Clinical Human Gait Classification: Extreme Learning Machine Approach

[1] Prithvi Patil, [2]K.Shusheel Kumar,[3]Neha Gaud and [4]Vijay Bhaskar Semwal*
[1]National Institute of Technology, Srinagar
[2]B.I.E.T. Jhansi
[3]Vikram University Ujjain
[4]Maulana Azad National Institute of Technology, Bhopal
prithvipatil1357@gmail.com,sus.iiita.932@gmail.com, gaud28neha@gmail.com,vsemwal@gmail.com*

Abstract: This study reports a novel approach for biometric gait pattern classification using Extreme Learning Machine (ELM) algorithm. Clinical gait analysis can be used for early detection of gait abnormality in brain or neurological disorder subjects. In many cases gait abnormality cannot be detected through visual observation alone, but becomes apparent only in a quantitative analysis of subject's gait. This can also help us understand the neuro-muscular mechanics associated with brain disorders. Human gait is also of profound interest to the research community in the field of biometric identification and bipedal robot locomotion due to its uniqueness and efficiency. This paper explores multi-class gait classification using four machine learning methods (KNN, SVM, ELM, MLP) and evaluates their performance for multi class gait classification. The proposed method achieves very good results. TheELM is used first time in to analyses the neuromuscular of patients suffering from multiple sclerosis and stroke

Keywords: Extreme Learning Machine (ELM), Biometric identification, Gait classification, Bipedal walk, Multiple sclerosis, Stroke gait, Clinical gait analysis, Cerebral palsy, Crouch gait, Anthropomorphic robot.

I. INTRODUCTION

Human gait is a complex locomotion for forward propulsion of center of gravity of the human body. It is achieved through a process of continuous learning as a child interacts with his environment [1]. It is highly nonlinear and complex due to varying configuration during different sub phases of human gait [2]. The human walk is designed as 8 discrete sub phases [3] with continuous behavior and researchers have started considering it as hybrid system [4]. Fig.1 shows the 8 different phases of bipedal walk. It is inherently unstable and nonlinear due to high degrees of non-linearity, high dimensionality, under actuation (in swing phase) and over actuation (in stance phase) [5]. Human gait is also unobtrusive compared to other biometric identifiers such as fingerprint [6]. Many Researchers have used human gait for purposes such as biometric identification, understanding the problem of elderly people, to understand the biomechanics of human walk, artificial limb development and bipedal walk [7] [8].

As on today, due to its inherent complexities, bipedal robots are not efficient to work outside laboratory environment i.e. unstructured environments designed for humans and are controlled as fully-actuated system, which is not energy efficient [9]. Among various types of humanoid locomotion, bipedal walking is the most natural, energy efficient and interesting, since bipedal robots have more potential to move in rugged terrains or complex environments, where wheeled or tracked robots cannot operate. On the other hand, bipedal robots are less stable and prone to falling down. To meet this challenge, over the years, many solutions have been proposed [10].

Gait can be studied as non-linear time series of kinematic trajectories [11],[12]. The analysis involves the investigating sensory motor interaction and understanding the biomechanics of locomotors system [13]. Gait classification allocates walking patterns into groups that can be identified and differentiated from one another based on a set of defined variables or features. We believe the problem being addressed so far using conventional mechanics based model and automated control theory can effectively be addressed using data driven computational theory. Throughout this work we tried to validate this hypothesis using machine learning techniques including ELM for patients suffering from multiple sclerosis and stroke. The ELM is used first time in to analyses the neuromuscular of patients suffering from multiple sclerosis and stroke

The paper is organized as follows. Section II is about methodology. Section III introduces ELM and briefly describes other machine learning approaches used in this study. Section IV discusses performance evaluation and results achieved by different machine learning algorithms. Section V is about Conclusion and Future work.

II. METHODOLOGY

A. The Proposed System:

Proposed system consists of eight phases which is shown in the Fig.1. They are gait data collection, gait data detection, trajectories smoothing, feature extraction, feature selection and classification using ML algorithms [14]. Figure 2 depicts a systematic approach for proposed gait recognition process.

978-1-7281-3445-1/19/$31.00 ©2019 IEEE

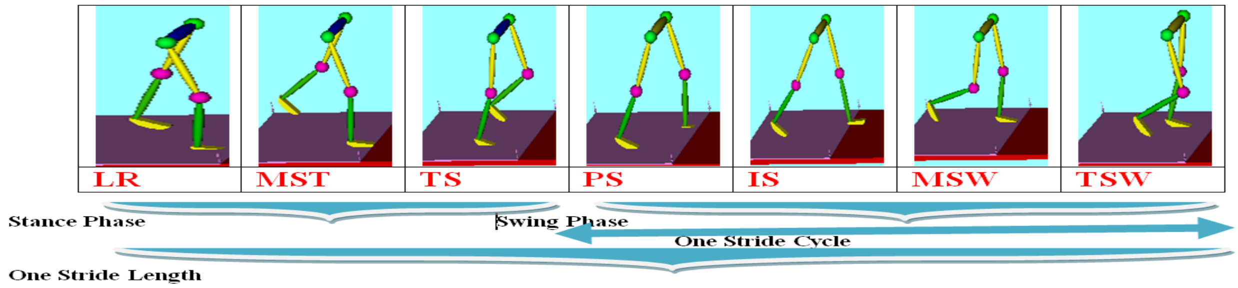

Fig.1. Breakdown of Human Gait into different Discrete Sub Phases

Fig.2. Proposed Architecture of the System

The Gait is divided into 2 phases, namely stance phase and swing phase. The swing phase is critical since only one leg needs to carry the entire body load. We consider it as discrete phase and the configuration during this phase will be over actuation. The swing phase can be considered as continuous phase as well. The configuration is under actuation during this phase.The Percentage-wise distribution of gait time series as following:

Stance phase:

1. Initial Contact– IC[0-2%]
2. Loading Response– LR[2-10%]
3. Mid Stance– MS[10-30%]
4. Terminal Stance– TS[30-50%]

Swing phase:

5. Pre Swing– PS[50-60%]
6. Initial Swing– IS[60-73%]
7. Mid Swing - MS[73-87%]
8. Terminal Swing – TS[87-100%]

B. Data Collection:

For class Multiple sclerosis and Stroke gait, data was collected and recorded using accelerometer for capturing acceleration in all the three planes for the 6 joints (left hip, right hip, left knee, right knee, left ankle, right ankle).This was later converted into joint angles using inverse kinematic solution [17].

C. Data Smoothing:

To remove unwanted noise from the data collected through accelerometers, we have cubic spline interpolation.

D. Dimension Reduction:

After data smoothing, we employed PCA (Principle Component Analysis) and IFA(Incremental Feature Analysis) to compute covariance between different features incrementally.

E. Feature selection:

After analyzing covariance between different set of features, we then proceed to select set of 6 features out of 10, which have highest covariance. This not only reduces the chance of over fitting, but also leads to improved generalization and classification accuracy as discussed in [15],[16].Table-1 below shows all the 10 kinematic features. F1-F6 represents the selected features.

Table1: List of 10 kinematic features

Feature Category	Feature Name
F1	Left Ankle
F2	Right Ankle
F3	Left Knee
F4	Right Knee
F5	Left Hip
F6	Right Hip
F7	CoP*
F8	CoG*
F9	GRF*
F10	Velocity

III. Algorithms:

A. Extreme Learning Machines(ELM)

ELM was proposed by Huang [18], [19]. It is different from MLP and SVM. MLP used to consider many hidden layer neuron and considered whole networks as a black box. ELM theories show that hidden neurons do not need tuning because its hidden nodes parameters (c_i ,a_i) are randomly assigned [20]. For N arbitrary distinct samples $(x_k, t_k) \in R^n \times R^m$, the single ELM classifier with N hidden nodes becomes a linear system as,

$$\sum_{i=1}^{N} \beta_i G(x_k;\ c_i;\ a_i) = t_k, k = 1, \dots \dots . N. \quad (1)$$

where c_i € Rn and a_i € R are the learning parameters of hidden nodes and randomly assigned weight β_i connecting the ith hidden node to the output node, xk are the training examples, t_k is the target output for k = 1, ..., N, and G (x_k ; c_i ; a_i) is the output of the i[th] hidden node with respect to the input xk. The output weights can be described in matrix form as

$$\beta = \left[\beta_1^T \dots . \beta_N^T\right]_{mxN}^T \qquad (2)$$

Equation (1) can be rewritten as:

$$H\beta = T \qquad (3)$$

Where H (c_1.......c_N, a1 ......, x_1, ..., x_N) =

$$\begin{bmatrix} G(x_1; c_{1,}a_1) & \dots & G(x_1; c_{N,}a_N) \\ \vdots & \ddots & \vdots \\ G(x_N; c_{1,}a_1) & \dots & G(x_N; c_{N,}a_N) \end{bmatrix}_{NxN} \qquad (4)$$

$$T = \left[t_1^T \dots t_N^T\right]_{mxN}^T (5)$$

The weights β can be obtained using equation 6 by taking the least-square solution:

$$\widehat{\beta} = H^{\dagger}T \qquad (6)$$

Here, H† is the Moore-Penrose generalized inverse [21] and H is the output matrix of the hidden layer. The ELM used to implement the multi-class classification using a network architecture which has output nodes equal to the number of pattern in the classes. The network output can be written as O = $(O_1; O_2; \dots O_n)^T$. For each training example say x, the target output tg is coded into n bits: $(tg_1; \text{---}\ tg_n)^T$ .For a pattern of class k, only the target output t_k is “1” and the rest is “-1”

B. MLP, SVM and KNN Classifiers:

Multi-layer perceptron (MLP) is the simplest neural network architecture which is used most frequently. It has configuration flexibility, good representational capabilities and programmable algorithms. The Complexity of neural network depends of no. of hidden layer and no. of neurons. The best model of neural network is derived by trial and error. The major challenge of neural network is to decide the no. of hidden layer and no. of neurons, as there is no as such mechanism of optimal selection of neurons and hidden layers. Most MLP are based on gradient descent techniques that require iterative tuning which makes MLP extremely slow and the algorithm can also converge to local minima and provide sub optimal results.

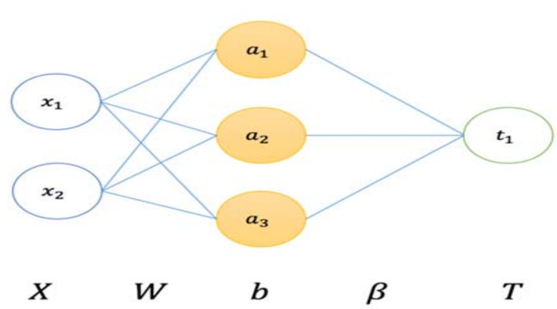

Fig.3. SLFNarchitecture of ELM

The objective of next classifiers SVM is to find the optimal decision boundary with maximum margin between classes in higher dimension feature space. It used to map the lower dimension data into higher dimension feature space using kernel function. The kernel function used to take linear time for this transformation. It converts the nonlinearly separable data into linearly separable data in high dimension feature space. Using SVM, an optimal separating hyper plane in the higher dimensional feature space can be computed by using kernel functions in the input space. In this study, we have used a Support Vector Machine classifier based on LibSVM. It uses a RBF(Radial Basis Function) kernel.

KNN used to match the similarity measure to classify the new class. It is simple algorithm. It is a simple non-parametric technique for classification. When the new testing data is fed, KNN computes the distance between the query data and training samples. Based on the defined threshold, number-k, k samples with least distances are selected and the class with more samples in the boundary is the result.

IV. Results and Discussions

Our experiments were conducted on a Windows 10 machine with a 2.50-GHz 7200U i5 CPU and 8.0-GB RAM. To find the best parameters for each of the classifiers, a grid search was run exhaustively with 10-fold cross validation to reduce variation associated with randomness of data shuffling and arrive at mean scores. KNN, SVM and MLP were implemented using scikit-learn python module. The best set of parameters was found to be as under. Table 2 represents the best set of parameters for MLP.

Table2: Set of Parameter for MLP

Hidden Layer Neurons	50
Activation function	‘relu’
solver	‘Lbfgs’
alpha	0.0001
momentum	0.9
iterations	1000

‘relu’ is an activation function such that,f(x)=max(0,x).‘lbfgs’ is an optimizer in the family of quasi-Newton methods.‘alpha’ is the L2 penalty parameter.‘multiquadric’ is an Radial basis activation function suchthat,f(x)=√1+(∈x)2. table 3,4 and 5 represents the best set of parameters for KNN, SVM, ELM.

Table 3:Set of Parameter for KNN

Neighbors	4
weights	‘uniform’
metric	‘manhattan’

Table 4:Set of Parameter for SVM

Parameter C	32768
kernel	‘rbf’
gamma	0.001953125

Table 5: Set of Parameter for ELM

Hidden neurons	370
Activation function	‘multiquadric’
Rbf-width	0.4

To evaluate results we computed confusion matrix, precision, recall, f1 score for each of the four classifiers and calculated overall classification accuracy. Overall accuracy is simply number of correct predictions divided by total number of predictions. F1 score is the harmonic mean of precision and recall. Additionally, we also recorded time taken for each classifier to train and predict labels on testing set. All performance measures were generated using scikit-learn python module. In order to eliminate high variation from results we have employed a 10-fold cross validation.

The data set consists of 5 classes namely,(Normal, Crouch1,Crouch2) obtained from OpenSim [22], Multiple sclerosis and Stroke. Data for classes Multiple sclerosis and Stroke were collected clinically capturing the same kinematic features as in the first three classes. Subjects didn't suffer from any other medical condition besides the said diseases. There are, a total of 945 samples across all 5 classes with their individual numbers given under 'support' column in the classification report tables. Crouch gait is a common movement abnormality among children with cerebral palsy, decreases walking efficiency due to the increased knee and hip flexion during the stance phase of gait [23][24].Crouch gait is of four types, out of which we have included type 1 and type 2 for classification. Table 6 and 7 represents the confusion matrix and classification report for ELM.

Table 6: ELM Confusion Matrix

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	202	0	0	0	0
Crouch1	0	54	19	0	0
Crouch2	0	25	45	0	0
MS	0	0	1	299	0
Stroke	0	9	7	0	284

Table 7: Classification Report using ELM

	Precision	Recall	F1-score	Support
Normal	1.00	1.00	1.00	202
Crouch1	0.61	0.74	0.67	73
Crouch2	0.62	0.64	0.63	70
MS	1.00	1.00	1.00	300
Stroke	1.00	0.95	0.97	300

The Overall classification accuracy using ELM is 93.54%. Table 8 and 9 represents the confusion matrix and classification report for MLP.

Table 8: MLP Confusion matrix

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	197	2	3	0	0
Crouch1	0	26	47	0	0
Crouch2	2	42	26	0	0
MS	0	0	0	300	0
Stroke	0	0	0	0	300

Table 9: Classification Report using MLP

	Precision	Recall	F1-score	Support
Normal	0.99	0.98	0.98	202
Crouch1	0.37	0.36	0.36	73
Crouch2	0.34	0.37	0.36	70
MS	1.00	1.00	1.00	300
Stroke	1.00	1.00	1.00	300

The Overall classification accuracy is 89.84% achieved using MLP. Table 10 and 11 represents the confusion matrix and classification report for KNN..

Table 10: KNN Confusion matrix

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	191	5	6	0	0
Crouch1	20	20	33	0	0
Crouch2	19	50	1	0	0
MS	0	0	0	300	0
Stroke	0	0	0	0	300

Table 11: Classification Report using KNN

	Precision	Recall	F1-score	Support
Normal	0.83	0.95	0.88	202
Crouch1	0.27	0.27	0.27	73
Crouch2	0.03	0.01	0.02	70
MS	1.00	1.00	1.00	300
Stroke	1.00	1.00	1.00	300

The Overall classification accuracy=85.92% using KNN. Table 12 and 13 represents the confusion matrix and classification report for SVM.

Table 12: SVM Confusion matrix:

	Normal	Crouch1	Crouch2	MS	Stroke
Normal	202	0	0	0	0
Crouch1	0	53	20	0	0
Crouch2	0	15	55	0	0
MS	0	0	0	300	0
Stroke	0	0	0	0	300

Table13: Classification Report using SVM

	Precision	Recall	F1-score	Support
Normal	1.00	1.00	1.00	202
Crouch1	0.78	0.73	0.75	73
Crouch2	0.73	0.79	0.76	70
MS	1.00	1.00	1.00	300
Stroke	1.00	1.00	1.00	300

We have achieved the overall classification accuracy 96.29% using SVM.

Comparative analysis

We have used three figures as shown below to bring out a comparative analysis of the four classifiers. From figure-4 we can see that although numerically the poorest performing classifier-KNN, achieves a good overall accuracy of 85.92%, but it's classification is very poor in the case of Crouch 1 and Crouch2.MLP too, fails to have a decent f1 score in the same two classes, while SVM and ELM have evidently performed well across all 5 classes. The overall accuracy comparison is provided by figure-5.Interestingly, although KNN is the fastest of all four algorithms, it's also the least accurate, while ELM, despite having a far more complex Neural network of 370 neurons against MLP's 50, predicts labels in less than 1/10th the time.

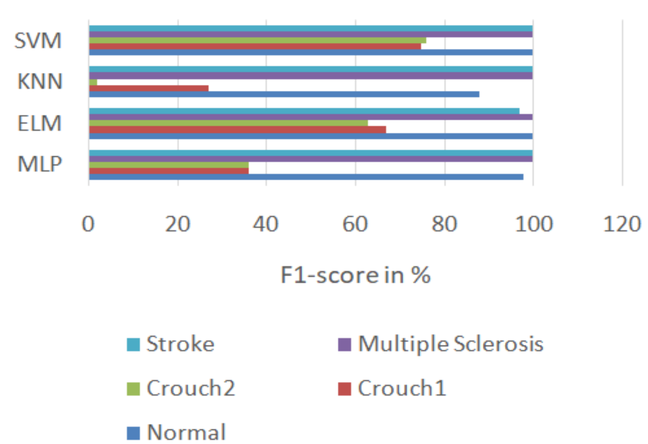

Fig.4. Comparison across classes

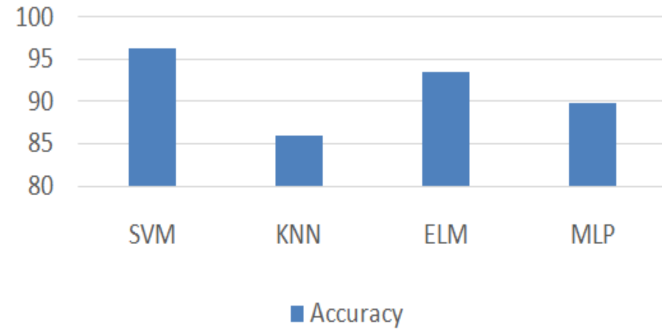

Fig.5. Different Classifier Classification Accuracy

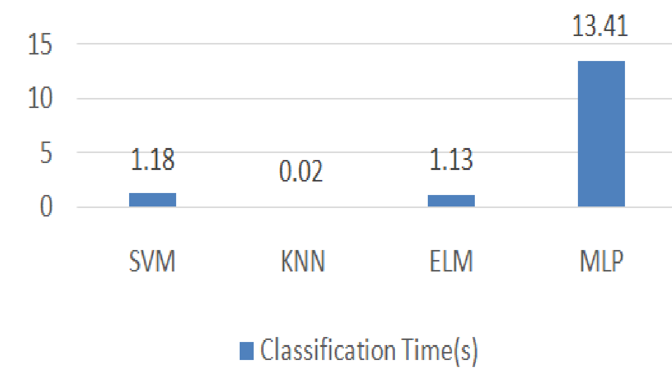

Fig.6. Different Classifier Classification time in seconds

VI. Conclusion

This paper presented the performance of ELM based classifiers with other machine learning classifiers named SVM, MLP and KNN for multi class gait classification. The paper has used kinematic features. The results show that ELM performs very good classification accuracy compare to other classifiers. The ELM has achieved the 93.54% overall classification accuracy. As the performance of machine learning algorithms dependent on the selected feature and parameter used, to improve the performance the paper has implemented increment feature analysis technique for kinematics features selection and various parameters is adjusted. In the future one could experiment with dynamic and kinematic features as this study was limited to kinematic gait characteristics alone.

References:

[1] Semwal, Vijay Bhaskar, et al. "Biped model based on human Gait pattern parameters for sagittal plane movement." IEEE,*International Conference onControl, Automation, Robotics and Embedded Systems (CARE),* 2013.

[2] Kale, A.;Sundaresan, et.al. "Identification of humans using gait." Image Processing, IEEE Transactions on 13, no. 9 (2004): 1163-1173.

[3] Kiichi_Katoh,, Makoto Katoh.: Using FS in Tuning a Biped Locomotion Surrogate Robot for Walking Training. In: 2011 4th International Conference on Biomedical Engineering and Informatics (BMEI).

[4] Nandi, Gora Chand, et al. "Modeling bipedal locomotion trajectories using hybrid automata." IEEE TENCON,2016.

[5] Mitsuru Yoneyama, Yosuke Kurihara, KajiroWatanabe,: Accelerometry-Based Gait Analysis and Its Application to Parkinson's Disease Assessment—Part 2: Ieee Transactions On Neural Systems And Rehabilitation Engineering, Vol. 21, No. 6, November 2013

[6] Venkataraman, S. T. "A simple legged locomotion gait model." Robotics and Autonomous Systems 22.1 (1997): 75-85.

[7] Semwal, Vijay Bhaskarand Gora Chand Nandi. "Biometric gait identification based on a multilayer perceptron." *Robotics and Autonomous Systems* 65 (2015): 65-75.

[8] Watelain, Eric, Franck Barbier, Paul Allard, André Thevenon, and Jean-Claude Angué. "Gait pattern classification of healthy elderly men based on biomechanical data." Archives of physical medicine and rehabilitation 81, no. 5 (2000): 579-586.

[9] Huang, Guang-Bin, Lei Chen, and Chee Kheong Siew. "Universal approximation using incremental constructive feedforward networks with random hidden nodes." IEEE Trans. Neural Networks 17.4 (2006): 879-892.

[10] Khusainov, R., A. Klimchik, and E. Magid. "Comparison of kinematic and dynamic leg trajectory optimization techniques for biped robot locomotion." Journal of Physics: Conference Series. Vol. 803. No. 1. IOP Publishing, 2017.

[11] Semwal, Vijay Bhaskar, et al. "Design of Vector Field for Different Subphases of Gait and Regeneration of Gait Pattern." IEEE Transactions on Automation Science and Engineering (2016).

[12] Khusainov, Ramil, et al. "Toward a human-like locomotion: modelling dynamically stable locomotion of an anthropomorphic robot in simulink environment." Informatics in Control, Automation and Robotics (ICINCO), 2015 12th International Conference on. Vol. 2. IEEE, 2015.

[13] Semwal, Vijay Bhaskar, and Gora Chand Nandi. "Generation of joint trajectories using hybrid automate-based model: a rocking block-based approach." IEEE Sensors Journal 16.14 (2016): 5805-5816.

[14] Huang, Guang-Bin, et al. "Extreme learning machine for regression and multiclass classification." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 42.2 (2012): 513-529.

[15] Semwal, Vijay Bhaskar, et al. "An optimized feature selection technique based on incremental feature analysis for bio-metric gait data classification." Multimedia Tools and Applications (2016): 1-19.

[16] Vijay Bhaskar Semwal and G.C. Nandi, Robust and more accurate feature and classification using deep neural network, Neural Computing and Application, Springer.

[17] Semwal, Vijay Bhaskar, and Gora Chand Nandi. "Toward developing a computational model for bipedal push recovery–a brief." *IEEE Sensors Journal* 15.4 (2015): 2021-2022.

[18] Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme learning machine: a new learning scheme of feedforward neural networks." Neural Networks, 2004. Proceedings. 2004 IEEE International Joint Conference on. Vol. 2. IEEE, 2004.

[19] Huang, Guang-Bin, Qin-Yu Zhu, and Chee-Kheong Siew. "Extreme learning machine: theory and applications." Neurocomputing 70.1 (2006): 489-501

[20] Huang, Guang-Bin. "An insight into extreme learning machines: random neurons, random features and kernels." Cognitive Computation 6.3 (2014): 376-390.

[21] Rao, Calyampudi Radhakrishna, and Sujit Kumar Mitra. "Generalized inverse of matrices and its applications." (1971): 601-620.

[22] Delp, Scott L., et al. "OpenSim: open-source software to create and analyze dynamic simulations of movement." IEEE transactions on biomedical engineering 54.11 (2007): 1940-1950.

[23] Wren, Tishya AL, Susan Rethlefsen, and Robert M. Kay. "Prevalence of specific gait abnormalities in children with cerebral palsy: influence of cerebral palsy subtype, age, and previous surgery." Journal of Pediatric Orthopaedics 25.1 (2005): 79-83.

[24] Semwal, Vijay Bhaskar, Neha Gaud, and G. C. Nandi. "Human gait state prediction using cellular automata and classification using ELM." Machine Intelligence and Signal Analysis. Springer, Singapore, 2019. 135-145.