COBAL - A NOVEL DESIGN OF CNN BASED GAIT FEATURE EXTRACTION USING Bat-ELM FOR HUMAN TRACKING SYSTEM

RATNALA VENKATA SIVA HARISH¹, Dr. P. RAJESH KUMAR²
Research Scholar¹, Professor²
Department of Electronics and Communications Engineering
AU COLLEGE OF ENGINEERING (AUTONOMOUS)
rvsharish@gmail.com¹, rajeshauce@gmail.com²
VISAKHAPATNAM-530 003-ANDHRA PRADESH.

ABSTRACT:

Human gait recognition is one of the promising non-invasive biometric traits. Owing to its uniqueness, effectiveness and possession of rich discriminative features even in low resolution data, it is considered among the popular choice for authentication purposes. The sensitivity of the model is the challenging attribute in gait recognition. Also, it is difficult to discriminate the view angle and motion of human gait in varying illumination conditions. Addressing these issues, a hybrid Convolutional Neural Network (CNN) adapted evolutionary Bat Optimized Extreme Learning Machine approach is proposed. This hybrid structure works in two folds, such that the novel spatiotemporal features including spatial and morphological features are extracted by the CNN structure. The output of the CNN is incorporated in to the Bat optimized learning scheme as a single feed forward loop to attain classification of gait movements for recognition. Extensive experiments shows reduced error rate, increasing the rate of recognition than most of the existing hybrid algorithms.

Keywords: Human Gaits, Biometric traits, Convolutional Neural Networks, Bat Optimized Extreme Learning Machines.

INTRODUCTION

Over the classical biometric schemes, human gesture has a great impact for identification of individuals. Earlier research in gait recognition addresses the robustness and necessity of the biometric scheme. The major advantages of gait as biometric, is the recognition can be accomplished even when the subject of interest (human) is at a distance point. This is more ambitious in fields such as defense, surveillance and other forensic applications. Though the low resolution data is not a challenging constraint, the environmental parameters and motion of the subject and the angle views are more challenging in order to build a recognition model. From the previous literatures on gait recognition, it is evident that both spatial and temporal features are needed for optimal recognition.

In general, the gait based recognition is grouped in to two classes namely model based approach and model free approach [2]. Clustering of the static and dynamic gait features along with their body segments such as arms, legs, fingers and limbs are considered in the model based approach. Here, the model based approaches are always looped with the reference and test segments that are more time consuming and computationally complex. The
model free approach in turn considers only the contours of the silhouettes or else the motion of subject of interest. Thus the model free method is less complex and simple to attain the classification. Though the model free approach is preferred over the other, there is a possibility for scale unfit of the features in certain circumstances.

Numerous methods were employed for the recognition of human gait for enhancing the data. Fuzzy based methods [5-7] are developed to balance the lighting conditions and environmental instability issues. Also a Fuzzy K- nearest neighbor (FKNN) [11] based classification is employed for conveying the patterns to specific sets. The variable and initialization parameter K selection is crucial in this scheme. This often leads to the uncertainty condition in order to govern appropriate K values for any chosen pattern. The promising knowledge for gait recognition is sensed by researchers in machine learning and deep learning approaches. Addressing all the above issues, a novel hybrid bat optimized extreme learning machine along with CNN based approach for gait classification is formulated. Extreme Learning Machines are considered as the high accurate and Single feed forward neural network characterized by the auto-tuning of hidden neurons. The auto-tuning property leads to the instability of the network which results in the inaccurate recognition rate. Hence the integration of bio-inspired extreme learning machines in CNN well fits for massive feature handling for better classification of gaits.

The contributions of the article are listed below:

1) First, the extraction of real time spatiotemporal features from the decomposition of silhouettes using Convolutional Neural algorithm. The spatiotemporal features extracted serves as a unique dataset with stable feature constraints.

2) Secondly, an effective learning scheme for optimal classification by adapting BAT optimized Extreme Learning Machine approach which delivers high classification rate with expected outcomes even in challenging environments.

The organization of the paper is as follows: Section I presented the related works. Section-II deals with preliminary views of convolutional neural networks, extreme learning machines and bat optimizer. Section-III deals with the overall working of the proposed architecture COBAL which deals with feature extraction using Convolutional Networks , training by the newly formed bat optimized extreme learning machines. Results and comparative analysis were discussed in Section-IV.

SECTION -I

1. RELATED WORKS:

Many works have been carried out so far to accomplish an efficient gait recognition system. The robustness and intelligent classification is the most important aspect to be considered. With recent advancements in deep learning for recognition attempts, the gait recognition becomes more attractive among biometric classes. Jiwen Lu et.al [14] developed a model to identify and recognize the human gender. The silhouettes are extracted based on background subtraction algorithm. The approach made few unrealistic assumptions such as fixation of direction of motion of the human. Lijia Wang et.al [15] exploited a model considering Gait Flow Image (GFI) based on silhouette extraction, direction estimation, GFI generation and recognition. The model uses Lacus-kennada’s approach for GFI generation.
The loss in quality of silhouettes is possible for large datasets. Manjunatha Guru et al [16] proposed a four directional vector model to develop a feature fusion approach. This considers the horizontal, vertical, forward and backward diagonal directions. A gait energy image is developed for each image frame and the gradients are calculated. The gradient of the gait energy image is obtained using neighborhood gradient computation. The classification is accomplished with the Support Vector Machine (SVM) classifier algorithm. This work is tested with a public Meta data CASIA-B (Chinese Academy of Sciences) with prominent results. Hansu et.al [17] developed an ideology of generating stick figures of subject of interest. Low quality video data is sufficient for this approach. This 2-D stick contributes both the appearance of gait and their kinematic information. The SVM classifier is employed to classify the pattern models. Worapan et.al [18] proposed a sparse regression model for gait recognition. A view transformation model is developed by selecting the region of interest (ROI) without redundancy in data. This model also supports multi-view gait recognition. Dong Ming et.al [19] emphasized boundary tracking approach using wavelet transform. This generates human skeleton model to extract dynamic features. The classification is carried out with SVM algorithm. Denton et.al [20] proposed a tensor low rank based algorithm to fasten convolutional computation for real time recognitions.

SECTION -II

II. PRELIMINARY OVERVIEW:

2.1 CONVOLUTIONAL NEURAL NETWORKS:

Convolutional Neural Network (CNN) is one of the descriptive network model in deep learning technology. CNN is grounded on artificial neural networks. Figure.1 shows the architecture of conventional CNN algorithm. For various classification tasks, the back propagation algorithm is used to repeatedly upkeep or reduces the weight of corresponding features. The filters are used to auto extract features in the convolution layer to attain image feature extraction. Among the various deep learning algorithms, CNN is preferred to extract the kinematic characteristics of human gaits sequences because of its effectiveness in identification.

CNN has the ability to learn the features using its convolution property to the input data sequence with a high level of abstraction. The various layers of CNN include the
network layers, the pooling layers and the normalization layers. A feature map is generated based on the raw input sequence in the convolution layer. With the adaption of a set of filter and weights, the pooling layer reduces the size of the generated feature map from the convolution layer making it more prominent for analysis. The normalization layer output comprises of the silhouette features as a single dataset to be feed to the classifier for the gait recognition purposes.

2.2 EXTREME LEARNING MACHINES (ELM)

Extreme Learning Machine (ELM) structures are a class of advanced neural network to achieve effective classification and prediction tasks [27-29]. The input weights and biases are automated without manual tuning and the single hidden layer in it is suitable for both classification and recognition purposes. The neurons ‘N’ in the hidden layer acts as a sigmoid function in ELM as they are tuned automatically.

For a single-hidden layer of ELM, the model is described as

$$f_{y}(n) = \sum_{i=1}^{N} \beta_{i} h_{i}(n) = h(n)\beta$$  \hspace{1cm} (1)

Where, ‘n’ is the input

$\beta$ is the weight of output which is given as

$$\beta = [\beta_{1}, \beta_{2}, \ldots, \ldots, \beta_{N}]^{T}$$  \hspace{1cm} (2)

$H(n)$ denotes the hidden output layer which is given by

$$h(n) = [h_{1}(n), h_{2}(n), \ldots, \ldots, h_{N}(n)]$$  \hspace{1cm} (3)

The output vector also known as target vector is calculated from the hidden layers as follows

$$H = \begin{bmatrix} h(n_{1}) \\ h(n_{2}) \\ \vdots \\ h(n_{N}) \end{bmatrix}$$  \hspace{1cm} (4)

Generally ELM uses the minimal non-linear least square methods that is represented as

$$\beta' = H^{*}O = H^{T}(HH^{T})^{-1}O$$  \hspace{1cm} (5)

The $H$‘is termed as the inverse of $H$ known as Moore–Penrose generalized inverse parameter. Equation (5) can be rewritten as

$$\beta' = H^{T}(\frac{1}{\zeta}HH^{T})^{-1}O$$  \hspace{1cm} (6)

The output function of simple ELM can be expressed as

$$f_{y}(n) = h(n)\beta = h(n)H^{T}(\frac{1}{\zeta}HH^{T})^{-1}O$$  \hspace{1cm} (7)

The advantages of ELM algorithms include less error rate and optimized output [17, 18]. It is opted for its auto tuning property of weights and biases and optimum prediction values..
The major setback of the ELM algorithms is the non-optimum choice of biases and weights in the hidden layer which may induce redundant nodes which in turn affects the accuracy of detection. Therefore, to overcome the instability issues the bat optimization based ELM is chosen over the conventional ELMs in our proposed method.

2.3 BAT OPTIMIZERS:

The bio-inspired Bat echolocation algorithm for developing a learning scheme is introduced by Xin-She Yang [30]. The three guideline frames by them includes

1. All bats use echolocation to sense the distance of their opponent object, and they also ‘know’ the difference between food and other background huddles in some magical way.

2. Bats flies non-uniformly with the velocity $v_i$ at the position $x_i$ with a static frequency and noise to locate food. The wavelength of the produced pulse and the frequency of pulse emission are attuned automatically by the bats based on the closeness of the target.

3. Even though the loudness can differ in many ways, it is supposed that the loudness varies from least constant value $A_{min}$ to a high value (positive) $A_0$.

The movement of each bat is related with the velocity $v_i^t$ and initial distance $x_i^t$ with the ‘n’ total number of iterations in a dimensional space. Among all the bats, the best bat has to be preferred depends on the three rules which are stated above. The updated velocity $v_i^t$ and initial distance $x_i^t$ using the three rules are given below

$$f_i^t = f_{min} + (f_{max} - f_{min}) \beta$$

$$x_i^t = x_i^{t-1} + v_i^t$$

where $\beta \in (0,1)$ $f_{min}$ is the smallest frequency =0 and $f_{max}$ is the upper bound frequency which primarily depends on the problem defined. Each bat are originally assigned for the frequency between the $f_{min}$ and $f_{max}$. Thus bat calculation can be considered as a frequency regulation calculation to give a reasonable combination of investigation and exploitation. The emission rates and volume mainly contributes a method for automatic control and auto-zooming into the region of interest with promising solutions.

SECTION -III
3.1 PROPOSED MODELS

Figure.2 illustrates the proposed hybrid CNN based Bat optimized ELM architecture. The input human gesture videos captured are preprocessed in order to attain a clarified silhouette model. The preprocessing stage includes conversion of the captured RGB video frame into a grayscale image and then background subtraction is carried out. The mechanism, used for background subtraction in pre reference adaptive thresholding algorithm. This model initiates the background subtraction by excluding the human in the frame.

Let $I(u)$ be the input image of the frame ‘u’, ‘B’ corresponds to the background area of image frame ‘u’, the selected pixel value $J(u)$ is determined as

$$P[J(u)] = P[A] - P[I(u)]$$

The pixel value has to be converted into binary scale for adaptive thresholding of the data. Thus the thresholding criteria becomes

$$P = \begin{cases} 1 & \text{if } P/n \geq T \\ 0 & \text{if } P/n < T \end{cases}$$

The data after background subtraction is allowed to pass through a median filter, in which the noise attributes are removed. The feature modelling can be observed as a key part of identifying gait movements from the data. In general feature modelling comprises feature extraction and feature selection. The input data of a CNN is processed consecutively through its layers to obtain the feature maps with dissimilar features. There are 55 convolution kernels in the first layer of each branch, and equivalent kernel size is $27\times27$. In the next layer, the number of convolution kernels is unchanged, but we assume a different convolution kernel size, i.e. $13\times13$ and $6\times6$ correspondingly. At the later layer, the feature maps are concatenated and 64 feature maps are obtained. The clusters of silhouette databases are made with the particular multiple angle views. The dissimilar features were extracted from the
group of databases that are used as the initial training and testing datasets for the classification mechanism.

Figure 3 shows the input sequence and its corresponding silhouettes after background subtraction operation.

3.2 MULTIPLE FEATURE EXTRACTION:

Feature extraction is subsequent phase of the approach for human gait analysis. In this phase, a set of 7 novel gait features were extracted from the silhouette images at various angles and these images are regularized in order to reduce the effects initiated due to the variable distance from the capturing device. Table 2 shows the tuned hyper parameter values for the convolution layer and pooling layer required to achieve feature extraction. The features that are extracted for ELM based classification and identification are listed below.

a) **Frames/Gait Cycle**: This constitutes the number of frames per gait cycles. This measurement is used in determining the speed of human movements. The difference between heel strike (Hₙ) and toe-off frame (Tₙ) gives the measure of frames per gait cycle.

b) **Swing Ratio**: The swing of hands in the human body while walking is measured as the swing ratio. The ratio between the maximum torsal widths to the minimum torsal width provides the swing ratio. As there are multiple swings in a single database, the swing ratio is estimated by averaging the values at regular intervals.

c) **Cadence**: Cadence is the measure of number of steps walked by human per minute. This solely depends on the frame rate of the capturing device.

d) **Velocity**: The distance moved by the human per unit time is termed as velocity. Since velocity varies with time, the average velocity is considered as the feature parameter. The average velocity can be measured by calculating the distance between two successive points of heel contact to the opposite feet.

\[
\text{Average Velocity, } V = S_i \times C
\]  

(12)

Where \(S_i\) is the step length is considered based on the space between the consecutive points of heel contact of the opposite feet and \(C\) is the cadence.

e) **Step Length**: Step length is nothing but the toe to toe measure. For normalization purposes, step length values are divided by the height of the subject. The normalized step length is calculated by

\[
\text{Normalized Step Length} = \frac{S_i}{H}
\]  

(13)

Where \(H\) is the height of the subject.

f) **Foot Length**: The measure of distance from heel to toe is termed as foot length. For normalization purposes, the foot values are divided by the height of the subject.

\[
\text{Normalized Foot Length} = \frac{D_f}{H}
\]  

(14)

Where, \(H\) is the height of the person and \(D_f\) is the foot length.
g) **Cycle Time:** This feature is used to estimate the time to complete one cycle by taking the number of camera frames in gait cycle in to consideration.

h) **Mass Location Point:** The relative position of human body in accordance with center of ROI is termed as Mass Location Point. The normalized Mass location is considered by dividing the mass location point by the height of the subject.

**Table. 2 Hyper parameters tuned for feature extraction stage of the proposed CNN based Bat Optimized ELM Model**

| Layers          | Output Layer | Filter-Pooling layer |
|-----------------|--------------|----------------------|
| Input Layer     | 48x48x1      | 2X2                  |
| Convolution1    | 48x48x32     | 2X2                  |
| Max-pooling     | 24x24x32     | 2x2                  |
| Convolution2    | 24x24x64     | 2X2                  |
| Max-pooling     | 12x12x64     | 2x2                  |
| Convolution-3   | 12x12x64     | 3x3                  |
| Max-pooling     | 6x6x128      | 2x2                  |
| Fully Connected Output Layer | 07 | n/A |

Once the feature parameters are extracted from the convolutional networks, the proposed features are then to used for the training the networks and training network is formulated by the newly proposed bat optimized extreme learning machines.

**3.3 BAT OPTIMIZED EXTREME LEARNING MACHINES LAYERS:**

The working principle of ELM and BAT algorithms are discussed in the previous section II. The proposed network assimilates the BAT algorithm with Extreme Learning machine to achieve optimized weight and bias factors rather than the standalone ELM structures. The advantage of bat algorithm in ELM is it upsurges the total minima points in search path. The proposed model generates the accuracy as the fitness function for the entire dataset. If the gained accuracy is equal to the threshold accuracy, the output parameters are validated and the iterations continue till the fitness criterion is achieved. The complete parameters used for optimizing the ELM are given in Table 3.

**Table .3 Illustrations of the BAT parameters used for the optimization of ELM**

| Sl.No | Bat parameters    | Description                       |
|-------|-------------------|-----------------------------------|
| 01    | No of BATS        | 07(Initial)                       |
| 02    | Initial Velocity  | 15%                               |
| 03    | No of Iterations  | BAT optimized                     |
| 04    | Initial Loudness  | 0.9                               |
| 05    | Initial Pulse rate| 0.9                               |
| 06    | Minimum Frequency | 0 KHZ                             |
| 07    | Active threshold  | 99.0% Accuracy in Prediction/Classification |
The complete working model for the proposed architecture is depicted in figure 4.

| Slno | Pseudo Code for the Proposed Algorithm |
|------|--------------------------------------|
| 01   | Input : Input Images – Gait Video Sequences |
| 02   | Define the frequency f and velocity v |
| 03   | Initialize the pulse rate r and loudness A |
|      | N = Bat_optimized_Learning(v, f, A, r, N1) where N1 is the number of Neurons / layers in Extreme Learning Machines |
| 04   | Output : Recognition Rate |
| 05   | While n = 1 to max_iteration |
| 06   | F = Convolutional_networks (Silhouette Images) where F = no of features |
| 07   | /*Training Phase*/ |
| 08   | I = f_θ(N) = h(n)β = h(n)H^T(CHHT)^{-1}Q/ Output layers using the equation 7. |
| 09   | /*Testing Phase*/ |
| 10   | Determine the Value of I |
| 11   | End |
| 12   | End |

SECTION -IV

4.1 EXPERIMENTAL ANALYSIS:

The experiment is carried out to validate the effectiveness of the hybrid CNN based Bat ELM model. The randomly chosen ‘n’ gait image sequences per subject of interest from the CASIA-B multi view gait database [16] to form the training sets and utilize the remaining to generate the testing set. This database was developed with 11 capturing device around the left hand side of the subject when he/she was walking, and the angle between two adjacent directions is 18. That is, from left to right, the viewing angles are 0, 18, 36, 54, 72, 90, 108, 126, 144, 162, and 180, respectively. Human gait data of 124 subjects of interest were all captured together, and among the subjects 93 were men and 31 were women. All the subject has six categorizations walking in a normal condition and thus there are a total 6 × 11 × 124 = 8184 of normal gait sequences in the database. It is a huge gait database in both the number of subjects and in the number of view angles. The pyTORCH is used for implementation of the proposed hybrid algorithms in Intel i7 CPU with 2TB HDD and 8 GB RAM.

The various features captured with the different viewing angles are used as the training datasets and testing data sets. The features with multiple view angles such as 0, 18, 36, 54, 72, 90, 108, 126, 144, 162, and 180, respectively has been recorded. The features which are extracted in different view angles are listed in table 4.
### Table. 4 Features being extracted for the View Angle of 90 degree

| Image Type | Step Length (cm) | Foot Length (cm) | No of Frames | Cadence | Swing Ratio | Distance(cm) | Velocity (cm/sec) |
|------------|------------------|------------------|--------------|---------|-------------|--------------|-------------------|
| 1          | 5.08             | 1.1              | 24           | 6       | 4.7:3.2     | 27.9         | 6.1               |
| 2          | 5.08             | 1.1              | 28           | 5       | 4.7:3.2     | 27.9         | 6.3               |
| 3          | 5.34             | 1.1              | 30           | 5       | 4.6:3.1     | 27.9         | 6.0               |
| 4          | 5.84             | 1.1              | 39           | 4       | 4.5:3.1     | 27.9         | 6.5               |
| 5          | 5.08             | 1.1              | 32           | 6       | 4.8:3.3     | 27.9         | 4.4               |
| 6          | 5.58             | 1.2              | 30           | 5       | 5.3:3.6     | 27.9         | 5.0               |
| 7          | 5.34             | 1.2              | 33           | 6       | 5.4:3.7     | 27.9         | 5.2               |
| 8          | 5.58             | 1.2              | 30           | 5       | 5.3:3.6     | 27.9         | 5.0               |
| 9          | 5.34             | 1.2              | 34           | 5       | 5.3:3.7     | 27.9         | 5.2               |
| 10         | 5.84             | 1.2              | 32           | 5       | 5.4:3.5     | 27.9         | 5.2               |
| 11         | 5.48             | 1.4              | 39           | 5       | 5.4:3.7     | 27.9         | 4.9               |
| 12         | 6.04             | 1.4              | 26           | 4       | 5.4:3.7     | 27.9         | 6.0               |
| 13         | 6.23             | 1.4              | 24           | 3       | 5.7:3.9     | 27.9         | 6.2               |
| 14         | 5.84             | 1.4              | 30           | 5       | 5.7:3.9     | 27.9         | 5.9               |
| 15         | 6.07             | 1.4              | 26           | 4       | 5.3:3.4     | 27.9         | 6.2               |
| 16         | 4.83             | 1.0              | 40           | 6       | 5.1:3.2     | 27.9         | 4.5               |
| 17         | 4.57             | 1.0              | 42           | 6       | 5.1:3.0     | 27.9         | 4.4               |
| 18         | 4.83             | 1.0              | 43           | 7       | 5.3:3.2     | 27.9         | 4.0               |
| 19         | 5.08             | 1.0              | 38           | 5       | 5.5:3.3     | 27.9         | 5.0               |
| 20         | 4.83             | 1.0              | 42           | 6       | 5.1:3.0     | 27.9         | 4.4               |
| 21         | 5.59             | 1.2              | 35           | 5       | 5.3:3.6     | 27.9         | 5.3               |
| 22         | 6.07             | 1.2              | 29           | 5       | 5.3:3.7     | 27.9         | 6.0               |
| 23         | 6.23             | 1.2              | 24           | 4       | 5.4:3.2     | 27.9         | 6.3               |
| 24         | 5.84             | 1.2              | 30           | 5       | 5.3:3.6     | 27.9         | 5.9               |
| 25         | 6.07             | 1.2              | 26           | 4       | 5.7:3.9     | 27.9         | 6.2               |
| 26         | 5.84             | 1.45             | 39           | 5       | 5.4:3.9     | 27.9         | 4.9               |
| 27         | 6.07             | 1.45             | 26           | 4       | 5.4:3.6     | 27.9         | 6.0               |
| 28         | 6.23             | 1.45             | 24           | 3       | 5.7:3.9     | 27.9         | 6.2               |
| 29         | 5.84             | 1.45             | 30           | 5       | 5.7:3.9     | 27.9         | 5.9               |
| 30         | 6.07             | 1.45             | 26           | 4       | 5.7:3.9     | 27.9         | 6.2               |
| 31         | 5.33             | 1.2              | 39           | 5       | 4.9:3.1     | 27.9         | 4.9               |
| 32         | 5.33             | 1.2              | 38           | 5       | 4.9:3.1     | 27.9         | 5.0               |
| 33         | 6.23             | 1.2              | 35           | 5       | 5.0:3.2     | 27.9         | 4.8               |
| 34         | 5.59             | 1.2              | 30           | 5       | 4.9:3.2     | 27.9         | 5.9               |
| 35         | 6.07             | 1.2              | 26           | 4       | 4.9:3.1     | 27.9         | 6.2               |
| 36         | 5.84             | 1.2              | 39           | 5       | 5.4:3.7     | 27.9         | 4.9               |
| 37         | 6.07             | 1.4              | 26           | 4       | 5.4:3.7     | 27.9         | 6.0               |
| 38         | 6.23             | 1.4              | 24           | 3       | 5.7:3.9     | 27.9         | 6.2               |
| 39         | 5.84             | 1.3              | 30           | 5       | 5.7:3.9     | 27.9         | 5.9               |
| 40         | 6.07             | 1.3              | 26           | 4       | 5.7:3.9     | 27.9         | 6.2               |
| 41         | 5.08             | 1.1              | 24           | 6       | 4.7:3.2     | 27.9         | 4.9               |
| 42         | 5.08             | 1.1              | 28           | 5       | 4.7:3.2     | 27.9         | 4.7               |
| 43         | 5.34             | 1.1              | 36           | 5       | 4.7:3.2     | 27.9         | 4.4               |
| 44         | 5.84             | 1.1              | 39           | 4       | 4.7:3.2     | 27.9         | 5.5               |
Likewise, the other features are also extracted for multiple view angles as said above in which features along with the angles are considered altogether as a complete single datasets for evaluation. Almost 3000 datasets were collected in which the 70% were used for generation of training sets and the remaining were used for testing process.

### 4.2 PERFORMANCE EVALUATION

The evaluation metrics for the proposed architecture are discussed as follows. The important performance attributes in recognition algorithms includes Accuracy of the system, Sensitivity and Specificity.

\[
\text{Accuracy} = \frac{\text{DR}}{\text{TNI}} \times 100
\]

\[
\text{Sensitivity} = \frac{\text{TP}}{\text{TP} + \text{FN}} \times 100
\]

\[
\text{Specificity} = \frac{\text{TN}}{\text{TP} + \text{FN}} \times 100
\]

Where TP and TN Represents True Positive and True Negative values and DR & TNI Represents Number of Detected Results and Total number of Iterations. Also we have used 10 cross validated matrix for validating of the different features. To prove superior performance of the proposed algorithm, we have compared the performance with the hybrid models such as CNN+RNN[36], DR+NN[38],3DCNN[35] and other non-hybrid models such as Multi layer perceptron’s [37] , Support vector machines(SVM)[34] and Extreme Learning Machines(ELM).

![Figure 4](image1.png)

**Figure 4** Training and testing performance of the Proposed Hybrid Algorithms
From the Figure 4, it is clear that the both training and testing accuracy has reached the peak value of 99% when the hidden neurons are 200 which are optimized by the proposed learning algorithms. This makes it suitable for the algorithm to produce the constant performance with the different multiple angle of features. The confusion matrix for the proposed algorithm is depicted in figure below.

| Angles | 0   | 18  | 36  | 54  | 72  | 90  | 108 | 126 | 144 | 162 | 180 |
|--------|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|-----|
| 0      | 2079| 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 18     | 0   | 2079| 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 36     | 0   | 0   | 2079| 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 54     | 0   | 0   | 0   | 2079| 0   | 0   | 0   | 0   | 0   | 0   | 0   |
| 72     | 0   | 0   | 0   | 0   | 2079| 0   | 0   | 0   | 0   | 0   | 0   |
| 90     | 0   | 0   | 0   | 0   | 0   | 2079| 0   | 0   | 0   | 0   | 0   |
| 108    | 0   | 0   | 0   | 0   | 0   | 0   | 2079| 0   | 0   | 0   | 0   |
| 126    | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 2079| 0   | 0   | 0   |
| 144    | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 2079| 0   | 0   |
| 162    | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 2079| 0   |
| 180    | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 0   | 2079|

Figure 5 Confusion Matrix Illustrating the True Positive Values for the Proposed algorithm

The integration of the Proposed Bat optimized ELM in CNN has proved the constant performance with the high recognition rate at the multiple angle views. Moreover as discussed in the previous section, the proposed algorithm is compared with the other algorithms with the analysis are given below.

Figure 6 Comparative Analysis of Prediction Accuracy of human tracking for the different hybrid algorithms.
Figure 7 Comparative Analysis of Sensitivity Analysis for human tracking for the different hybrid algorithms.

Figure 8 Comparative Analysis of Specificity Analysis of human tracking for the different hybrid algorithms.

Figure 6-8 discusses about the different performance metrics analysis for the different hybrid algorithms and non-hybrid models. It has been noticed from the Figure 6, rate recognition has drastically reduced when the viewing angle is increased whereas the proposed COBAL architecture has maintained the constant recognition rate of 99% in multiple viewing angles due to the optimized learning structure incorporated with the convolutional networks. These proposed deep learning models can be used for the human tracking systems when surveillance plays important applications.

SECTION-V

5. CONCLUSION

The proposed novel human gait recognition model is accomplished by extracting spatial and morphological feature parameters. The efficiency of CNN in feature extraction upholds the model for accurate feature selection process. Likely the ELM adapted with the bio-inspired bat algorithm shows prominent accuracy in classification of the feature gaits. The Bat algorithm maintained the stability of the model hence, that the recognition is robust and accurate. CASIA-B datasets were used for analysis and the gait features at different views of
angles were evaluated in which the proposed hybrid learning method has outperformed other existing hybrid models and non-hybrid models. Future work attempts in intelligent feature selection methods at real-time for the extrema to reduce the complexity of the learning metrics and also formation of the new hybrid deep learning models to increase the accuracy when interfaced with noisy input videos.

REFERENCES:

[1]. Gafurov, D.: ‘A survey of biometric gait recognition: approaches, security and challenges’. IEEE Int. Conf. on Biometrics: Theory, Applications and Systems, Norwegian, 2007, pp. 1–12
[2]. Zheng, L., Zhang, Z., Wu, Q., et al.: ‘Enhancing person re-identification by integrating gait biometric’, Neurocomputing, 2015, 168, pp. 1144–1156 [7] Maodi, H., Wang, Y., Zhang, Z., et al.: ‘Incremental learning for video-based gait recognition with LBP flow’, IEEE Trans. Cybern., 2013, 43, (1), pp. 77–89
[3]. Kusakunniran, W., Wu, Q., Li, H., et al.: ‘Multiple views gait recognition using view transformation model based on optimized gait energy image’. 12th IEEE Int. Conf. on Computer Vision, 2009, pp. 1058–1064
[4]. Lu, J., Wang, G., Moulin, P.: ‘Human identity and gender recognition from gait sequences with arbitrary walking directions’, IEEE Trans. Inf. Forensics Sec., 2014, 9, (1), pp. 51–61
[5]. Lu, J., Zhang, E.: ‘Gait recognition for human identification based on ICA and fuzzy SVM through multiple views fusion’, Pattern Recognition. Lett., 2007, 28, (16), pp. 2401–2411
[6]. Liu, N., Tan, Y.: ‘View invariant gait recognition’. IEEE Int. Conf. on Acoustics Speech and Signal Processing, March 2010, pp. 1410–1413
[7]. Kim, Y., Han, J.: ‘Fuzzy KNN algorithm using modified K-selection’. IEEE Int. Conf. on Fuzzy Systems, Japan, 1995, pp. 1673–1680
[8]. Rhee, F., Hwang, C.: ‘An interval type-2 fuzzy K-nearest neighbor’. Int. Conf. on Fuzzy Systems, USA, 2003, pp. 802–807
[9]. Qilian, L., Mendel, J.: ‘Interval type-2 fuzzy logic system: theory and design’, IEEE Trans. Fuzzy Syst., 2000, 8, (5), pp. 535–550
[10]. Venkata, G., Jilani, S.: ‘Fuzzy principal component analysis based gait recognition’, J. Computer. Sci. Inf. Technol., 2012, 3, (3), pp. 4015–4020
[11]. Jiwen Lu, Gang Wang, Pierre Moulin, “Human Identity and Gender Recognition From Gait Sequences With Arbitrary WalkingDirections,” IEEE Transactions on Information Forensics and Security, Vol.9, pp.51-61, 2014.
[12]. Lijia Wang, Songmin Jia, Xiuzhi Li, Shuang Wang, “Human gait recognition based on gait flow image considering walking direction,” IEEE International Conference on Mechatronics and Automation, pp.1990–1995, 2012.
[13]. V G Manjunatha Guru, V N Kamalesh, R Dinesh, “Human gait recognition using four directional variations of gradient gait energy image,” International Conference on Computing, Communication and Automation (ICCCA), pp.1367-1371, 2016.

[14]. Nirattaya Khamsemanan, Cholwich Nattee, Nitchan Jianwattanapaisarn, “Human Identification From Freestyle Walks Using Posture-Based Gait Feature,” IEEE Transactions on Information Forensics and Security, Vol.13, Issue 1, pp.119-128, 2018.

[15]. Saad M. Darwish, “Design of adaptive biometric gait recognition algorithm with free walking directions,” IET Biometrics, Vol.6, Issue 2, pp.53-60, 2017.

[16]. Makoto Shinzaki, Yumi Iwashita, Ryo Kurazume, Koichi Ogawara, “Gait-based person identification method using shadow biometrics for robustness to changes in the walking direction,” IEEE Winter Conference on Applications of Computer Vision, pp.670-677, 2015.

[17]. Ihn-Sik Weon, Soon-Geul Lee, “Recognition of User’s Gait Intention for a Walking Assist System Using Pressure-Sensing Handlebars,” IEEE International Conference on Robotic Computing (IRC), pp.326-329, 2017.

[18]. Fang Wang, Marjorie Skubic, Marilyn Rantz, Paul E. Cuddihy, “Quantitative Gait Measurement With Pulse-Doppler Radar for Passive In-Home Gait Assessment,” IEEE Transactions on Biomedical Engineering, Vol.61, pp.2434-2443, 2014.

[19]. Han Su, Feng-Gang Huang, “Markerless human gait recognition by shape and motion analysis,” International Conference on Intelligent Sensing and Information Processing, pp.161-165, 2005.

[20]. Tanmay T. Verlekar, Paulo L. Correia, Luis D. Soares, “View-invariant gait recognition system using a gait energy image decomposition method,” IET Biometrics, Vol.6, Issue 4, pp.299-306, 2017.

[21]. Liang Wang, Tieniu Tan, Huazhong Ning, Weiming Hu, “Silhouette analysis-based gait recognition for human identification,” IEEE Transactions on Pattern Analysis and Machine Intelligence, Vol.25, pp.1505-1518, 2003.

[22]. Worapan Kusakunniran, Qiang Wu, Hongdong Li, Jian Zhang, “Multiple views gait recognition using View Transformation Model based on optimiz
ed GaitEnergy Image.” IEEE International Conference on Computer Vision Workshops, ICCV Workshops, pp. 1058-1064, 2009.

[23]. WorapanKusakunniran, Qiang Wu, Jian Zhang, Hongdong Li, “Gait Recognition Under Various Viewing Angles Based on Correlated Motion Regression,” IEEE Transactions on Circuits and Systems for Video Technology, Vol. 22, Issue. 6, pp. 966-980, 2012.

[24]. Dong Ming, Cong Zhang, Yanru Bai, Baikun Wan, Yong Hu, K.D.K. Luk, “Gait recognition based on multiple views fusion of wavelet descriptor and human skeleton model,” IEEE International Conference on Virtual Environments, Human-Computer Interfaces and Measurements Systems, pp. 246-249, 2009.

[25]. R. F. Mansour, “A Robust Approach to Multiple Views Gait Recognition Based on Motion Contours Analysis,” WIAR 2012; National Workshop on Information Assurance Research, pp. 1-7, 2012.

[26]. Rohit, K., Pathak, V.: ‘Gait recognition based on energy deviation image using fuzzy component analysis’, J. Innov. Manage. Technol., 2013, 4, (1), pp. 43–46

[27]. Yang X-S. A New Metaheuristic Bat-Inspired Algorithm. Stud ComputIntell (NICSO) 2010; 284: 65–74

[28]. Wang B, Huang S, Qiu J, et al. Parallel online sequential extreme learning machine based on MapReduce. Neurocomputing 2015; 149: 224–32.

[29]. Lu S, Lu Z. A pathological brain detection system based on kernel-based ELM. Multimed Tool Appl 2016; 1–14.

[30]. Yang X-S. A New Metaheuristic Bat-Inspired Algorithm. Stud ComputIntell (NICSO) 2010; 284: 65–74

[31]. Javier, C., Josep, A., Fernando, D.: ‘Robust normalization of silhouettes for recognition applications’, J. Pattern Recognit., 2004, 25, (5), pp. 591–601

[32]. Wang, L., Tan, T., Hu, W.: ‘Silhouette analysis-based gait recognition for human identification’, IEEE Trans. Pattern Anal. Mach. Intell., 2003, 25, (12), pp. 1505–1518

[33]. Qinyong, M., Shenkang, W., Jianfeng, Q.: ‘Gait recognition at a distance based on energy deviation image’. 1st Int. Conf. on Bioinformatics and Biomedical Engineering, China, 2007, pp. 621–624

[34]. T. Whytock, A. Belyaev, and N.M. Robertson, “Dynamic distance-based shape features for gait recognition,” J. Math. Imaging Vis., vol. 50, pp. 314–326, 2014

[35]. T. Wolf, M. Babaei, and G. Rigoll, “Multi-view gait recognition using 3D convolutional neural networks,” in Proc. IEEE Int. Conf. Image Process. (ICIP), Sep. 2016, pp. 4165–4169. doi: 10.1109/ICIP.2016.7533144.

[36]. D. Ravì, C. Wong, B. Lo, and G.-Z. Yang, “A deep learning approach to on-node sensor data analytics for mobile or wearable devices,” IEEE J. Biomed. Health Inform., vol. 21, no. 1, pp. 56–64, Jan. 2017. doi: 10.1109/JBHI.2016.2633287
[37]. M. Gadaleta, L. Merelli, and M. Rossi, “Human authentication from ankle motion data using convolutional neural networks,” in Proc. IEEE Stat. Signal Process. Workshop (SSP), Jun. 2016, pp. 1–5. doi: 10.1109/SSP.2016.755181

[38]. A. Murad and J.-Y. Pyun, “Deep recurrent neural networks for human activity recognition,” Sensors, vol. 17, no. 11, p. 2556, 2017. doi: 10.3390/s1711255