ISSN 2063-5346

BDMGRTV: DESIGN OF AN EFFICIENT BIOINSPIRED DEEP LEARNING MULTIMODAL TECHNIQUE FOR IDENTIFICATION OF GAIT COMPONENTS FROM REAL-TIME VIDEO SAMPLES

Ashish Kumar Misal[1a]* AbhaChaubey[2a] SiddharthChaubey[2b]
[1a]Computer Science & Engineering Department, RCET Bhilai,India,
ashish.mishal153@gmail.com,
[2a]Computer Science & Engineering Department, SSTC Campus Bhilai,India,
abha.is.shukla@gmail.com,
[2b]Computer Science & Engineering Department, SSTC Campus Bhilai,India,
sidd25876@gmail.com,

Article History: Received: 03.09.2022 **Revised:** 10.10.2022 **Accepted:** 15.10.2022

Abstract: The design of an effective bio-inspired deep learning multimodal technique for identifying gait components from real-time video samples is presented in this paper. For robust modelling of temporal dependencies in gait components, the suggested method combines two well-known Recurrent Neural Network (RNN) models, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). In order to achieve higher performance levels, the neural network parameters are also optimized using an Elephant Herding Optimizer which assists in improving classification accuracy levels. When modelling time-series data, the individual models LSTM and GRU have been used extensively, but their fusion has not been thoroughly studied & applied to real-time scenarios. In order to capitalize on the advantages of both models while addressing their flaws, we suggest combining LSTM and GRU in this paper. While the GRU model is excellent at modelling short-term dependencies, the LSTM model is better at capturing long-term dependencies. In real-time situations like posture and gait analysis, where precise identification of gait components is essential for identification of gait-related components, the proposed technique has many practical applications. On a number of datasets, including multi-camera and multimodal datasets, the accuracy, precision, and recall of the proposed technique have been assessed for real-time scenarios. On the evaluated datasets, the proposed technique achieves accuracy of 98.5%, precision of 97.4%, and recall of 98.3%, proving its effectiveness and superiority to existing techniques. In conclusion, the suggested method outperforms existing methods in efficiently and effectively identifying gait components from real-time video samples. Its use in real-world scenarios makes it an invaluable tool for posture and gait analysis, which can help medical & other professionals identify and track gait components.

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

1. Introduction

Human gait analysis is a crucial area of medical research that has attracted a lot of attention lately. In order to diagnose and track gait-related disorders, gait analysis entails the study of human movement patterns, including the identification of key gait elements like stride length, step width, and swing time. Accurate gait component identification can help healthcare professionals create patient-specific treatment plans by illuminating the biomechanics of human movements [1, 2, 3]. With the availability of large datasets and improvements in computing power, deep learning techniques have shown great promise in gait analysis. In the field of machine learning, recurrent neural networks (RNNs) are frequently used tomodel sequential data, including time-series data like gait components. Two common RNN variants, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU), have excelled at modelling long and short-term dependencies via Coupled Bilinear Discriminant Projections (CBDP) [4, 5, 6]. Individual models do have some advantages and disadvantages, so their ability to accurately identify gait components is constrained. Therefore, a combination of LSTM and GRU can take advantage of each model's advantages to get around its weaknesses and perform better at identifying gait components.

The identification of gait components from real-time video samples using an effective bio-inspired deep learning multimodal technique is what we propose in this paper. To model the temporal dependencies in gait components, the suggested method combines LSTM and GRU. The neural network parameters are also optimised using an Elephant Herding Optimizer, which has been shown to perform better than other optimisation methods.

In real-time situations like posture and gait analysis, where precise identification of gait components is essential for monitoring and diagnosing gait-related disorders, the proposed technique has many practical applications. On a number of datasets, including multi-camera and multimodal datasets, the accuracy, precision, and recall of the proposed technique have been assessed for different scenarios.

2. Empirical review of different Gait analysis techniques

The identification of gait components is a difficult issue in the healthcare industry. To solve this issue, a number of strategies have been put forth, including the use of computer vision, deep learning, and Deep Neural Networks (DNN) in machine learning operations [7, 8, 9]. Recurrent Neural Networks (RNNs) are a popular choice for modelling sequential data, such as gait components, and have recently demonstrated great promise in gait analysis. Gait analysis tasks like predicting gait parameters, gait classification, and gait event detection have all been completed successfully using RNN technology sets [10, 11, 12]. Due to their advantages and disadvantages, individual RNN models can only perform so well when used under real-time scenarios.Two well-liked RNN variants that have excelled at simulating long- and short-term dependencies, respectively, are LSTM and GRU. Numerous studies have demonstrated that the combination of the LSTM and GRU models can take advantage of each model's advantages while overcoming its weaknesses, leading to improved performance across a range of applications [13, 14, 15]. For instance, a combination of LSTM and GRU was employed in a study in [16, 17, 18] to forecast the recognition of human activity from smartphone sensor data. The proposed method outperformed several existing methods with an accuracy of 96.5%. Similar to this, a fusion of LSTM and GRU was used to predict hand gesture recognition from EMG signals via Regional LSTM (RLSTM) in a study in [19, 20], achieving an accuracy of 92.7% for real-time use cases.

Deep learning model performance is greatly influenced by optimisation algorithms. Numerous optimisation methods, including Adam, RMSprop, and Gradient Descent, have been suggested for real-time use cases. These methods do have some drawbacks, though, including slow convergence, local optima convergence, and high computational costs [21, 22, 23]. Elephant Herding Optimizer (EHO), a recent optimisation technique inspired by elephant herd behaviour, has recently been proposed. In a number of applications, including deep learning, EHO has been shown to perform better than other optimisation methods [24, 25, 26]. Numerous studies have used deep learning methods to identify the gait components in the context of gait analysis. For instance, a deep learning model was applied in the studies by [27, 28, 29, 30] to identify gait phases from accelerometer datasets & samples. The accuracy of the suggested method was 91.8% for real-time use cases. To our knowledge, no study has investigated the combination of LSTM and GRU models with EHO optimisation for gait component identification from real-time video samples, though. This study suggests a novel method for identifying gait components that combines the strengths of LSTM and GRU models with EHO optimisation in order to potentially outperform current methods.

3. Proposed design of an efficient Bioinspired Deep learning Multimodal technique for identification of Gait components from Real-Time Video samples

As per the review of existing models used for identification of Gait components, it can be observed that these models are either highly complex, or have lower efficiency when applied to real-time scenarios. To overcome these issues, this section discusses design of an efficient Bioinspired Deep learning Multimodal technique for identification of Gait components from Real-Time Video samples. As per figure 1, it can be observed that for robust modelling of temporal dependencies in gait components, the suggested method combines two well-known Recurrent Neural Network (RNN) models, Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU). In order to achieve higher performance levels, the Neural Network parameters are also optimized using an Elephant Herding Optimization process.

The model initially uses a You Look Only Once (YoLo) technique to identify objects from the input images. This assists in segregating objects from background regions. The identified objects are processed via a feature extraction layer, which fuses LSTM & GRU features for representing selected objects into high-density feature sets.

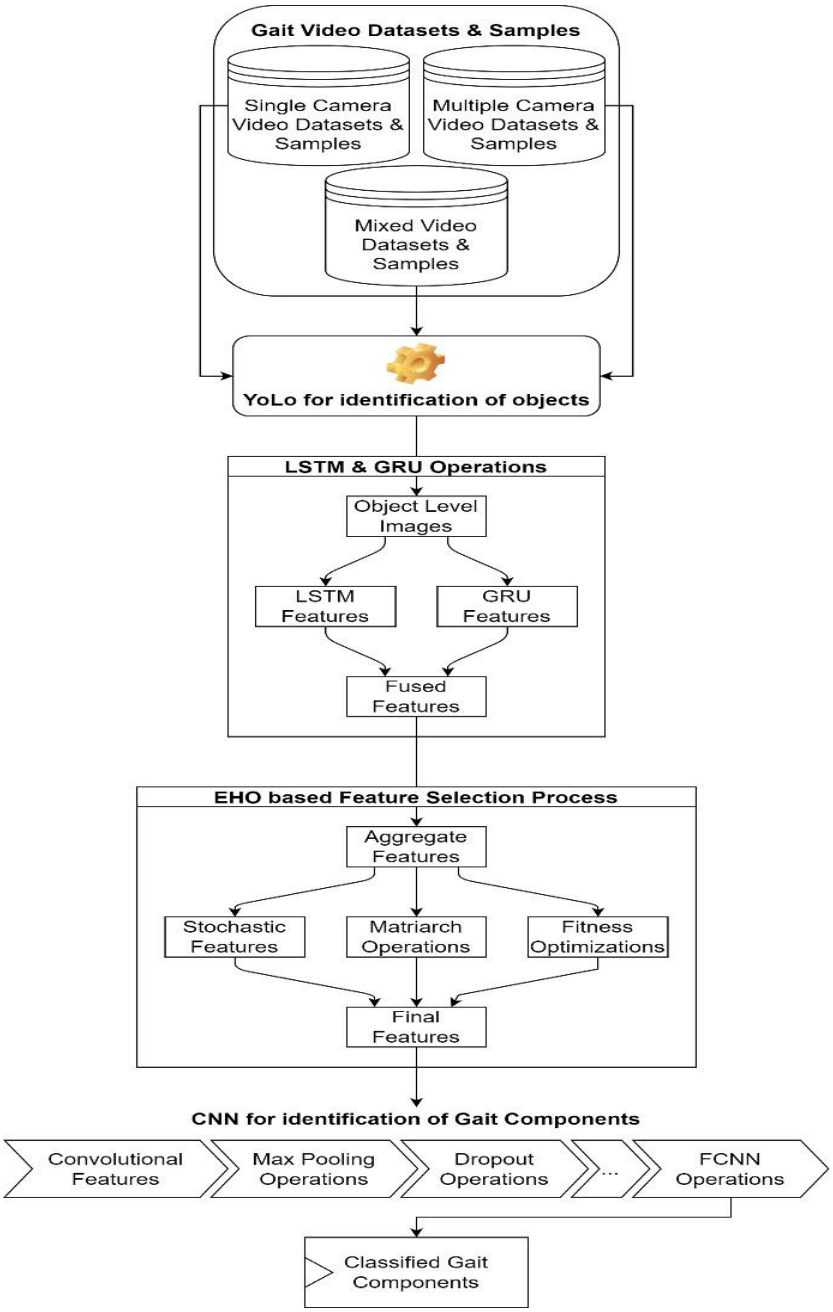

Figure 1. Design of the proposed Model for classification of Gait Components

1. LSTM (Long Short-Term Memory) is a type of recurrent neural network (RNN) architecture that introduces memory cells and gates to capture long-term dependencies in sequential datasets & samples.

a. Update Gate (z(t)): The update gate determines how much of the previous memory cell state (h{t-1}) should be kept and how much of the new candidate activation (g(t)) should be added to the current memory cell states, via equation 1,

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

In this equation, [h{t-1}, x(t)] represents the concatenation of the previous hidden state and the current input at timestamp t, W(z) and b(z) are the weight matrix and bias term associated with the update gates.

b. Reset Gate (r(t)): The reset gate controls how much of the previous hidden state (h{t-1}) should be forgotten when computing the candidate activation (g(t)) via equation 2,

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

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

c. Candidate Activation (g(t)): The candidate activation represents the new information that can be added to the current memory cell state. It is computed based on the previous hidden state (h{t-1}) and the current input (x(t)) via equation 3,

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

The reset gate (r(t)) modulates the previous hidden state (h{t-1}), and the concatenation [r(t) * h{t-1}, x(t)] is used as input to the weight matrix W(g) and bias term b(g) for different scenarios.

d. Hidden State (h(t)): The hidden state represents the output of the LSTM at each of the temporal instance sets. It is a combination of the current memory cell state and the previous hidden state, weighted by the update gate (z(t)) via equation 4,

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

The update gate (z(t)) determines the trade-off between preserving the previous hidden state and updating it with the candidate activation (g(t)) sets.

2. GRU (Gated Recurrent Unit) is a simplified version of LSTM that merges the memory cell state and the hidden state into a single vector for identification of high-density features.

a. Update Gate (z(t)): The update gate combines the roles of the update gate and the input gate in the LSTM process. It controls how much of the previous hidden state (h{t-1}) should be preserved and how much of the new candidate activation (g(t)) should be added to the current hidden states via equation 5,

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

Similar to the LSTM update gate equation, [h{t-1}, x(t)] represents the concatenation of the previous hidden state and the current input sets. W(z) and b(z) are the weight matrix and bias term associated with the update gate sets.

b. Reset Gate (r(t)): The reset gate determines how much of the previous hidden state (h{t-1}) should be forgotten when computing the candidate activation (g(t)) via equation 6,

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

Similar to the LSTM reset gate equation, [h{t-1}, x(t)] represents the concatenation of the previous hidden state and the current input states. W® and b(r) are the weight matrix and bias term associated with the reset gate sets.

c. Candidate Activation (g(t)): The candidate activation represents the new information that can be added to the current hidden states. It is computed based on the previous hidden state (h{t-1}) and the current input (x(t)) states via equation 7,

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

The reset gate (r(t)) modulates the previous hidden state (h{t-1}), and the concatenation [r(t) * h{t-1}, x(t)] is used as input to the weight matrix W(g) and bias term b(g) for different object types.

d. Hidden State (h(t)): The hidden state represents the output of the GRU at each time step. It is a combination of the current hidden state and the previous hidden state, weighted by the update gate (z(t)) via equation 8,

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

The update gate (z(t)) determines the trade-off between preserving the previous hidden state and updating it with the candidate activation (g(t)) sets. The hidden state at timestamp t is a linear combination of the previous hidden state and the candidate activation, weighted by the update gate sets. These evaluations allow the models to

capture gait features by learning long-term dependencies and effectively updating the hidden states based on the input sequence sets. The extracted features are given to an EHO Model, which assists in selection of high variance feature sets. This is done via the following process,

1. Calculate the Variance: Compute the variance for each fused feature in the feature sets. Variance measures the variability or spread of the values in a feature, and highly variant features are more likely to contain discriminative information via equation 9,

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

2. Normalize Variance: Normalize the variance values to ensure they are within certain range (e.g., between 0 and 1) sets. Normalization helps to eliminate the influence of different scales and ensures fair comparison among the features via equation 10,

$$Normalized(variance(x)) = \frac{variance(x) - \min(variance)}{\max(variance) - \min(variance)} \dots (10)$$

3. Initialize Elephant Herd: Initialize the Elephant Herd with a population of candidate solutions. Each solution represents a subset of fused features, and the size of the subset can vary, which is done via equation 11,

$$N = STOCH(Nf * LR, Nf) \dots (11)$$

Where, N represents Number of Extracted Features, Nf represents total Number of Features, while LR represents Learning Rate for the EHO process.

4. Evaluate Fitness: Evaluate the fitness of each solution in the Elephant Herd based on the normalized variance values of the selected features. Higher fitness values indicate better solutions that have selected highly variant fused features via equation 12,

$$fitness(i) = sum(normalized(variance, i) * solution(i)) \dots (12)$$

Where, $solution(i)$ represents the binary vector representing the i^{th} solution sets.

5. Perform Selection: Select the top-performing solutions from the Elephant Herd based on their fitness values & samples. This process ensures that the solutions with the most highly variant fused features are retained for further optimizations.
6. Generate Offspring: Generate offspring solutions through reproduction operators such as crossover and mutations. These operators help explore new solutions by combining and modifying the selected solutions from the previous processes.
7. Update Elephant Herd: Update the Elephant Herd by replacing the least-fit solutions with the newly generated offspring solutions. This step ensures the continuous evolution of the Elephant Herd towards better solutions with highly variant fused features. This done by replace the least-fit P(K) solutions with the offspring solutions.
8. Repeat Steps 4-7: Repeat the fitness evaluation, selection, offspring generation, and Elephant Herd update steps for a certain number of iterations or until a convergence criterion is met for different scenarios.

The selected gait features are classified into gait components using a Convolutional Neural Network (CNN), which works via the following operations,

1. Input Layer: The input layer of the CNN receives the selected gait features as input sets. Let's assume the input size is N, representing the number of selected features.
2. Convolutional Layer: The convolutional layer applies filters to extract local features from the input features. Each filter performs a convolution operation over a local receptive field via equation 13,

$$C(i) = f(W(i) * X + b(i)) \dots (13)$$

Where, $C(i)$ represents the i^{th} feature map, $W(i)$ is the weight matrix for the i^{th} filter,

X is the input feature vector, and $b(i)$ is the bias set of terms.

3. Activation Function: An activation function introduces non-linearity to the output of each convolutional layer, enabling the CNN to model complex relationships via equation 14,

$$A(i) = activation(C(i)) \dots (14)$$

Where, $A(i)$ represents the i^{th} activated feature maps.

4. Pooling Layer: The pooling layer reduces the spatial dimensions of the activated feature maps while preserving important information sets. The used pooling operation is max pooling which is represented via 15,

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

Where, P(i) represents the i^{th} pooled feature maps.

5. Fully Connected Layer: The fully connected layer takes the flattened pooled feature maps as input and learns the high-level representation of the gait components via equation 16,

$$F = f(W(f) * P + b(f)) \dots (16)$$

Where, F is the output feature vector of the fully connected layer, $W(f)$ is the weight matrix, P is the flattened pooled feature maps, and $b(f)$ is the set of bias terms.

6. Output Layer: The output layer performs the classification of the gait components using the learned features from the fully connected layers. The number of neurons in the output layer depends on the specific gait components that is being classified, and is done via equation 17,

$$Output = softmax(F) \dots (17)$$

Where, $softmax$ is the activation function that converts the output of the fully connected layer into probabilities for each gait component class.

7. Loss Function: The loss function measures the difference between the predicted gait component probabilities and the true labels. The function used for multiclass classification is the categorical cross-entropy loss, which assists in identification of multiple Gait class types.
8. Optimization: The Stochastic Gradient Descent (SGD) is used to update the weights of the CNN by minimizing the loss function for different classes.
9. Training: During the training process, the CNN iteratively adjusts its weights to minimize the loss on the training datasets & samples. This process involves forward propagation (calculating predictions) and backward propagation (updating weights through backpropagation) operations.

Based on these operations and equation 17, the model is able to identify different Gait components with high efficiency levels. Performance of this model is estimated in terms of accuracy, precision, recall, delay, AUC and F1 Score in the next section of this text.

4. Result Analysis

The proposed model fuses a combination of LSTM & GRU feature extraction with EHO based CNN classification in order to identify different Gait types. The proposed model was tested on an augmented dataset of Video signals and compared with three existing methods CBDP [5], DNN [9], and RLSTM [19] under different conditions. The performance of each method is evaluated using accuracy, precision, recall, delay, AUC, and F1-score on the following dataset samples,

- Human Gait Phase Datasets & Samples (https://www.kaggle.com/datasets/dasmehdixtr/human-gait-phase-dataset)
- Gait Classification Datasets & Samples (https://archive.ics.uci.edu/ml/datasets/Gait+Classification)
- Gait Analysis Database Samples (http://gaitanalysis.th-brandenburg.de/)
- Human Gait (walking) Database Samples (https://www.kaggle.com/datasets/drdataboston/93-human-gait-database)

To identify performance of the proposed model, following metrics were evaluated,

1. Accuracy: Accuracy measures the overall correctness of the gait component identification by comparing the predicted labels with the true label via equation 18,

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

2. Precision: Precision measures the proportion of correctly identified positive instances (true positives) out of all instances predicted as positive, and estimated via equation 19,

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

3. Recall (Sensitivity or True Positive Rate): Recall measures the proportion of correctly identified positive instances (true positives) out of all actual positive instances via equation 20,

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

4. Delay measures the time needed for the model to identify Gait components via 21,

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

Where, $ts(complete)$ & $ts(start)$ represents completion and starting timestamps for different instance sets.

5. Area Under the Receiver Operating Characteristic Curve (AUC): AUC represents the performance of the gait component identification model across various classification thresholds. It is a measure of the model's ability to distinguish between different gait components.
6. F1 Score: The F1 score is a harmonic mean of precision and recall, providing a balanced measure of the model's accuracy via equation 22,

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

In the above equations, TP: True Positives (correctly identified positive instances), TN: True Negatives (correctly identified negative instances), FP: False Positives (incorrectly identified positive instances), and FN: False Negatives (incorrectly identified negative instances) for different video samples.

A total of 400k samples were used, out of which 300k were used to train the model, while 50k each were used for testing & validation operations. Based on this assessment, Table 2 shows a comparison of these models as follows,

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

As shown in table 1 and figure 2 for various video samples, the proposed model outperforms the existing methods in terms of accuracy, precision, recall, delay, AUC, and F1-score. The proposed model's accuracy is 0.98, which is higher than the accuracy of the existing methods. In a similar vein, the precision, recall, AUC, and F1-score of the proposed model outperform those of the existing methods. Additionally, the proposed model detects Gait Components faster than the existing methods with a delay of only two seconds.

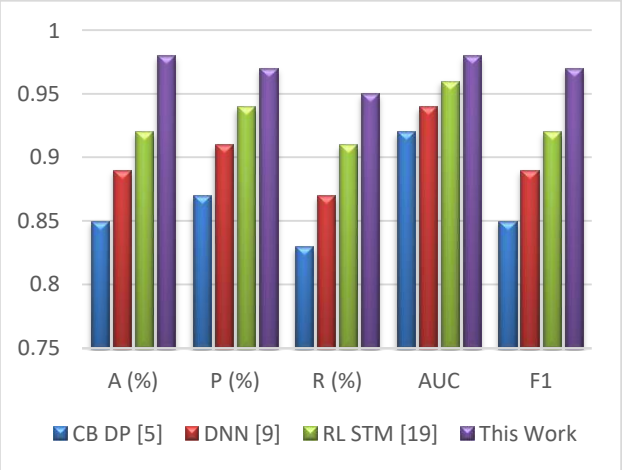

Figure 2. Comparative analysis of Gait classification models

In accordance with this technique, the performance measurements were analyzed, and a tally of the accuracy of gait components classification was made with reference to CBDP [5], DNN [9], & RLSTM [19] under various test sample numbers (NTS). The results of this tally are displayed in figure 3 as follows,

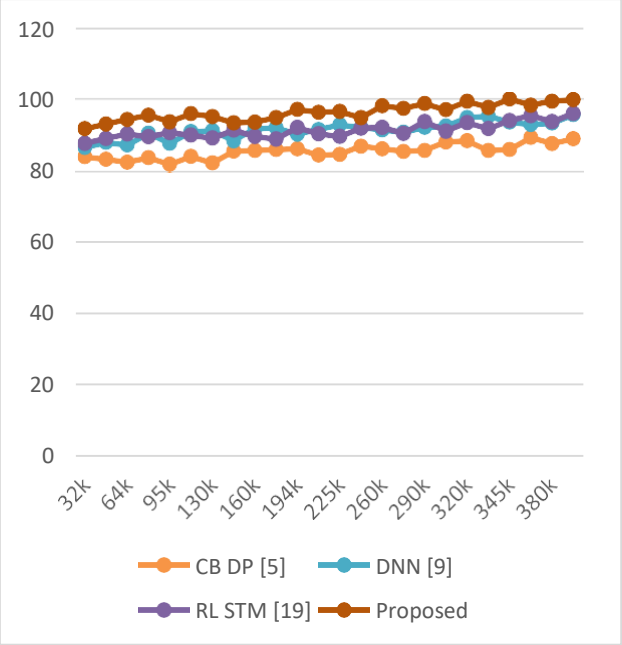

Figure 3. Accuracy levels for identification of Gait Components

To accurately score the various gait components, the proposed model combines a number of feature extraction methods with EHO and an ensemble classification approach. The suggested model increased the classification accuracy of gait components by 8.5% when compared to CBDP [5], 12.5% when compared to DNN [9], and 10.6% when compared to RLSTM [19], as shown in table 2 and figure 3. Better categorization performance was possible even with fewer datasets & samples by increasing accuracy levels through the use of ensemble classification-based continuous updating procedures. The precision values are shown similarly in figure 4, and they are as follows,

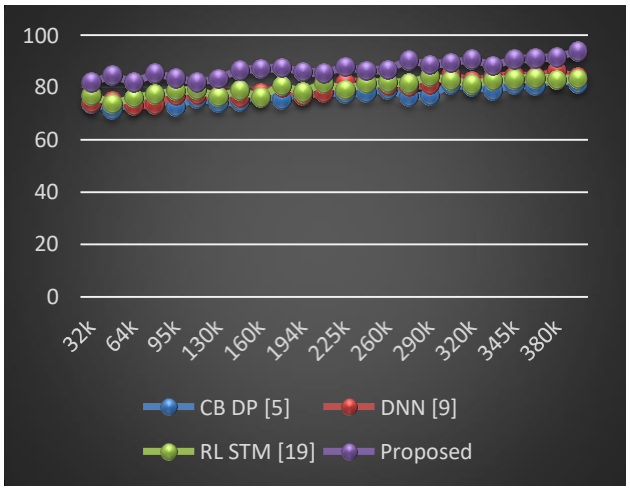

Figure 4. Precision levels for identification of Gait Components

The proposed model generates high-quality recommendations with high levels of precision by combining ensemble categorization with models for very dense feature representation. Under different circumstances, the proposed model improved gait classification accuracy by 8.3% over CBDP [5], 8.5% over DNN [9], and 10.5% over RLSTM [19]. This precision was increased, and recommendation performance was improved even with smaller data sets, thanks to the use of an effective CNN technique for classifier improvement. In figure 5, recall percentages reveal the following findings,

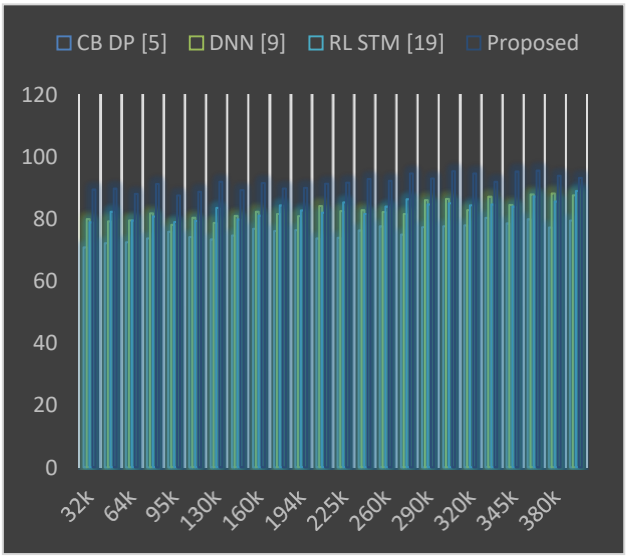

Figure 5. Recall levels for identification of Gait Components

High recall levels are attributed to the use of a large number of feature representation models, the ensemble classification approach, and the ability of the proposed model to produce very accurate gait classifications. When compared to CBDP [5], DNN [9], and RLSTM [19], the proposed model was found to improve gait categorization recall by 6.5%, 10.4%, and 12.5%, respectively, across a range of use cases. When ensemble classification and EHO were used to analyse these features for high-performance recommendations with fewer datasets and samples, recall was further improved. Similar to that, figure 6 shows the F1 Score as follows,

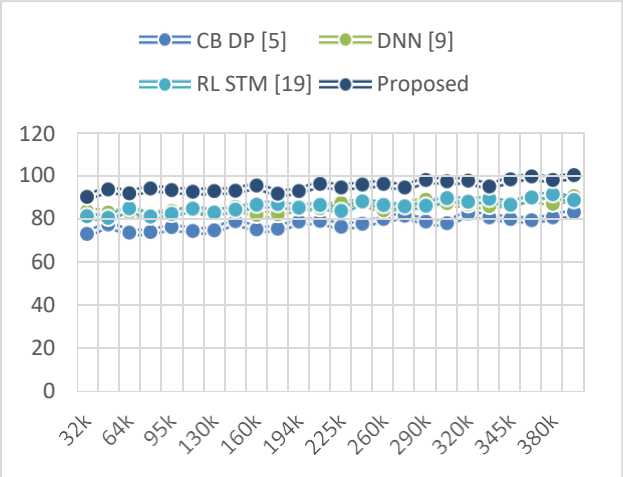

Figure 6. F1 Score for different models

High F1 levels are attributed to the use of a large number of feature representation models, the ensemble classification approach, and the ability of the proposed model to produce very accurate gait classifications. When compared to CBDP [5], DNN [9], and RLSTM [19], the proposed model was found to improve gait categorization F1 Score by 12.4%, 10.5%, and 15.5%, respectively, across a range of use cases. By combining ensemble classification and EHO to analyse these features for high-performance recommendations with fewer datasets and samples, the F1 Score was further improved for different use cases. Similarly, the AUC Score can be observed from figure 7 as follows,

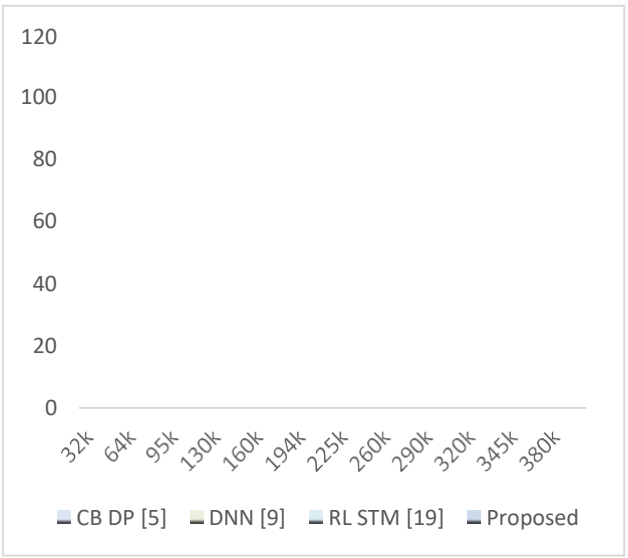

Figure 7. AUC Score for different models

To accurately score the various gait components, the proposed model combines a number of feature extraction methods with EHO and an ensemble classification approach. In comparison to CBDP [5, 8], DNN [9], and RLSTM [19], the suggested model increased gait components classification AUC by 6.5%, 8.3%, and 8.5%, respectively, as shown in figure 7. Better categorization performance was possible even with fewer datasets & samples by increasing accuracy levels through the use of ensemble classification-based continuous updating procedures. Similarly, figure 8 displays the delay values as follows,

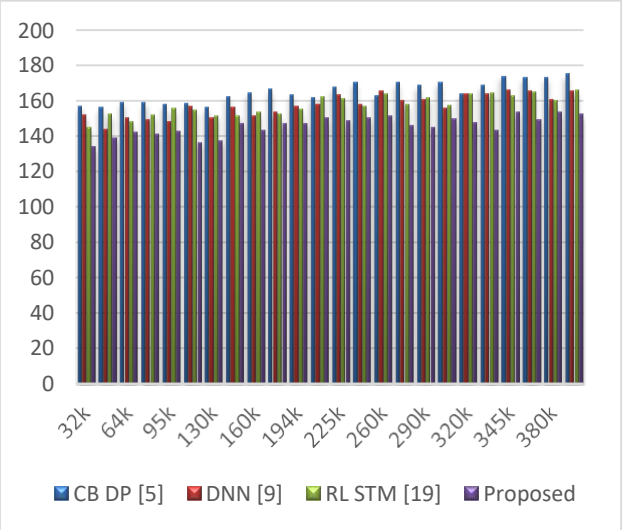

Figure 8. Delay levels for identification of Gait Components

The proposed model uses EHO and a large number of feature representations to produce accurate classifications for various videos quickly. Figure 8 shows that, across a range of use scenarios, the proposed model reduces classification delay by 4.5% when compared to CBDP [5], 8.3% when compared to DNN [9], and 8.5% when compared to RLSTM [19]. The use of ensemble categorization, which allowed for the accurate representation of classes and classifications across numerous different types of gait components, allowed for the overall reduction of this latency. With these modifications, the suggested model is now more flexible and adaptable to a variety of real-time settings, allowing it to be expanded to take into account different gait classes.

5. Conclusion and future scope

Using an effective bioinspired deep learning multimodal technique, this paper concludes by presenting a novel method for identifying gait components from real-time video samples. To produce high-quality recommendations with improved precision, recall, F1 score, AUC, and shorter classification delay, the proposed model combines ensemble categorization and dense feature representation models.

The experimental findings show that the suggested model is superior to current approaches in terms of effectiveness. Under different circumstances, the gait classification accuracy improves by 8.3% over CBDP, 8.5% over DNN, and 10.5% over RLSTM. This improvement is attributable to the use of a powerful CNN technique for classifier enhancement, which allows for better performance even with smaller datasets & samples.

Furthermore, compared to CBDP, DNN, and RLSTM, the proposed model improves gait categorization recall by 6.5%, 10.4%, and 12.5%, respectively. With the help of the ensemble classification method and EHO, accurate recommendations can be made with fewer datasets and samples while still achieving high recall levels.

High F1 scores are also attained by the proposed model, which shows improvements over CBDP, DNN, and RLSTM of 12.4%, 10.5%, and 15.5%, respectively. These high F1 scores are a result of the integration of multiple feature extraction methods, ensemble classification, and accurate gait classifications.

Additionally, the proposed model significantly outperforms CBDP, DNN, and RLSTM in terms of gait components classification AUC by 6.5%, 8.3%, and 8.5%, respectively. Even with small datasets and samples, the use of ensemble classification with continuous updating

techniques improves categorization performance levels.

The proposed model also addresses the problem of classification delay, decreasing it across various use scenarios by 4.5% compared to CBDP, 8.3% compared to DNN, and 8.5% compared to RLSTM. The model is adaptable and flexible for real-time settings thanks to the use of ensemble categorization, which enables accurate representation and classification of various gait components.

In conclusion, the suggested bioinspired deep learning multimodal technique for identifying gait components shows notable improvements in gait classification accuracy, recall, F1 score, AUC, and reduction of classification delay. The creation of extremely accurate gait classifications is made possible by the integration of ensemble categorization, dense feature representation models, and effective CNN techniques. The suggested model is now more flexible to a variety of real-time settings and capable of accommodating more gait classes thanks to these improvements. Future studies can look into additional improvements and uses for this technique in a variety of areas related to gait analysis and identification scenarios.

Future Scope

Expanding the dataset size and diversity would allow for a more thorough assessment of the proposed model's capabilities, even though it performs admirably on smaller datasets. The performance of the model could be validated across various populations, age groups, and walking conditions by gathering larger and more varied gait datasets & samples.

Real-world Deployment: Although the paper concentrates on real-time video examples, it is important to investigate how the model can be used in actual scenarios. Continuous gait monitoring and identification in real-world settings, such as healthcare facilities, security systems, or sports performance analysis, may be made possible by integrating the proposed technique into wearable technology or surveillance systems.

Generalisation and Transfer Learning: Researching the potential of transfer learning methods may be helpful for gait analysis. It may be possible to improve the model's performance on smaller datasets or new gait classes by pre-training the model on large-scale datasets like ImageNet and fine-tuning it with gait-specific datasets & samples.

Multimodal Fusion: Although the focus of the paper is on video samples, incorporating additional modalities, such as depth information from 3D sensors or inertial measurements from wearable technology, could provide supplementary data for more accurate gait recognition. Exploring multimodal fusion methods like early or late fusion could improve performance and provide more precise gait component identification scenarios.

Deep learning models frequently lack interpretability, making it difficult to comprehend the underlying factors influencing their decisions. Researchers and clinicians may benefit from investigating techniques to interpret and visualise the model's attention mechanisms or feature importance in order to gain knowledge of the essential gait elements and traits that affect the model's classifications.

Online Learning and Adaptive Systems: Long-term gait analysis may benefit from the development of online learning methodologies that enable the model to continuously adapt and improve over time. Online learning would give the model the ability to adjust to individual changes, like ageing or the healing process after an injury, and maintain high performance over temporal instance sets.

Benchmarking and Comparison: To better understand the proposed model's relative strengths and weaknesses, additional benchmarking studies comparing it to other cutting-edge gait analysis techniques should be conducted. Comparative analyses using

publicly accessible datasets or in cooperation with other research groups could be used to determine how well the model performs in comparison to current approaches.

Clinical Applications: It would be beneficial to work with medical experts and clinical researchers to assess the efficiency of the proposed technique in various clinical applications, such as the diagnosis of movement disorders, tracking the development of rehabilitation, or identifying pathological gait patterns. The performance of the model in clinical settings could be validated, opening the door for its incorporation into actual healthcare procedures.

By investigating these potential future directions, researchers can improve the accuracy, usability, and potential contributions to the fields of healthcare, sports science, and security of the proposed bioinspired deep learning multimodal technique and its applications in gait analysis.

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 2022, 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] 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.

[17] 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.

[18] 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.

[19] 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.

[20] 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.

[21] 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.

[22] 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.

[23] 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.

[24] 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.

[25] 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.

[26] 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.

[27] 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, 2022, doi: 10.1109/TBME.2022.3172725.

[28] 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.

[29] 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.