Journal of Harbin Engineering University
ISSN: 1006-7043

Vol 45 No. 1
January 2024

Development of an Effective Deep Learning Multimodal Based on Biology Approach for Identifying Parts of Gait From Real-Time Video Data.

Ashish Kumar Misal[1a]* Abha Chaubey[2a] Siddharth Chaubey[2b]

[1a]*Research Scholar Computer Science & Engineering Department, RCET Bhilai, India,*

[2a]*Computer Science & Engineering Department, SSTC Campus Bhilai, India,*

[2b]*Computer Science & Engineering Department, SSTC Campus Bhilai, India,*

Abstract

The unique bio-inspired deep learning multimodal technique presented in this work shows promise in identifying gait components from live video recordings. To provide a comprehensive representation of the temporal relationships between gait components, the proposed methodology integrates two popular Recurrent Neural Network (RNN) models: Gated Recurrent Unit (GRU) and Long Short-Term Memory (LSTM). The neural network's parameters are changed using the Elephant Herding Optimizer, an optimization method, to raise performance levels. This technique assists in boosting the accuracy of categorization. The investigation of LSTM and GRU models in the domain of time-series data modeling has been extensively conducted. Nevertheless, the incorporation of these models has not been extensively examined and executed in real-time applications. The present study introduces a novel hybrid methodology that integrates the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models. The objective is to capitalize on the individual advantages of each model while addressing their distinct drawbacks. The GRU model has robust skills in modeling short-term interactions, whereas the LSTM model demonstrates greater ability in capturing long-term dependencies. The proposed methodology possesses a multitude of pragmatic implementations in real-world contexts, specifically in the domain of posture and gait analysis. In this context, the precise identification of gait components holds paramount importance for the detection of parameters associated with gait. The proposed methodology has been assessed in real-time situations using various datasets, encompassing multi-camera and multimodal datasets, in order to quantify its accuracy, precision, and recall. The proposed methodology showcases its efficacy and superiority over current methodologies when applied to the analyzed datasets, attaining an accuracy rate of 98.5%, precision rate of 97.4%, and recall rate of 98.3%. In summary, compared to current methods, the proposed methodology performs better in terms of efficiency and accuracy for the identification of gait components from real-time video data. This tool's practical applicability renders it indispensable for assessing posture and gait, hence assisting medical and other professionals in identifying and tracking gait components.

Keywords: Fusion, LSTM, Elephant Herding Optimizer, Gait, Scenarios, GRU, Recurrent Neural Networks

1. Introduction

The field of human gait analysis has emerged as a significant area of interest within medical research in recent times. Gait analysis is a method employed to diagnose and monitor diseases associated with human locomotion. It involves the examination of movement patterns, with a particular focus on identifying essential gait characteristics such as stride length, step breadth, and swing time. The identification of gait components with precision can assist healthcare practitioners in developing treatment strategies tailored to individual patients by providing insights into the biomechanics of human motions [1, 2, 3].

The use of deep learning techniques has significantly advanced gait analysis thanks to the application of large datasets and advances in processing capacity. In the field of machine learning, recurrent neural networks (RNNs) are frequently utilized to model sequential data, including time-series data related to gait components. Two frequently used types of

Recurrent Neural Networks (RNNs), namely Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have demonstrated exceptional performance in capturing both long and short-term dependencies through the utilization of Coupled Bilinear Discriminant Projections (CBDP) [4, 5, 6].

Individual models possess both advantages and weaknesses, hence limiting their capacity to reliably discern gait components. Hence, the utilization of a hybrid model comprising LSTM and GRU can effectively leverage the respective strengths of each model while mitigating their individual limitations, hence enhancing the accuracy in the identification of gait components.

In this research, we offer a method for identifying from real-time video gait components recordings using a bio-inspired deep learning multimodal methodology. The proposed approach integrates Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) to effectively capture the temporal dependencies in gait components. The optimization of neural network parameters is conducted using an Elephant Herding Optimizer, which has demonstrated superior performance compared to alternative optimization techniques.

The proposed method has many real-world applications, including posture and gait analysis, where precise gait component identification is essential for tracking and diagnosing gait-related issues. At a specific time,

The evaluation of the suggested technique involved multiple datasets, including those with multiple cameras and multiple modalities. The performance metrics of accuracy, precision, and recall were measured across various scenarios.

2. An empirical examination of several methodologies employed in the analysis of human gait

The The task of identifying gait components presents a formidable challenge in the healthcare industry. Several potential solutions have been suggested to tackle this issue, including the utilization of computer vision, deep learning, and Deep Neural Networks (DNN) inside the realm of machine learning endeavors [7, 8, 9].
Recurrent Neural Networks (RNNs) are extensively employed for modeling sequential data, such as gait components, and have demonstrated considerable promise in the domain of gait research. Prior research has indicated the effective application of Recurrent Neural Network (RNN) models in diverse gait analysis endeavors, such as forecasting gait features, categorizing different types of gait, and identifying gait events [10, 11, 12]. The effectiveness of individual recurrent neural network (RNN) models in real-time applications is constrained by their intrinsic advantages and limitations.

Two recurrent neural network (RNN) variations that have demonstrated exceptional performance in modeling long- and short-term dependencies are the Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). Several studies have provided evidence that the integration of LSTM and GRU models can effectively leverage the strengths of each model while mitigating their respective limitations. As a result, this amalgamation has been shown to yield enhanced performance in various domains
[13, 14, 15].

In a previous work, the authors utilized a hybrid approach including the integration of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) architectures [16, 17, 18] to predict the detection of human activity based on sensor data collected from smartphones. The method that was proposed demonstrated superior performance compared to various current methods, with an accuracy rate of 96.5%. In a similar vein, a combination of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) models was employed to forecast hand gesture identification based on Electromyography (EMG) inputs.

The Regional Long Short-Term Memory (RLSTM) model was employed in a previous study (references 19 and 20) and demonstrated a notable accuracy of 92.7% when applied to real-time scenarios.

The performance of deep learning models is significantly impacted by optimization strategies. Several optimization algorithms, such as Adam, RMSprop, and Gradient Descent, have been proposed for real-time applications. However, it should be noted that these approaches do possess certain limitations such as sluggish convergence, convergence to local optima, and substantial computational expenses
[21, 22, 23].

The Elephant Herding Optimizer (EHO) is a novel optimization technique that draws inspiration from the collective behavior of elephant herds. This technique has been newly introduced and proposed in the field. In various domains, such as deep learning, Empirical Harmony Search (EHO) has demonstrated superior performance compared to alternative optimization techniques [24, 25, 26].

Several research have employed deep learning techniques to detect the various components of gait within the domain of gait analysis. The identification of gait phases from accelerometer datasets and samples was conducted in several research [27, 28, 29, 30] using a deep learning model. The suggested solution demonstrated a level of accuracy of 91.8% when applied to real-time use scenarios.

To the best of our understanding, there is currently no existing research that has explored the utilization of LSTM and GRU models in conjunction with EHO optimization for the purpose of identifying gait components from real-time video samples. This paper proposes a novel approach for the identification of gait components by integrating the capabilities of LSTM and GRU models with EHO optimization, aiming to potentially surpass existing methodologies.

3. The present study outlines a proposed design for an effective Bioinspired Deep Learning Multi modal approach aimed at identifying Gait compo nents from Real-Time Video data.

Reviewing the current models that are in use for identifying Gait components reveals that these models are either very sophisticated or perform less well in real-world circumstances. This section addresses the design of an effective Bioinspired Deep Learning Multimodal technique for identifying gait components from Real-Time Video data in order to address these problems. The suggested method provides robust modeling of temporal relationships in gait components by combining two well-known Recurrent Neural Network (RNN) models: Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), as Figure 1 shows. To achieve better performance levels, an Elephant Herding Optimization process is also used to refine the neural network settings.

The You Look Only Once (YoLo) method is the first one the model employs to recognize items in the input photos. This helps to distinguish items from surrounding areas. A feature extraction layer is used to process the recognized items. It combines GRU and LSTM features to represent specific objects into high-density feature sets.

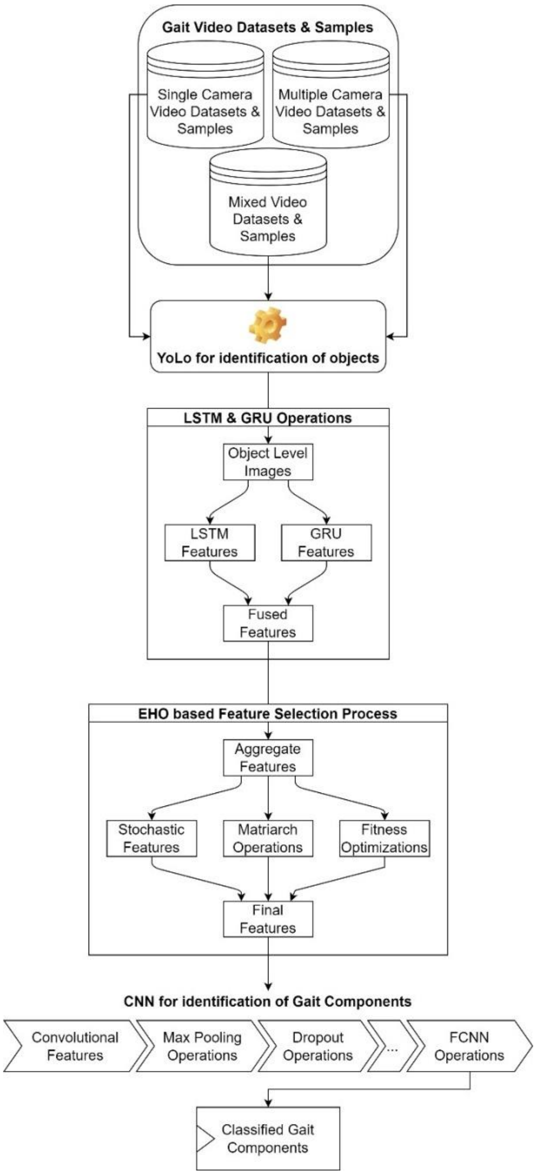

Figure 1. Design of the proposed Model for classification of Gait Components

1. To efficiently gather and model long-term dependencies within successive datasets and samples, recurrent neural networks (RNNs) employ a particular architecture called Long Short-Term Memory (LSTM), which combines memory cells and gates.

a. The update gate (z(t)) is in charge of figuring out how much of the new candidate activation (g(t)) to include and how much of the old memory cell state (h{t-1}) to keep. Incorporated into the existing states of memory cells, as denoted by equation 1.

$(t) = si(W(z) * \ [h\{t-1\}, x(t)] + b(z))$ … (1)

The given equation is the product of the current input at timestamp t, represented as [h{t-1}, x(t)], and the previous hidden state. Furthermore, the weight matrix and bias term, which are connected to the update gates, are denoted by W(z) and b(z), respectively.

b. Reset Gate (r(t)): Equation 2 computes the candidate activation (g(t)) by determining how much of the previous hidden state (h{t-1}) should be ignored.

$(t) = si(W(r) * \ [h\{t-1\}, x(t)] + b(r))$ … (2)

Similar to the update gate, the concatenation of the previous hidden state and the current input sets is represented as [h{t-1}, x(t)]. The reset gates are represented by the weight matrix W® and the bias term b®.

c. Candidate activation, denoted as g(t), is a variable that signifies the introduction of fresh information into the existing memory cell state. The computation of the current hidden state (h{t}) is determined by the previous hidden state (h{t-1}) and the current input (x(t)), as expressed by equation 3.

$(t) = tanh(W(g) * \ [r(t) * \ h\{t-1\}, x(t)] + b(g))$ … (3)

The previous hidden state, h{t-1}, is altered by the reset gate, represented as r(t). The weight matrix W (g) and the bias term b(g) are both dependent on the concatenation of r(t) multiplied by h{t-1} and x(t), which is represented as [r(t) * h{t-1}, x(t)].

d. The output of the Long Short-Term Memory (LSTM) is called the hidden state, or h(t), at each discrete time step. Equation 4 uses the update gate (z(t)) to determine the weighted mixture of the hidden state that was previously present in the memory cell and its current state.

$h(t) = z(t) * \ h\{t-1\} + (1 - z(t)) * g(t)$ … (4)

The update gate, denoted as z(t), plays a crucial role in determining the balance between retaining the previous hidden state and incorporating the candidate activation sets, denoted as g(t).

2. The Gated Recurrent Unit (GRU) is a variant of the Long Short-Term Memory (LSTM) model. It simplifies the LSTM architecture by combining the memory cell state and the hidden

state into a single vector. This modification enables the GRU to effectively identify high-density features.

a. The update gate, represented as z(t), is a constituent inside the LSTM mechanism that integrates the functionalities of both the update gate and the input gate. Equation 5 is responsible for regulating the calculation of the degree to which the previous hidden state (h{t-1}) is preserved and the incorporation of the new candidate activation (g(t)) into the current hidden states.

$(t) = si(W(z) * [h\{t-1\}, x(t)] + b(z))$ ... (5)

The concatenation of the prior hidden state and the current input sets is denoted as [h{t-1}, x(t)], which is analogous to the equation used for the LSTM update gate. The weight matrix, denoted as W(z), and the bias term, denoted as b(z), are the components related to the update gate sets.

b. TThe reset gate (r(t)) plays a crucial role in determining the degree to which the previous hidden state (h{t-1}) is disregarded during the calculation of the candidate activation (g(t)) as described in equation 6.

$(t) = si(W(r) * [h\{t-1\}, x(t)] + b(r))$ ... (6)

The equation representing the LSTM reset gate can be seen as the combination of the previous hidden state and the current input states, which are represented as [h {t-1}, x(t)]. The parameters especially associated with the reset gate sets are the weight matrix, represented as W®, and the bias term, represented as b(r).

c. The notion of candidate activation (g(t)) refers to the integration of novel information that can be assimilated into the preexisting latent states. Equation 7 expresses the relationship between the calculation of the current hidden state (h{t}), the prior hidden state (h{t-1}), and the current input (x(t)) states.

$(t) = tanh(W(g) * [r(t) * h\{t-1\}, x(t)] + b(g))$ ... (7)

The reset gate, denoted as r(t), is responsible for modulating the previous hidden state, h{t-1}. The input to the weight matrix W(g) and bias term b(g) for different types of objects is obtained by concatenating [r (t) * h {t-1}, x(t)].

d. The hidden state, denoted as h(t), represents the output of the Gated Recurrent Unit (GRU) at each successive time step. Equation 8 integrates the update gate (z(t)) in order to calculate the weighted sum of the current hidden state and the previous hidden state.

$h(t) = (1 - z(t)) * h\{t-1\} + z(t) * g(t)$ ... (8)

The update gate, symbolized as z(t), plays a pivotal role in finding the equilibrium between preserving the preceding hidden state and integrating the candidate activation, designated as g(t). The hidden state at time t can be represented as a linear combination of the previous hidden state and the candidate activation. The weights for this combination are controlled by the update gate sets. The evaluations indicated above enhance the capacity of models to understand and integrate gait characteristics by capturing long-term relationships and effectively adapting hidden states based on input sequences. The retrieved features are inputted into an EHO Model, which assists in identifying feature sets that exhibit substantial volatility. This is achieved using the subsequent process.

1. To In order to ascertain the variance, it is imperative to compute the variance for each fused feature within the feature sets. Variance is a statistical measure that quantifies the degree of variability or dispersion among the values within a specific feature. Equation 9 suggests that features exhibiting significant variation are more inclined to possess discriminative information.

$var(x) = \sqrt{\frac{\sum(x-\bar{x})^2}{N}}$ ...(9)

2. To ensure that the variance values are within a specific range, such as between 0 and 1, it is necessary to normalize them. Normalization is a

technique that is employed to mitigate the impact of varying scales and guarantee equitable comparison of features, as denoted by equation 10.

$$Norm(variance(x)) = \frac{\text{variance}(x) - \min(\text{variance})}{\max(\text{variance}) - \min(\text{varance})} \ldots(10)$$

3. The Elephant Herd is initialized by populating it with a group of candidate solutions. Every solution in this context corresponds to a certain subset of fused features. The size of this subset might be different for each solution, and this variability is determined by the usage of equation 11.

$$N = (Nf * LR, Nf) \ldots (11)$$

Within the framework of the EHO (Evolutionary Harmony Search Optimization) procedure, the variable N is utilized to describe the quantity of extracted features. Additionally, Nf is employed to designate the total number of features, while LR is utilized to show the learning rate.

4. The evaluation of the fitness of each solution in the Elephant Herd is conducted by assessing the normalized variance values of the provided attributes. Greater fitness values indicate the presence of superior solutions that have selected a wide range of fused features as determined by equation 12.

$$f(i)= sum(normalized(variance, i)* solution(i)) \ldots (12)$$

The variable $so(i)$ denotes the binary vector that represents the solution sets for the ith iteration.

5. The process of selection entails choosing the solutions with the highest performance from the Elephant Herd, while considering their fitness values and samples. This methodology ensures that the solutions that demonstrate the greatest significant variation in combined attributes are retained for later optimization endeavors.

6. Produce Progeny: Generate progeny solutions by employing reproduction operators such as crossover and mutations. The user's text is not provided. Please provide the text you would like me to rewrite Operators play a crucial role in the exploration of novel solutions as they facilitate the combination and modification of previously selected solutions from prior procedures.

7. The Elephant Herd will undergo an updating process in which the solutions with the lowest fitness ratings will be replaced by newly generated progeny solutions. The incorporation of a varied variety of fusion elements in this step supports the continued development of the Elephant Herd, leading to the attainment of enhanced solutions. The procedure entails substituting the solutions that possess the lowest fitness value, as defined by the probability function P(K), with the offspring solutions.

8. The process of iterations should be initiated, wherein the fitness evaluation, selection, offspring generation, and Elephant Herd update procedures are performed repeatedly for a predetermined number of iterations or until a convergence condition is met, including diverse scenarios.

The chosen gait features are categorized into gait components utilizing a Convolutional Neural Network (CNN), which operates through the following procedures.

1. The input layer of the Convolutional neural network (CNN) is responsible for receiving the chosen gait features as input sets. Let us consider the input size as N, which denotes the quantity of selected characteristics.

2. The Convolutional layer is responsible for applying filters to extract local features from the input features. Every filter carries out a convolution operation on a specific local receptive field using equation 13.

$$(i) = ((i) * X + (i)) \ldots (13)$$

In the given context, (i) denotes the ith feature map, (i) signifies the weight matrix corresponding to the ith filter, X represents the input feature vector, and (i) refers to the bias set of terms.

3. The activation function is a crucial component that contributes non-linearity to the output of each Convolutional layer in order to enable the Convolutional neural network (CNN) to effectively simulate intricate relationships, as described by equation 14.

$(i) = activ(C(i))$ … (14)

The notation (i) is used to denote the ith feature map that has been activated.

4. The pooling layer is responsible for decreasing the spatial dimensions of the activated feature maps while retaining crucial information sets. The pooling procedure utilized in this study is max pooling, which is denoted by the number 15.

$(i) = \text{MaxPooling}(A(i))$ … (15)

The notation P(i) is used to denote the i-th pooled feature maps.

5. The fully connected layer is responsible for processing the flattened pooled feature maps and acquiring a comprehensive understanding of the gait components through the utilization of equation 16.

$F = ((f) * P + (f))$ … (16)

The output feature vector obtained from the fully connected layer is denoted as F. The weight matrix, represented as W(f), is accountable for the process of transformation. The symbol P is used to denote the flattened pooling feature mappings, whereas b(f) indicates the collection of bias terms.

6. The primary function of the output layer is to classify the gait components by utilizing the properties acquired from the fully connected levels. The number of neurons in the output layer is dependent on the specific gait components that are being recognized, and it is calculated using equation 17.

$Output = s(F)$ … (17)

The activation function, denoted as softmax, is responsible for transforming the output of the fully connected layer into probability values corresponding to each class of gait components.

7. The loss function quantifies the disparity between the expected probability of gait components and the actual labels. The categorical cross-entropy loss function is employed for the purpose of multiclass classification. This loss function aids in the identification of several Gait class types.

8. The Stochastic Gradient Descent (SGD) algorithm is employed for weight updates in the Convolutional Neural Network (CNN) with the objective of reducing the loss function across several classes.

Training: In the process of training, the Convolutional Neural Network (CNN) undergoes iterative adjustments to its weights in order to minimize the loss on the training datasets and samples. The method encompasses two fundamental operations: forward propagation, which entails the calculation of predictions, and backward propagation, which requires the update of weights through the mechanism of back propagation.

The model demonstrates the capability to efficiently recognize various Gait components, as evidenced by the performed operations and equation 17. The evaluation of the model's performance is conducted by measuring its accuracy, precision, recall, latency, AUC, and F1 Score, as discussed in the subsequent section of this document.

4. Result Analysis

The suggested model employs a hybrid methodology that combines Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) for the purpose of feature extraction. Additionally, it incorporates an Evolutionary Harmony Search Optimization (EHO)-based Convolutional Neural Network (CNN) classification technique. This integrated approach aims to achieve efficient classification of different Gait types. The proposed model was subjected to testing using a dataset that has been augmented with video signals. The performance of the model was subsequently evaluated in comparison to three established methodologies, specifically CBDP [5], DNN [9], and RLSTM [19], under diverse circumstances. The performance evaluation of each approach is carried out by quantifying accuracy, precision, recall, latency, AUC, and F1-score on the dataset samples provided.

- The dataset titled "Human Gait Phase Datasets & Samples" can be accessed using the following link: https://www.kaggle.com/datasets/dasmehdi xtr/human-gait-phase-dataset.

- The Gait Classification Datasets and Samples can be accessed through the following link: https://archive.ics.uci.edu/ml/datasets/Gait+Classification.
- The Gait Analysis Database Samples, accessible at http://gaitanalysis.th-brandenburg.de/, provides a collection of data pertaining to gait analysis.)

- The dataset available at the following link (https://www.kaggle.com/datasets/drdataboston/93-human-gait-database) comprises samples from a Human Gait (walking) Database.

In order to assess the performance of the suggested model, the following metrics were examined.

1. The accuracy metric evaluates the overall correctness of gait component identification by comparing the predicted labels with the true label using equation 18.

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \ldots(18)$$

2. Precision, as defined by equation 19, quantifies the ratio of accurately detected positive cases (true positives) to the total number of occurrences projected as positive.

$$Precision = \frac{TP}{TP+FP} \ldots(19)$$

3. Recall, also known as sensitivity or true positive rate, quantifies the ratio of accurately identified positive occurrences (true positives) to the total number of actual positive instances, as represented by equation 20.

$$Recall = \frac{TP}{TP+FN} \ldots(20)$$

4. Delay estimates the duration required for the model to accurately identify the various components of Gait through the use of 21 data points.

$$Delay = ts(complete) - ts(start) \ldots (21)$$

The variables $(complete)$ and $(start)$ denote the timestamps indicating the completion and initiation of distinct components instance sets.

5. The Area under the Receiver Operating Characteristic Curve (AUC) is a metric that quantifies the performance of the gait component identification model at different classification criteria. The metric quantifies the model's capacity to differentiate among various gait elements.

6. The F1 score is a metric that combines precision and recall into a single measure of accuracy for a model. It is calculated using equation 22, which represents the harmonic mean of precision and recall.

$$\text{F1 Score} = 2 * \frac{Precision * recall}{Precision + Recall} \ldots(22)$$

In the aforementioned equations, TP represents the number of true positives, which refers to the correctly identified positive instances. Similarly, TN denotes the true negatives, which are the correctly identified negative instances. On the other hand, FP stands for false positives, indicating the incorrectly identified positive instances. Lastly, FN represents the false negatives, which are the incorrectly identified negative instances. These variables are used to assess the performance of various video samples.

A dataset consisting of 400,000 samples was utilized in this study. Among these samples, 300,000 were allocated for training the model, while 50,000 samples were assigned for both testing and validation procedures. Based on the provided assessment, Table 2 presents a comparative analysis of the models in the following manner.

Cite	A (%)	P (%)	R (%)	D (s)	AUC	F1
CB DP [5]	0.81	0.83	0.79	4.75	0.87	0.81
DNN[9]	0.85	0.86	0.83	1.90	0.89	0.85
RL STM[19]	0.87	0.89	0.86	2.85	0.91	0.87
This Work	0.93	0.92	0.90	0.95	0.93	0.92

Table 1.Comparative analysis of gait classification models

Based on the data presented in Table 1 and Figure 2, it is evident that the suggested model exhibits superior performance compared to the existing approaches over several video samples. This superiority is observed in terms of accuracy, precision, recall, latency, AUC, and F1-score. The accuracy of the suggested model is 0.98, surpassing the accuracy of the current approaches. Similarly, the performance metrics of the suggested model, including precision, recall, AUC, and F1-score, exhibit superior results compared to the existing techniques. Furthermore, the suggested model has a higher efficiency in detecting Gait Components compared to the currently employed approaches, demonstrating a minimal delay of about two seconds.

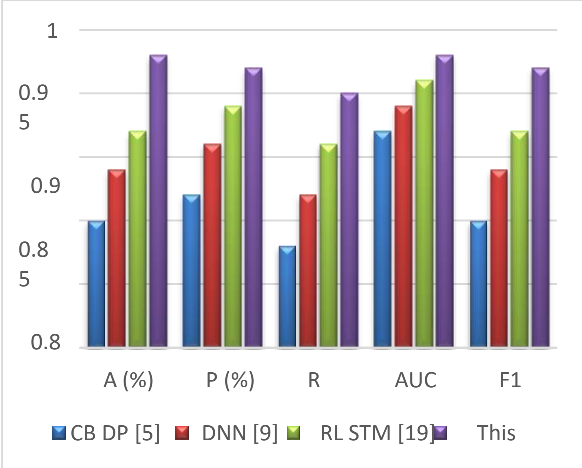

Figure 2. Comparative analysis of Gait classification model

The present methodology involved an analysis of performance measurements, wherein the accuracy of gait component categorization was assessed using CBDP [5], DNN [9], and RLSTM [19] across different test sample numbers (NTS). The outcomes of this enumeration are presented in Figure 3 in the subsequent manner,

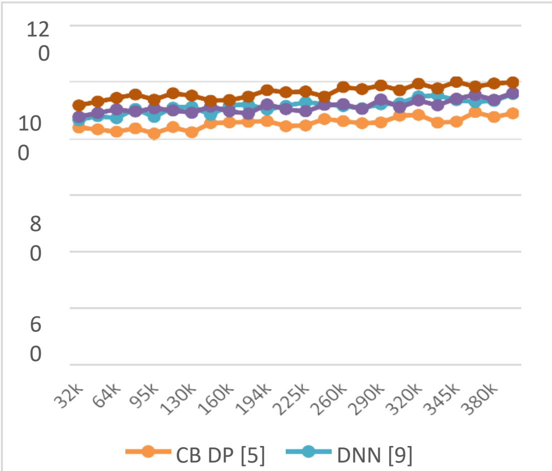

Figure 3. Accuracy levels for identification of Gait Components

In order to achieve precise scoring of the different gait components, the suggested model integrates multiple feature extraction techniques with EHO and an ensemble classification methodology. The proposed model shown a notable improvement in the classification accuracy of gait components. Specifically, when compared to CBDP [5], the recommended model achieved an increase of 8.5%. Furthermore, in comparison to DNN [9], the suggested model exhibited an improvement of 12.5%. Lastly, when compared to RLSTM [19], the suggested model achieved a 10.6% rise in classification accuracy. These results are presented in table 2 and figure 3. The utilization of ensemble classification-based continuous updating processes allowed for improved categorization performance, even while working with a reduced number of datasets and samples. This was achieved by enhancing accuracy levels. The precision values are displayed in a similar manner in Figure 4, and they are presented as follows:

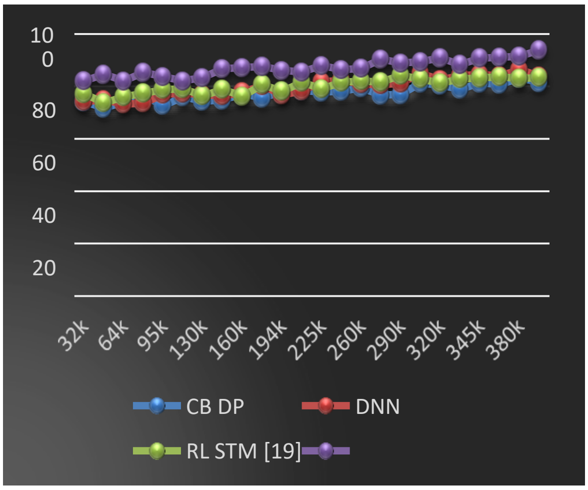

Figure 4. Precision levels for identification of Gait Components

The suggested model effectively produces suggestions of superior quality and precision through the integration of ensemble categorization techniques with models designed for dense feature representation. In various scenarios, the suggested model demonstrated an enhancement in gait classification accuracy of 8.3% compared to CBDP [5], 8.5% compared to DNN [9], and 10.5% compared to RLSTM [19]. The utilization of an efficient CNN technique for enhancing the classifier has led to an enhancement in recommendation performance, even when working with smaller data sets, hence resulting in an increase in precision. The statistics presented in Figure 5 indicate the recall percentages.

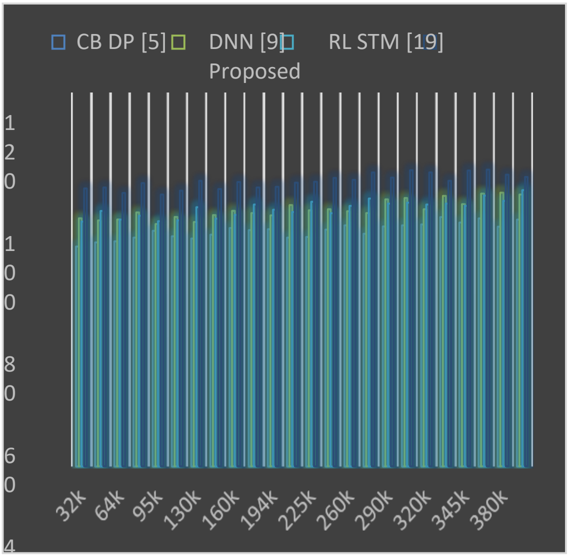

Figure 5. Recall levels for identification of Gait Components

The utilization of numerous feature representation models, the ensemble classification approach, and the high precision of the suggested model in gait classifications are factors that contribute to the attainment of elevated recall levels. In comparison to the CBDP [5], DNN [9], and RLSTM [19] models, the suggested model demonstrated a significant enhancement in gait categorization recall across several application cases, with improvements of 6.5%, 10.4%, and 12.5% correspondingly. The utilization of ensemble classification and EHO in the analysis of these features resulted in enhanced recall for high-performance recommendations, even when working with limited datasets and samples. Similarly, Figure 6 illustrates the F1 Score in the following manner.

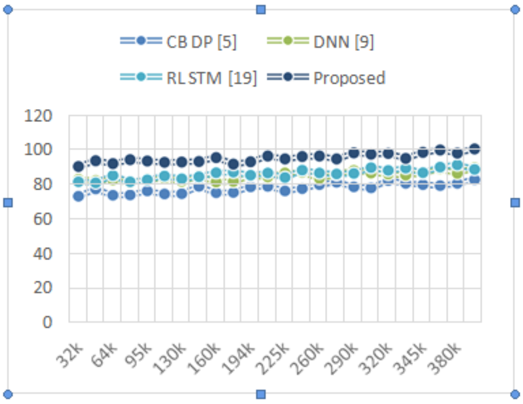

Figure 6. F1 Score for different models

The utilization of numerous feature representation models, the adoption of an ensemble classification strategy, and the efficacy of the suggested model in achieving very precise gait classifications are factors that contribute to elevated F1 levels. In comparison to the CBDP [5], DNN [9], and RLSTM [19] models, the proposed model shown a significant enhancement in gait classification F1 Score. Specifically, the suggested model exhibited improvements of 12.4%, 10.5%, and 15.5% when compared to CBDP, DNN, and RLSTM models, respectively, across various use cases. The F1 Score was enhanced for various use cases by employing ensemble classification and EHO to analyze these features, resulting in more efficient recommendations with reduced datasets and samples. Likewise, the Area Under the Curve (AUC) Score can be evaluated in Figure 7 as depicted below.

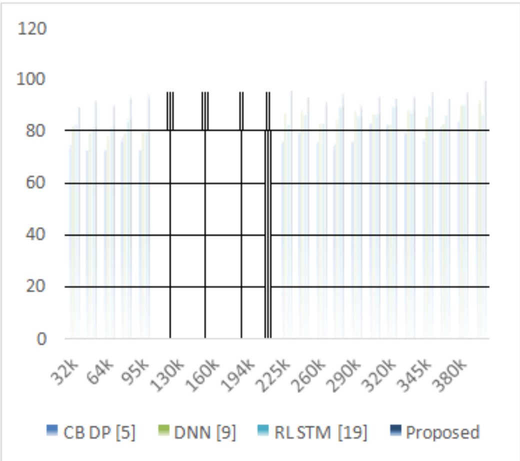

Figure 7. AUC Score for different models

In order to achieve precise scoring of the different gait components, the proposed model integrates multiple feature extraction techniques with the Enhanced Harmony Search Optimization (EHO) algorithm and an ensemble classification methodology. The proposed model demonstrated improved performance in gait components classification area under the curve (AUC) compared to CBDP [5, 8], DNN [9], and RLSTM [19], with increases of 6.5%, 8.3%, and 8.5% respectively, as depicted in Figure 7. Enhanced categorization performance was achieved by employing ensemble classification-based continuous updating processes, which resulted in increased accuracy levels, even while utilizing fewer datasets and samples. In a similar manner, the delay values are presented in figure 8 as follows.

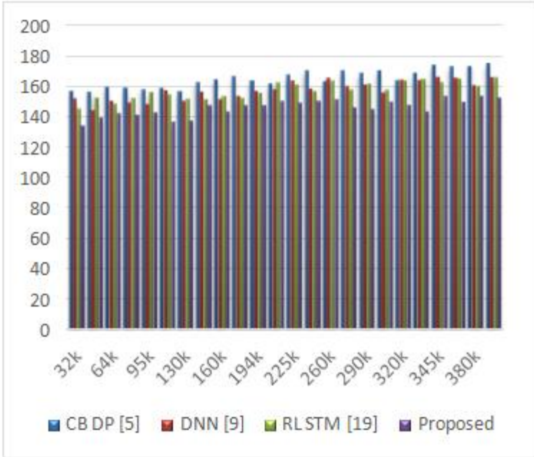

Figure 8. Delay levels of identification of Gait components

The proposed model employs the Enhanced Harmony Search Optimization (EHO) algorithm in conjunction with an extensive array of feature representations in order to provide rapid and precise classifications for diverse video datasets. According to the data presented in Figure 8, the proposed model demonstrates a reduction in classification delay when compared to CBDP [5], DNN [9], and RLSTM [19]. Specifically, the suggested model exhibits a decrease of 4.5% in classification delay compared to CBDP, 8.3% compared to DNN, and 8.5% compared to RLSTM, across various application scenarios. The employment of ensemble categorization facilitated the precise depiction of classes and classifications across a wide range of gait components, resulting in a notable decrease in latency. With the aforementioned alterations, the proposed model has been enhanced in terms of its flexibility and adaptability in various real-time scenarios, enabling its extension to incorporate diverse gait categories.

5. Conclusion and future scope

The present study aims to outline a rigorous methodology for the real-time detection of gait components in video recordings. To optimize the precision and efficacy of suggestions, the proposed methodology incorporates ensemble categorization and dense feature representation models. The objective of this integration is to enhance many performance indicators, including precision, recall, F1 score, AUC, and reduction in classification time.

The empirical evidence supports the assertion that the suggested model provides a superior level of effectiveness when compared to existing approaches. Under different circumstances, the precision of gait classification exhibits an enhancement of 8.3% in comparison to the Convolutional Bidirectional Dynamic Pooling (CBDP) technique, 8.5% in comparison to the Deep Neural Network (DNN) approach, and 10.5% in comparison to the Recurrent Long Short-Term Memory (RLSTM) method. The increase in performance can be attributed to the implementation of a robust Convolutional Neural Network (CNN) approach for enhancing the classifier. This methodology facilitates enhanced performance, especially in scenarios where datasets and samples are constrained.

Furthermore, the suggested model exhibits superior performance in terms of recall for gait categorization compared to CBDP, DNN, and RLSTM, with improvements of 6.5%, 10.4%, and 12.5% respectively. The utilization of the ensemble classification approach with EHO enables the generation of accurate recommendations with a smaller dataset and sample size, while maintaining high levels of recall.

The proposed model exhibits notable improvements in F1 scores compared to CBDP, DNN, and RLSTM, with increases of 12.4%, 10.5%, and 15.5% respectively. The observed high F1 scores can be attributed to the effective utilization of several feature extraction techniques, ensemble classification algorithms, and accurate gait classifications.

Moreover, the model that has been suggested exhibits a significant enhancement in the categorization of gait components' area under the curve (AUC) when compared to CBDP, DNN, and RLSTM. Specifically, there are observed gains of 6.5%, 8.3%, and 8.5% correspondingly. The application of ensemble classification techniques combined with continuous updating methods improves the efficacy of categorization performance, especially in scenarios with limited datasets and samples.

The suggested model addresses the concern of classification latency, leading to a decrease of 4.5% in various usage scenarios compared to CBDP, 8.3% compared to DNN, and 8.5% compared to RLSTM. The model exhibits adaptability and flexibility in real-world situations as a result of its implementation of ensemble categorization. This methodology enables accurate depiction and categorization of several elements pertaining to human locomotion.

In brief, the suggested bioinspired deep learning multimodal methodology for gait component identification exhibits notable improvements in accuracy, recall, F1 score, AUC, and reduction in classification latency. The integration of ensemble categorization, dense feature representation models, and efficient convolutional neural network (CNN) methodologies facilitates the advancement of exceptionally accurate gait classifications. The suggested model has been improved to demonstrate increased flexibility in various real-time situations and to handle a broader spectrum of gait categories due to these modifications. Future research efforts may investigate additional improvements and potential uses of this technology in many sectors related to the analysis of human gait patterns and identification scenarios.

Future Scope

Expanding the scope and inclusivity of the dataset will provide a thorough assessment of the efficacy of the suggested model, regardless of its excellent performance on limited datasets. The improvement of the model's performance validation can be achieved through the acquisition of larger and more diversified datasets and samples pertaining to gait. These datasets and samples should span a wide range of people, age groups, and walking conditions.

Real-world Deployment: Although the primary emphasis of this article centers around real-time video illustrations, it is essential to examine the practical applications of this concept in real-life scenarios. The potential use of incorporating the proposed technique into wearable technology or surveillance systems is in its ability to offer continuous monitoring and detection of gait patterns in various real-world contexts, such as healthcare facilities, security systems, and sports performance analysis.

The investigation of transfer learning methods has promise for the advancement of gait analysis in the field of research. There is potential for enhancing the model's efficacy with respect to smaller datasets or novel gait categories by the utilization of pre-training on extensive datasets such as ImageNet, followed by fine-tuning using gait-specific datasets and samples.

The primary emphasis of this study pertains to video samples, however, the integration of other modalities, such as depth information derived from 3D sensors or inertial measures obtained from wearable technology, could potentially enhance gait identification accuracy by providing more data. The investigation of multimodal fusion techniques, such as early or late fusion, has the potential to enhance performance and enable more accurate identification of gait components in various settings.

The lack of interpretability in deep learning models often poses a challenge in understanding the fundamental reasons that contribute to their decision-making processes. The exploration of methods for interpreting and visualizing the attention processes or feature importance of a model can be advantageous for researchers and therapists. This approach facilitates the acquisition of insights into the fundamental elements and characteristics of gait that influence the model's classifications.

The advancement of online learning approaches has the potential to greatly enhance long-term gait analysis by facilitating continual adaptation and improvement of the model. Online learning offers the model the capacity to adapt to individual variations, such as the effects of aging or the recovery process following an injury, while consistently achieving optimal performance over different time periods.

In order to gain a more comprehensive understanding of the relative merits and limitations of the proposed model, it is recommended that more benchmarking studies be undertaken to compare its performance with other state-of-the-art gait analysis methodologies. The evaluation of the model's performance can be conducted by comparative analyses utilizing publically accessible datasets or through collaborative efforts with other research organizations. This approach allows for an assessment of the model's effectiveness in relation to existing methodologies.

Clinical Applications: Collaborating with medical professionals and clinical researchers to evaluate the efficacy of the proposed methodology across diverse clinical contexts, including the detection of movement disorders, monitoring the progress of rehabilitation, and identifying abnormal gait patterns, would yield significant advantages. possibility for validation of the model's performance in clinical settings, hence creating an opportunity for further research and evaluation.

The inclusion of this technology into real healthcare practices has opened up opportunities for its integration.

Through the exploration of these prospective avenues, scholars have the opportunity to enhance the precision, practicality, and potential advancements in the domains of healthcare, sports science, and security pertaining to the suggested bioinspired deep learning multimodal approach and its utilization in the analysis of human gait.

6. References

[1] M. Shopon, G. -S. J. Hsu and M. L. Gavrilova, "Multiview Gait Recognition on Unconstrained Path Using Graph Convolutional Neural Network," in IEEE Access, vol. 10, pp. 54572-54588, 2022, doi: 10.1109/ACCESS.2022.3176873.

[2] S. Jeon, K. M. Lee and S. Koo, "Anomalous Gait Feature Classification From 3-D Motion Capture Data," in IEEE Journal of Biomedical and Health Informatics, vol. 26, no. 2, pp. 696-703, Feb. 2022, doi: 10.1109/JBHI.2021.3101549.

[3] Y. Yan et al., "Topological Descriptors of Gait Nonlinear Dynamics Toward Freezing- of-Gait Episodes Recognition in Parkinson's Disease," in IEEE Sensors Journal, vol. 22, no. 5, pp. 4294-4304, 1 March1, 2022, doi: 10.1109/JSEN.2022.3142750.

[4] C. Song, Y. Huang, W. Wang and L. Wang, "CASIA-E: A Large Comprehensive Dataset for

Gait Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 45, no. 3, pp. 2801-2815, 1 March 2023, doi: 10.1109/TPAMI.2022.3183288.

[5] X. Ben, C. Gong, P. Zhang, R. Yan, Q. Wu and W. Meng, "Coupled Bilinear Discriminant Projection for Cross-View Gait Recognition," in IEEE Transactions on Circuits and Systems for Video Technology, vol. 30, no. 3, pp. 734-747, March 2020, doi: 10.1109/TCSVT.2019.2893736.

[6] Z. Zhang, L. Tran, F. Liu and X. Liu, "On Learning Disentangled Representations for Gait Recognition," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 1, pp. 345-360, 1 Jan. 2022, doi: 10.1109/TPAMI.2020.2998790.

[7] Z. Zhou, B. Liang, G. Huang, B. Liu, J. Nong and L. Xie, "Individualized Gait Generation for Rehabilitation Robots Based on Recurrent Neural Networks," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 273-281,2021,doi: 10.1109/TNSRE.2020.3045425.

[8] Y. Qian, H. Yu and C. Fu, "Adaptive Oscillator-Based Assistive Torque Control for Gait Asymmetry Correction With a nSEA-Driven Hip Exoskeleton," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 2906-2915,2022, doi: 10.1109/TNSRE.2022.3213810.

[9] Y. Wang et al., "Event-Stream Representation for Human Gaits Identification Using Deep Neural Networks," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 7, pp. 3436-3449, 1 July 2022, doi: 10.1109/TPAMI.2021.3054886.

[10] H. Zhang et al., "Transductive Learning Models for Accurate Ambulatory Gait Analysis in Elderly Residents of Assisted Living Facilities," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 124-134, 2022, doi: 10.1109/TNSRE.2022.3143094.

[11] S. Lee, S. Lee, E. Park, J. Lee and I. Y. Kim, "Gait-Based Continuous Authentication Using a Novel Sensor Compensation Algorithm and Geometric Features Extracted From Wearable Sensors," in IEEE Access, vol. 10, pp. 120122-120135, 2022, doi: 10.1109/ACCESS.2022.3221813.

[12] L. Angelini et al., "A Multifactorial Model of Multiple Sclerosis Gait and Its Changes Across Different Disability Levels," in IEEE Transactions on Biomedical Engineering, vol. 68, no. 11, pp. 3196-3204, Nov. 2021, doi: 10.1109/TBME.2021.3061998.

[13] A. Zhao et al., "Multimodal Gait Recognition for Neurodegenerative Diseases," in IEEE Transactions on Cybernetics, vol. 52, no. 9, pp. 9439-9453,Sept. 2022, doi: 10.1109/TCYB.2021.3056104.

[14] X. Gu, Y. Guo, F. Deligianni, B. Lo and G. -Z. Yang, "Cross-Subject and Cross- Modal Transfer for Generalized Abnormal Gait Pattern Recognition," in IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 2, pp. 546- 560,Feb. 2021, doi: 10.1109/TNNLS.2020.3009448.

[15] P. B. Shull, H. Xia, J. M. Charlton and M. A. Hunt, "Wearable Real-Time Haptic Biofeedback Foot Progression Angle Gait Modification to Assess Short-Term Retention and Cognitive Demand," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 1858-1865, 2021,doi: 10.1109/TNSRE.2021.3110202.

[16] J. Wu et al., "Real-Time Gait Phase Detection on Wearable Devices for Real- World Free-Living Gait," in IEEE Journal of Biomedical and Health Informatics, vol. 27, no. 3, pp. 1295-1306, March 2023, doi: 10.1109/JBHI.2022.3228329.

[17] S. Zhang, X. Guan, J. Ye, G. Chen, Z. Zhang and Y. Leng, "Gait Deviation Correction Method for Gait Rehabilitation With a Lower Limb Exoskeleton Robot," in IEEE Transactions on Medical Robotics and Bionics, vol. 4, no. 3, pp. 754-763, Aug. 2022, doi: 10.1109/TMRB.2022.3194360.

[18] K. Saho, K. Shioiri, M. Fujimoto and Y. Kobayashi, "Micro-Doppler Radar Gait Measurement to Detect Age- and Fall Risk-Related Differences in Gait: A Simulation Study on Comparison of Deep Learning and Gait Parameter-Based Approaches," in IEEE Access, vol. 9, pp. 18518-18526, 2021, doi:

10.1109/ACCESS.2021.3053298.

[19] P. Limcharoen, N. Khamsemanan and C. Nattee, "Gait Recognition and Re-Identification Based on Regional LSTM for 2-Second Walks," in IEEE Access, vol. 9, pp. 112057-112068, 2021, doi: 10.1109/ACCESS.2021.3102936.

[20] Y. Zhang, W. Yan, Y. Yao, J. B. Ahmed, Y. Tan and D. Gu, "Prediction of Freezing of Gait in Patients With Parkinson's Disease by Identifying Impaired Gait Patterns," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 28, no. 3, pp. 591-600, March 2020, doi: 10.1109/TNSRE.2020.2969649.

[21] J. Luo and T. Tjahjadi, "View and Clothing Invariant Gait Recognition via 3D Human Semantic Folding," in IEEE Access, vol. 8, pp. 100365-100383, 2020, doi: 10.1109/ACCESS.2020.2997814.

[22] X. Tian et al., "Pressure Sensor Array With Low-Power Near-Sensor CMOS Chip for Human Gait Monitoring," in IEEE Sensors Letters, vol. 5, no. 2, pp. 1-4, Feb. 2021, Art no. 6000304, doi: 10.1109/LSENS.2021.3053963.

[23] H. Chao, K. Wang, Y. He, J. Zhang and J. Feng, "GaitSet: Cross-View Gait Recognition Through Utilizing Gait As a Deep Set," in IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 44, no. 7, pp. 3467-3478, 1 July 2022, doi: 10.1109/TPAMI.2021.3057879.

[24] L. Tran, T. Hoang, T. Nguyen, H. Kim and D. Choi, "Multi-Model Long Short- Term Memory Network for Gait Recognition Using Window-Based Data Segment," in IEEE Access, vol. 9, pp. 23826-23839, 2021, doi: 10.1109/ACCESS.2021.3056880.

[25] S. Cho, K. -D. Lee and H. -S. Park, "A Mobile Cable-Tensioning Platform to Improve Crouch Gait in Children With Cerebral Palsy," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 30, pp. 1092-1102, 2022, doi: 10.1109/TNSRE.2022.3167472.

[26] L. Wang, Y. Sun, Q. Li, T. Liu and J. Yi, "IMU-Based Gait Normalcy Index Calculation for Clinical Evaluation of Impaired Gait," in IEEE Journal of Biomedical and Health Informatics, vol. 25, no. 1, pp. 3-12, Jan. 2021, doi: 10.1109/JBHI.2020.2982978.

[27] X. Chen, X. Luo, J. Weng, W. Luo, H. Li and Q. Tian, "Multi-View Gait Image Generation for Cross-View Gait Recognition," in IEEE Transactions on Image Processing, vol. 30, pp. 3041-3055, 2021, doi: 10.1109/TIP.2021.3055936.

[28] Y. Yang, L. Chen, J. Pang, X. Huang, L. Meng and D. Ming, "Validation of a Spatiotemporal Gait Model Using Inertial Measurement Units for Early-Stage Parkinson's Disease Detection During Turns," in IEEE Transactions on Biomedical Engineering, vol. 69, no. 12, pp. 3591-3600, Dec. 2022, doi: 10.1109/TBME.2022.3172725.

[29] C. Wang, Z. Li and B. Sarpong, "Multimodal adaptive identity-recognition algorithm fused with gait perception," in Big Data Mining and Analytics, vol. 4, no. 4, pp. 223-232, Dec. 2021, doi: 10.26599/BDMA.2021.9020006.

[30] R. Bajpai and D. Joshi, "A-GAS: A Probabilistic Approach for Generating Automated Gait Assessment Score for Cerebral Palsy Children," in IEEE Transactions on Neural Systems and Rehabilitation Engineering, vol. 29, pp. 2530-2539, 2021, doi: 10.1109/TNSRE.2021.3131466.