a leaf) may additionally grow out two or more branches i.e. a node that is now not a leaf node. A node represents a sure attribute while the branches signify a range of values. These ranges of values act as partition points for the set of values of the given characteristic. Figure 2 describes the structure of a tree [29, 30]. J48 is an extension of ID3. The additional features of J48 are accounting for missing values, decision trees pruning, continuous attribute value ranges, derivation of rules, etc. [35, 36].

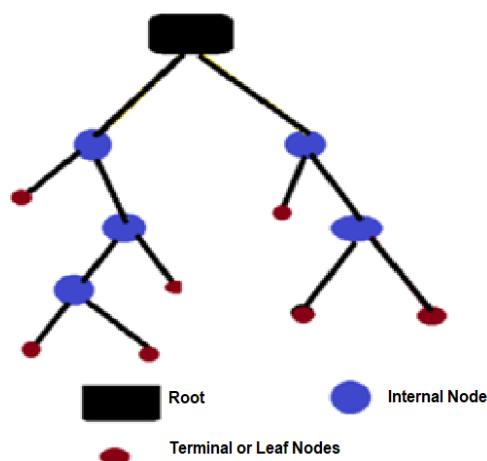


Figure 2. Structure of a tree [34]

### 4.3. Naïve bayes

The Naive Bayes classifier is a simple probabilistic classifier based on applying Bayes' Theorem with strong independence assumptions, which assumes all of the features are equally independent. It uses a Bayesian algorithm for the total probability procedure, the principle is according to the probability that the text belongs to a category of prior probability, and the text would be assigned to the category of posterior probability. In simple terms, a naive Bayes classifier assumes that the presence (or absence) of a particular feature of a class is unrelated to the presence (or absence) of any other feature [37].

### 4.4. Part

PART is a separate-and-conquer rule learner. The algorithm producing sets of rules called „decision lists“ which are planned set of rules. A new data is compared to each rule in the list in turn, and the item is assigned the class of the first matching rule. PART builds a partial C4.5 decision tree in each iteration and makes the “best” leaf into a rule [37].

### 4.5. Random trees

Random trees have added via Leo Breiman and Adele Cutler [38]. The random tree is a tree drawn at random from a set of feasible trees. In this context “at random” ability that each tree in the set of timber has an equal danger of being sampled. Another way of announcing this is that the distribution of trees is “uniform”. Random timber can be generated correctly and the combination of large units of random timber typically leads to accurate models. The algorithm can deal with each classification and regression issues [39, 40].

### 4.6. SMO algorithm

The Sequential Minimal Optimization (SMO) algorithm was proposed by John C. Platt in 1998 and became the fastest quadratic programming optimization algorithm, especially for linear SVM and sparse data performance [41]. SMO algorithm is derived by taking the idea of the decomposition method to its extreme and optimizing a minimal subset of just two points at each iteration. The power of this technique resides in the fact that the optimization problem for two data points admits an analytical solution, eliminating the need to use an iterative quadratic programming optimizer as part of the algorithm [42, 35].

## 5. RESULTS AND DISCUSSION

The dataset that mentioned bove have examined with six classifiers, Sequential Minimal Optimization (SMO), Naïve Bayes, DecisionTree (J48), Random Tree, Rule (PART), and Bayes Ne, respectively. The results indicate that by using PART classifier the system gives the best performance with higher accuracy as shown in Table 1. The rules can be generated to represent a base to classify as a negative and positive classifier, as well as extract the rules from the non-pruned tree. Instead, order the rules, subset rules are ordered (class ordering) and compute the description length of each subset, where the classes that have small length given a high priority. PART is easy to generated rule, therefore, the system generates 430 rules from the mentioned dataset which gives power to the system to classify new instances rapidly.

The performances of DT and the random tree were quite good, DT has based on information theory, therefore, with this huge number of attributes and different classes with various instance, it is more complex to calculate the gain perfectly in order to split the instances. Comparison with the other classifiers, the SMO had good performance, where, SMO is resilient to the overfitting dataset because it depends on only a small number of points in the dataset. Whereas the system gives bad performance when using naïve Bayes and Bayes net classifiers, where the Naïve Bayes classifier is not updatable because the estimator value chosen based on analyzing training set. Naïve net classifier using different types of search algorithms and quality measures, so, with such a huge dataset the system gives weakness performance.

Mean Absolute Error (MAE) are calculated for each classifier as shown in Table 1 to measure how close prediction is to eventual outcomes, which represent the average of the absolute errors, Relative Absolute Error (RAE) is calculated as well in order to measure the uncertainty compare to actual values and clarify the relative in actual value how much space dose error take up, which is represented as percentage. These measurements indicate that the SMO classifier has the highest RAE that the rest classifiers, which explain how this classifier chooses only the small samples of the dataset to create support vectors. The RAE ratio is also high because it handles the attributes of the dataset independently.

Table 1. Measurement of the performance of the classifiers

Classifiers	Accuracy (%)	Mean Absolute Error	Relative Absolute Error (%)
Sequential Minimal Optimization (SMO)	72.39	0.03	98.01
Naïve Bayes	51.065	0.02	50.61
DecisionTree (J48)	75.27	0.01	27.73
Random Tree	72.77	0.01	27.80
Rule (PART)	76.75	0.01	25.20
Bayes Net	68.15	0.03	34.20

### 5.1. Discretized filter

For transforming numerical values of all attributes into categorical counterpart a discretize filter has been used, where, the values have usually discretized in modeling method based n frequencies tables, which may improve the performance of the system and accuracy of the prediction model. This filter is a tool to reduce non-linearity and noise as well, therefore it considered to identify the outliers and missing value of attributes. The performance of the system is increased through using the discretize filter, where, the values of the attributes converted from continues to discrete values, this is clearly visible and demonstrated by the accuracy ratio mentioned in Table 2. Therefore, this filter can be considered as a pre-processing stage for the raw data, this data passed to the classifiers to be tested, the results showed this filter is suitable to such data and capable to increase the performance of the system that using SMO, Naive Bayes, and Naive net classifiers. on the contrary, this filter effect negatively on the system that using decision tree and random tree classifier, where it is quite difficult to those classifiers to distinguish among classes, because, all values discretized simultaneously, and the cases of different classes grouped into the same instance, where it usually there would be mixture of data from several classes in each interval.

From the Table 2, the results indicate that the SMO classifier achieved highest recognition accuracy rate, this is due to the mechanism of this filter, where it maximizes the interdependence between the variables value and the class labels, minimize information loss, and reduce the number of values as continuous variable assumes by grouping them into a number of intervals or bins, which corresponds to the nature of SMO classifier for dealing with data which minimize the optimization of all data.

Table 2 Measurement of the performance of the classifiers using discretize filter

Classifiers	Accuracy (%)	Mean Absolute Error	Relative Absolute Error (%)
Sequential Minimal Optimization (SMO)	91.43	0.03	98.2
Naïve Bayes	68.64	0.01	33.42
DecisionTree (J48)	69.72	0.01	34.12
Random Tree	63.86	0.01	37.25
Rule (PART)	76.75	0.01	25.20
Bayes Net	69.53	0.01	32.39

### 5.2. Resample filter

The dataset is consisting of 39 classes, some of these classes have very fewer classes, and accordingly, they are considered as an unbalanced dataset. Therefore, using the Resample filter may increase the number of instances of the classes has few instances, the produced dataset is strongly biased in terms of class for which only a few samples are available. Table 3 shows that apply this filter on the dataset leads to

improve the performance of the classifiers where this filter gives them the power to recognize the classes efficiently by generating instances for minority classes that make the classes almost balanced. The Decision tree, Random tree [43, 44], and PART classifiers are results of more accuracy by using this filter, where, increasing the numbers of instances give the classifiers wide area to calculating the gain for building the tree. On the other side, this filter adversely affects the performance of a system that used Naïve Bayes and Bayes net classifiers because these classifiers treat the data independently. Furthermore, utilizing this filter helps to decrease the MAE and RAE for all classifiers.

Table 3. Measurement all classifiers performance using resample filter

Classifiers	Accuracy (%)	Mean Absolute Error	Relative Absolute Error (%)
(SMO)	73.21	0.03	98.01
Naïve Bayes	51.82	0.02	49.97
DecisionTree(J48)	84.14	0.001	17.62
Random Tree	85.55	0.009	14.74
Rule (PART)	85.73	0.006	15.88
Bayes Net	70.55	0.01	31.04

### 5.3. Spread sub-sample

From Table 4 it can observe that the accuracies were decreased. Where, through using the spread sub-sample filter, randomly undersampling the majority of classes, therefore, the number of instances of one class is becoming equal to the number of instances of another. The data we use are unbalanced instances of classes; therefore, employing this filter going to undersampling the majority classes, and reduce the oversampling classes, so, number of instances be the same as the fewer instance classes causing decreasing the performance of the system due to losing a lot of instances from different classes to be balanced. Consequently, balancing a dataset has many classes with a big gap between the numbers of instances affect negatively on the efficiency of the system performance. Obviously, using under-sampling on the dataset include many classes, needed to reduce the values of MAE and RAE as illustrated in Table 4.

Table 4. Measurement of the performance of the classifiers using Spread sub-sample filter

Classifiers	Accuracy (%)	Mean Absolute Error	Relative Absolute Error (%)
(SMO)	66.36	0.03	0.13
Naïve Bayes	51.38	0.02	0.13
DecisionTree(J48)	68.85	0.01	0.10
Random Tree	66.29	0.01	0.11
Rule (PART)	70.06	0.01	0.10
Bayes Net	63.47	0.01	0.11

## 6. CONCLUSION

Through the construction of an integrated system without using descriptors, to distinguish humans based on a set of points obtained by using a Kinect sensor, the system can identify them through their gait despite the difference of speeds, directions, ages, and genders. The results indicate the system that using the discretized filter with SMO classifier gives excellent performance more than the rest, where gives 91.3% as recognition rate for all the given data.

## REFERENCES

- [1] Ahmed, M.H., "Kinect-Based Human Gait Recognition Using Static and Dynamic Features," *International Journal of Computer Science and Information Security*, vol. 14, no. 12, pp. 425, 2016.
- [2] Chowdhury, S. M. A., Patwary, T. W. U., and Rahman, R., "Human recognition using wireless router signal," Doctoral dissertation, Brac University, 2019.
- [3] Wang, K., Liu, L., Lee, Y., Ding, X., and Lin, J., "Nonstandard Periodic Gait Energy Image for Gait Recognition and Data Augmentation," In *Chinese Conference on Pattern Recognition and Computer Vision (PRCV)*, Springer, Cham, pp. 197-208, Nov. 2019.
- [4] Choi, S., Kim, J., Kim, W., and Kim, C., "Skeleton-based Gait Recognition via Robust Frame-level Matching," *IEEE Transactions on Information Forensics and Security*, vol. 14, no. 10, pp. 2577-2592, 2019.
- [5] Wang, Jin, She, Mary, Nahavandi, Saeid and Kouzani, Abbas, "A review of vision-based gait recognition methods for human identification," in *DICTA 2010: Proceedings of the Digital Image Computing: Techniques and Application*, IEEE, Piscataway, N.J., pp. 320-327, 2010.

- [6] Jameel, S.K. and Manza, R.R., "Color image segmentation using wavelet," *Int. J. Appl. Inform. Syst. (IJ AIS)*, vol. 1, no. 6, pp.1-4, 2012.
- [7] Jameel, S., "Face recognition system using PCA and DCT in HMM," *Int. J. Adv. Res. Comput. Commun. Eng*, vol. 4, no. 1, pp.13-18, 2015.
- [8] Bashir K., Xiang T. and Gong S., "Gait recognition without subject cooperation," *Pattern Recognition Letters*, vol. 31, no. 13, pp. 2052-2060, 2010.
- [9] Wang, L., Tan, T., Ning, H. and Hu, W., "Silhouette analysis-based gait recognition for human identification," *IEEE transactions on pattern analysis and machine intelligence*, vol. 25, no. 12, pp. 1505-15[023], 2003.
- [10] Kozlowski, L.T. and J.E., "Cutting, Recognizing the sex of a walker from a dynamic point-light display," *Perception & Psychophysics*, vol. 21, no. 6, pp. 575-580, 1977.
- [11] Cutting, J.E. and L.T. Kozlowski, "Recognizing friends by their walk: Gait perception without familiarity cues," *Bulletin of the psychonomic society*, vol. 9, no. 5, pp. 353-356, 1977.
- [12] J. Perry and Others, "Gait analysis: normal and pathological function," *Journal of Pediatric Orthopaedics*, vol. 12, no. 6, pp. 815, 1992.
- [13] M. P. Murray, "Gait as a total pattern of movement," *American Journal of Physical Medicine*, vol. 46, no. 1, pp. 290, 1967.
- [14] D. A. Winter, "Biomechanics and motor control of human gait," *Normal Elderly and Pathological*, pp. 143, 1991.
- [15] Z. Liu and S. Sarkar, "Effect of silhouette quality on hard problems in Gait recognition," *IEEE transactionson systems, man, and cybernetics, Part B, Cybernetics: a publication of the IEEE Systems, Man, and Cybernetics Society*, vol. 35, no. 2, pp. 170-183, Apr. 2005.
- [16] M. Nixon, "New Advances in Automatic Gait Recognition," *Information Security Technical Report*, vol. 7, no. 4, pp. 23-35, Dec. 2002.
- [17] J. Cutting, D. Proffitt, and L. Kozlowski, "A biomechanical invariant for gait perception," *Journal of Experimental Psychology: Human Perception and Performance*, vol. 4, no. 3, pp. 357-372, 1978.
- [18] J. Cutting and L. Kozlowski, "Recognizing friends by their walk: gait perception without familiarity cues," *Bull. Psychonomic Soc.*, vol. 9, no. 5, pp. 353-356, 1977.
- [19] L. Kozlowski and J. Cutting, "Recognizing the sex of a walker from a dynamic point-light display," *Perception and Psychophysics*, vol. 21, no. 6, pp. 575-580, 1977.
- [20] A. F. Bobick and A. Y. Johnson, "Gait recognition using static, activity-specific parameters," in *Proceedings of the 2001 IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, vol. 1, pp. I-423-I-430, 2001.
- [21] N. Goddard, "*The Perception of Articulated Motion: Recognizing Moving Light Displays*," PhD thesis, University of Rochester, 1992.
- [22] C. Bregler and J. Malik, "Tracking people with twists and exponential maps," In *Proc. of IEEE CVPR*, pp. 8-15, 1998.
- [23] D. DiFranco, T. Cham, and J. Rehg, "Recovery of 3D articulated motion from 2d correspondences," *Technical report, Compaq Cambridge Research Lab*, pp. 71-73, 1999.
- [24] H. Sidenbladh, M. Black, and D. Fleet, "Stochastic tracking of 3d human figures using 2D image motion," In *European Conference on Computer Vision (ECCV)*, vol. 2, pp. 702-718, 2000.
- [25] R. Cutler and L. Davis, "Robust real-time periodic motion detection, analysis, and applications," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 781-796, 2000.
- [26] R. Polana and R. Nelson, "Detecting activities," In *CVPR*, 1993.
- [27] J. Little and J. Boyd, "Recognizing people by their gait: The shape of motion," *Videre*, vol. 1, no. 2, 1998.
- [28] M. Nixon, J. Carter, J. Nash, P. Huang, D. Cunado, and S. Stevenage, "Automatic gait recognition," In *Motion Analysis and Tracking (Ref. No. 1999/103)*, *IEE Colloquium on*, pp. 3/1-3/6, 1999.
- [29] J. Shutler, M. Nixon, and C. Harris, "Statistical gait recognition via velocity moments," In *Visual Biometrics, IEEE*, pp. 10/1-10/5, 2000.
- [30] J. Little and J. Boyd, "Recognizing people by their gait: The shape of motion," *Videre: Journal of Computer Vision Research*, vol. 1, no. 2, 1998.
- [31] Cayci, A., Eibe, S., Menasalvas, E. and Saygin, Y., "Bayesian networks to predict data mining algorithm behavior in ubiquitous computing environments," In *Analysis of Social Media and Ubiquitous Data*, Springer, Berlin, Heidelberg, pp. 119-141, 2010.
- [32] Sharma, H., Aakash, A., and Aggarwal, A. G., "A Supervised Guest Satisfaction Classification with Review Text and Ratings," In *Data Science*, CRC Press, pp. 267-284, 2019.
- [33] Muralidharan, V. and Sugumaran, V., "A comparative study of Naïve Bayes classifier and Bayes net classifier for fault diagnosis of monoblock centrifugal pump using wavelet analysis," *Applied Soft Computing*, vol. 12, no. 8, pp. 2023-2029, 2012.
- [34] Majidpour, J. and Jameel, S.K., "Automatic image annotation base on Naïve Bayes and Decision Tree classifiers using MPEG-7," In *2019 5th Conference on Knowledge Based Engineering and Innovation (KBEI)*, IEEE, pp. 007-012, 2019.
- [35] Patil, T.R. and Sherekar, S.S., "Performance analysis of Naive Bayes and J48 classification algorithm for data classification," *International journal of computer science and applications*, vol. 6, no. 2, pp. 256-261, 2013.
- [36] Kaleem, A., and Prabha, I. S., "Boosting of Urban Sound Classification using Preprocessing Techniques," *Journal of Advanced Research in Dynamic and Control Systems*, vol. 11(05-Special Issue), pp. 2293-2304, 2019.

- [37] Parsania, V., Bhalodiya, N. and Jani, N.N., "Applying Naïve bayes, BayesNet, PART, JRip and OneR Algorithms on Hypothyroid Database for Comparative Analysis, *International Journal of Darshan Institute on Engineering Research & Emerging Technologies*, vol. 3, no. 1, 2014.
- [38] [Online]. Available: <http://www.stat.berkeley.edu/users/breiman/RandomForests/>
- [39] Wendl, C., "Novelty Detection in Convolutional Neural Networks Using Density Forests (No. STUDENT)," SIE Master Project, *EPFL*, 2018.
- [40] Zhao, Y. and Zhang, Y., "Comparison of decision tree methods for finding active objects," *Advances in Space Research*, vol. 41, no. 12, pp. 1955-1959, 2008.
- [41] Platt, J., "Sequential minimal optimization: A fast algorithm for training support vector machines," 1998.
- [42] Fatama, A. Z., Haque, A., and Khan, M. A. "A Multi Feature Based Islanding Classification Technique for Distributed Generation Systems," In *2019 International Conference on Machine Learning, Big Data, Cloud and Parallel Computing (COMITCon)*, IEEE, pp. 160-166, Feb. 2019.
- [43] Toppireddy, H. K. R., Saini, B., and Hada, P. S., "Academic Enhancement System using EDM Approach," In *2019 Second International Conference on Advanced Computational and Communication Paradigms (ICACCP)*, IEEE, pp. 1-10, Feb. 2019.
- [44] Ho, T.K., "Random decision forests," In *Proceedings of 3rd international conference on document analysis and recognition*, IEEE, vol. 1, pp. 278-282, Aug. 1995.