The Systematic Review On Gait Analysis: Trends And Developments

Nithyakani P1, M Ferni Ukrit2
Assistant Professor, School of Computing, SRMIST.
1nithykp@srmist.edu.in
2ferniukm@srmist.edu.in

ABSTRACT - The research on gait analysis has a drastic growth in recent years for various applications like authentication, video surveillance, health monitoring system and animation. Gait recognition is an emerging intelligent system because of its recognition at a distance and it works well with low resolution videos. This paper provides the various gait features extracted in model based approach and model free approach. This paper describes the survey of gait recognition techniques which include the four major steps such as data acquisition, preprocessing, feature extraction and classification. Descriptions about wearable sensor, Vision base and pathological database is summarized. In recent years, performance of human recognition has been impressively enhanced by deep architectures. This paper offers an up-to-date survey of deep architectures on gait recognition, mainly focusing on the performance of convolutional neural network combined with other architectures. Furthermore, general challenges faced in gait recognition and directions for future research are discussed.

1. INTRODUCTION

Human analysis and identification has been an attractive technology in many fields. In the digital imaging, a human can be analyzed by his/her unique feature of face, iris, sweat, hair, periocular region of eyes, Gait, smell, finger and palm[18]. Gait defines the style of walking. Human Gait analysis is a study of human movement to analyze the style of walking and its related features which are applied in the field of biometrics, surveillance, diagnosis of gait disease, therapy and rehabilitation, etc. Gait recognition can be performed through involvement and non-involvement of a person to the recognition system. In human involvement approach, person is directly contacted to the recognition system with the help of different sensors, accelerometer or monitoring devices. Non-involvement approach uses camera at a distance to identify a person identity with or without their knowledge based on its application. Gait analysis is prevalent than other features since the features can be extraction from low resolution and without the cooperation of any individual.

“On the Gait of Animals” is the first book to report the gait analysis written by Aristotle in 350BC[1]. In 1680, Giovanni Alfonso Borelli made the extensive study on the animal locomotion mechanism. In 1890, Christian Wilhelm Braune and Otto Fischer performed the investigation of gait activity over human and published a series of papers. They have extended their research on the center of gravity, moments of inertia of the human body and its segments. In 1880s Eadweard Muybridge and Étienne-Jules Marey developed the animated photography to capture the several movement phases in a photographic surface. Muybridge used multiple cameras to record the locomotion and presented through his public lectures and demonstrations. In 1970s, video camera systems were available to study the several human pathological conditions. Human individuals were recognized with their gait style in 1970s. Later gait analysis was impressed by the forensic and animation experts [3-5]. Advancements in technologies of artificial intelligence and computing, Gait recognition has become the most popular data driven and artificial surveillance tool which can recognize the human at a distance. This survey aims to highlight the technologies developed for gait analysis. Section II provides the application of gait recognition in various fields. Section III describes about the human gait features considered for model based and appearance based recognition. Section IV focuses on the role of machine learning algorithms in recognizing human. Section V presents the description and
comparison of recently implemented deep architecture which gives the high accuracy. Section VI discusses about future directions in the field of human gait analysis and Section VII concludes with the challenges and the future work.

2. APPLICATIONS OF GAIT RECOGNITION

2.1. Application-Based Systems
Authentication is necessary to avoid violation in secured protected areas, home, public security, etc. Authentication processes[3] are carried out in three methodologies such as knowledge based (credentials, PIN, MPIN), object or token based (E.g. bank card, identity card etc.) and biometric based. Knowledge based authentication process faces memorability issues specifically robust passwords. Security breach arises when the password is reused in many services. Object based method usually combined with knowledge based method for accessing the services provided with physical keys. Biometric authentication is used to identify a person by his/her unique features like fingerprints, face, and periocular region of eye, retina, palm, sweet, body odor and gait. These features are unique in all subjects and also remain stable for a specific period. Behavioural features are difficult to imitate. Gait recognition method includes computer vision techniques, signal analysis, motion estimation, processing the video data, and ubiquitous applications.

2.2. Pattern Recognition
Biometrics is divided into physical and behavioral features. Physical features are the pattern of height, structure of face, retina, eye etc. Physical features has universality, unique features for a long term period. Behavioural features includes the way of walking, speaking, singing, movement etc. Though behavioral features varies on emotional factor, difficult to forge or imitate the individuality. A set of biometric images are recorded for verifying the similarity of captured image. Due to the recognition accuracy in various capturing modes, gait recognition has the highest popularity among the other biometric approach.

2.3. Life And Medical Sciences
Gait recognition helps to identify the quality of walking of an individual with several common gait disease such as parkinson’s disease, multiple sclerosis, cardiopathies, post-stroke Hemiplegic gait, spastic diplegic, neuropathic, myopathic, choreiform, ataxic (cerebellar) and sensory [9-10]. Study of the gait sequences helps for early diagnoses of their musculoskeletal and neurological complications to provide the best treatment[62]. Kinematic and kinetic data are extracted from gait analysis helps the clinicians to interpret, assess, and evaluate the individual gait disorder. Data collection is performed by videotaping the individual before the examination, examining physically to analyze the muscle strength and tone, motion of joint, abnormalities in bone, stubborn muscular contractions, knee and ankle width, distance between right and left sacroiliac spine, etc.

2.4. Animation
Animation of human gait is focused by the fields of biomechanics, signal processing, kinematics, dynamics and robotics[6-7,11]. Interest in human gait animation is growing rapidly in the Computer Animation Community despite of the complex techniques. The first computer animation approach for human was developed in 1970s which depicts the early sequential pictorial representation of series. Later Humanoid animation (H-Anim) standard was approved in 1990s and greatly influenced by the research experts of gaming, movies, virtual avatars, graphics, ergonomics and simulations industry. The challenge in human animation is to simulate the natural motion style into the virtual characteristics. Animation is performed by the motion sequence extraction of skeleton representation of human connected hinges points like shoulder, neck, elbow, waist, knee and ankle. Higher quality is obtained by modelling the skeleton with surfaces. Methods of human gait animation are knowledge based kinematic animation, dynamic animation[61] and interactive synthetic walking motion data.
3. Gait Recognition Features

Gait is a synchronized movement of bones, muscles and nerves of hip, knee and ankle joints rotations. The cyclic nature of the walking style provides the periodic leg movement. Gait of a person is generally represented as a walk cycle describes in the Figure 1 which has two phases: stance phase and swing phase; gait cycle has eight stages \cite{2} as (i) initial contact of the foot or toe. (ii) Loading response as placing the entire foot over the floor with the limb segments reaction. (iii) Mid-stance as the intermediate support. (iv) Terminal stance is denoted as heel off. V) Pre-swing is the initialization of knee movement. VI) Initial Swing . Parameters extracted from gait sequence are step length, stride length, cadence, speed, dynamic base, progression line, foot angle, hip angle and squat performance. The measurements extracted were analyzed, interpreted to recognize the individual through the techniques like temporal or spatial, kinematics, Marker less gait capture, pressure analyzer, kinetics and dynamic electromyography.

![GAIT CYCLE](image)

Fig. 1. A single Gait cycle is also termed as stride. Stance phase begins with the heel strike and toe off the same foot. Average seconds of stance phase is 0.64s. Swing phase begins with another foot heel strike and toe off. Average seconds of swing phase is 0.40s. A stride can also be represented with 8 stages: Initial contact, loading response, mid stance, terminal stance, pre-swing, initial swing, mid swing, terminal swing.

Human gait recognition system depends on the video has two major categories as model free gait recognition and model based gait recognition. Model based approaches extract the structural pattern of model from human body using various methodologies such as pendulum, leg inclination, stride, trajectories, joint angles, 2D and 3D temporal models \cite{19-22} with their length, width and position dimensions. It helps to track the individual with set of mathematical parameters and its quantifiable relationships. Model based approach used in handling self-occlusions. Issues faced in Model based approach were viewpoint invariant, physiological changes such as age, weight, height, walking speed, etc., and external factors like clothes, foot wear, load, etc., and environmental factors. In model based approach kinematics track joints of an individual are head, shoulder, spine, foot, knee, ankle, thumb, neck, hand, etc shown in the Figure 2a. It also defines the dynamics or kinematics of the body movements in the terms of gait cycle time, stride length, Fourier magnitude, footstep, ground reaction force, pressure mat features, etc. Model based gait recognition utilizes spatio temporal volume, linear regression analysis, trigonometric-polynomial interpolant functions, biped locomotion feature, Principal component Analysis, canonical analysis, 1NN, KNN, Hidden markov Models. Accuracy of the model based approach depends on the silhouette quality.
Fig. 2. Gait measurements and types of wearable sensors

a) Measurements of human body used for feature extraction
b) Wearable sensors used to extract gait features

The other categories of gait recognition approaches were image and video based approach, wearable sensors, floor sensor and radar sensor[26]. Image and video based approach which involves remote capturing of human movement. Image and video processing capture the gait sequence of an individual with the help of several optic sensors like digital camera, RGB, Low RGB-D Microsoft Kinect, laser range scanners, infrared sensors and time to flight cameras. Table 1 presents the description about various gait datasets collected using camera and wearable sensors. Wearable sensors which are attached to the human body measures movement time, calculate joint or inclination angle, walking speed, step or stride length, segment position, etc. Wearable sensors used for human recognition are accelerometer, gyroscope, force sensors, strain gauges, inclinometers, goniometer, Magnetoresistive sensor, electromagnetic tracking system, fabric sensor, and Electromyography sensor (Figure 2b). Sensors can be combined with the respect to application of gait analysis. Floor sensor based approach measures pressure, step length, etc. with the help of pressure under floor mat and floor accelerometer. In gait recognition, various factors can modify the gait pattern permanently or transiently. External factors are terrain, footwear, clothing and cargo; Internal factors are sex, weight, height, age, physique, etc.

3.1. Model Based Features
Model based approach model[19] generate the structure for human body and the arbitrary 2D views from the 3D model using the techniques like pendulum, cadence, stride, 2d Stick, kinematic features, angular trajectories etc. 3D model uses more than two calibrated cameras to capture the gait sequences and to avoid the problems of variations in surfaces, self-occlusion and fixed viewpoint. Structural method uses geometrical patterns of time variations and motion variations of gait sequence. This method has overcome the problem of variation in light, background, view and segmentation. The features extracted in model based approaches were classified using K-Nearest Neighbor, Back Propagation Neural network, Multilayer Perceptron, Principal Component Analysis, Support Vector machine and Linear Discriminant Analysis.

3.2. Model Free Features
Model free approach extract the gait features for classification and analysis from the captured video or image without the person direct involvement with the system. The general process involves person...
detection, background subtraction, silhouette extraction, feature extraction and classification. Model free approach is preferred for its low computational cost and better performance in low quality video. Background subtraction is a process of classifying the pixel as foreground and background. Basic background techniques are frame differencing and filtering which uses the concept of min, max, mean, median and histogram filter. Gait image descriptors are body length and height, GEI, gradient histogram energy image, Fourier descriptors, normal vectors, Principal Component Analysis, k-means, HMM, standard deviation and frequency domain features, Gabor filters Multi-linear Discriminant Analysis, Chrono-Gait Image, Linear Discriminant Analysis, Microsoft Kinect distance method.

### TABLE 1: Datasets available for gait biometric analysis

<table>
<thead>
<tr>
<th>METHOD</th>
<th>DATASET</th>
<th>DEVICE</th>
<th>SUBJECTS</th>
<th>DATA FORMAT</th>
</tr>
</thead>
<tbody>
<tr>
<td>WEARABLE SENSOR</td>
<td>OU-ISIR Inertial Sensor</td>
<td>tri-axial accelerometer and gyroscope, smartphone</td>
<td>744 – Males, 389 – females, Age- 2 to 87</td>
<td></td>
</tr>
<tr>
<td></td>
<td>ZJU-gaitacc</td>
<td>Wii Remote</td>
<td>175 Subjects 22-session 0 153- session 1 153- session 2</td>
<td></td>
</tr>
<tr>
<td></td>
<td>BWRMultiDevice</td>
<td>3 mobile devices with accelerometer</td>
<td>25 subjects (session 1&amp; 2)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CASIA D – FOOT PRINT</td>
<td>Foot pressure measurement plate + foot scanner</td>
<td>88 subjects</td>
<td></td>
</tr>
<tr>
<td>VISION BASED</td>
<td>OU-ISIR Biometric Database</td>
<td>camera</td>
<td>62 528, Single view</td>
<td></td>
</tr>
<tr>
<td></td>
<td>OU-ISIR Multiview Database</td>
<td>7 cameras</td>
<td>10,307, Multi view</td>
<td></td>
</tr>
<tr>
<td></td>
<td>CASIA A – STANDARD DATASET</td>
<td>camera</td>
<td>20 person ; each with 12 sequence in three directions</td>
<td></td>
</tr>
<tr>
<td>Dataset</td>
<td>Description</td>
<td>Subjects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>---------</td>
<td>-------------</td>
<td>----------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>CASIA B</td>
<td>Multiview Gait dataset</td>
<td>Captured in 11 views</td>
<td>124 subjects</td>
<td></td>
</tr>
<tr>
<td>CASIA C</td>
<td>INFRARED</td>
<td>Infrared thermal camera</td>
<td>153 subjects</td>
<td></td>
</tr>
<tr>
<td>Tum Gaid Dataset</td>
<td>RGB Video, audio and depth</td>
<td>305 subjects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DBMHI-DB2</td>
<td>Camera (front and back)</td>
<td>100 subjects</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Pathological Dataset**

- **Gait in Neurodegenerative Disease Database**
  - force-sensitive resistors
  - 64 Subjects (Parkinson Disease - 15, Huntington Disease - 20, amyotrophic lateral sclerosis - 13, Healthy - 16)

- **Daphnet Freezing of Gait Data Set**
  - Accelerator Sensor
  - 10 Parkinson disease Patient (3 measurements - ankle, knee and hip)
4. MACHINE LEARNING IN GAIT RECOGNITION

Machine learning concepts are required to build the intelligent system to learn, recognize, analyze the gait characteristics with large number of image sequence collected to create the gait dataset. Figure 3 gives the general process of gait analysis using machine learning algorithms. Artificial intelligence techniques like data mining, machine learning, deep learning and automation are considered for the innovation in gait analysis fields. In machine learning paradigm, learning models are like supervised learning, unsupervised learning, reinforcement learning used in various design perspective to capture the gait sequence, finding the gait pattern from the characteristic features extracted with respect to the implementing field.

Machine learning approach follows the sequential procedure of data collection, feature extraction, learning algorithm and the evaluation methodology. Feature selection is a process used to identify the unique characteristics and eliminate the unnecessary information to reduce the problem of over fitting and training time thereby increasing the accuracy of the model. Feature selection techniques which are commonly used are univariate selection, feature importance, correlation matrix with heatmap and filter methods. Filter and wrapper feature selection method is specially used for gait sequence images. Wrapper methods is quite expensive since the learning algorithm repeats the random resampling technique usage for obtaining the feature subset accurately. Filter method use heuristic search procedure to obtain the undesirable feature subset. Without the help of learning algorithm, undesirable features were removed in terms of correlation between the examined data. Popular filter method is recursive feature elimination algorithm which removes the low weighted features. Feature subset were fed into evaluation procedure which determines the quality by comparing with the previous best value given by predefined stopping criterion and validated [22].

Model selection is the technique of selecting the appropriate model from the set of models which has high classification or prediction accuracy. Model is selected with the help of statistical technique, exploratory data analysis, scientific techniques and specification of a model. It is recommended that learning approach should be selected along with model selection [23, 24]. Generally, dataset is divide into three portions namely training data, cross-validation data and testing data. Training data is utilized to train the model with the input and known outcome. Cross validation data helps for frequent evaluation to improve the effectiveness of the model. Testing data is used for validating the specific model. Selected model is estimated to perform well with the training data and computed with similar learning model with various hyper parameters, complexities. Validation dataset is opted using the methods: holdout method, k-fold cross validation, leave-one cross validation, random subsampling cross validation and bootstrapping.
4.1. Classification Network
Classification Network has statistical and derivative phase. Statistical phase extracts the statistical measurements from the non-linear transformed input. Derivative phase improves the accuracy with mathematical model. Classification networks are classified into shallow network and deep neural network. Shallow neural network uses solitary hidden layer whereas deep neural network uses more than one hidden layer. Shallow networks used for gait recognition are K-Neighbor, Hidden Markov Model, Bayesian Network, Support Vector machine and BPNN. Deep neural network increases the accuracy rate when compared to shallow network and has become more interesting approaches in extracting features in recent years. Wang et al[33] used nearest neighbour approach to generate the silhouette signal to identify the person and achieved 77.4% classification rate with CASIA B dataset. Jeevan et al[34] recognizes the gait of an individual with Gait Pal and Pal Entropy using Principal Component Analysis with Support Vector machines. Implementing with treadmill dataset and CASIA A, B and C datasets, achieved 92.3% accuracy. Xiao et al[35] recognizes the individual by their gait sequence using Zernike moments and Back Propagation Neural Network (BPNN)

4.2. Deep Learning Approaches
In the family of machine learning algorithms, deep learning has become the most important methods due to its wide applications. Deep architectures represent the high level abstractions using multiple
stacked learning neural network layers and separate network segment is used to learn separate data representation. The energetic factors of deep neural network are feature learning, hierarchical and distributed representations, computational speed with large dataset varies in the automation process and analysis of depth. Multiple neural network layers are convolutional layer, sampling pooling layers, activation layer and fully connected layer. The multiple arrangement of these layer and dimensions varies corresponding to the application and architecture.

5. DEEP ARCHITECTURES
In deep learning approaches, there is no need to create a separate feature set since the discriminative features were found while training the model. Deep architectures which are used widely for gait recognition is convolutional neural network, multiple stacked auto encoder and deep Boltzmann machine and Restricted Boltzmann Machine.

5.1. Deep Boltzmann Machine
A Deep Boltzmann Machine uses hidden random variables for multiple layers of symmetrically couple stochastic Markov Random Field[44]. It has visible unit set \( \mathbf{v} \in \{0,1\}^D \) and hidden layer series units \( \mathbf{h}^{(1)} \in \{0,1\}^F_1, \mathbf{h}^{(2)} \in \{0,1\}^F_2, \ldots, \mathbf{h}^{(L)} \in \{0,1\}^F_L \). Units in adjacent layers have connection between them. For instance, consider three hidden layers for DBM as shown in Figure 4. The probability of a visible vector \( \mathbf{v} \) assigned by Deep Boltzmann Machine is given as (1).

\[
P(\mathbf{v}; \theta) = \frac{1}{Z(\theta)} \sum_{\mathbf{h}} \exp \left( \sum_{ij} W_{ij}^{(1)} u_i h_j^{(1)} + \sum_{jl} W_{jl}^{(2)} h_j^{(1)} h_l^{(2)} + \sum_{lm} W_{lm}^{(3)} h_l^{(2)} h_m^{(3)} \right)
\]

In equation (1) \( \mathbf{h} \) is a set of hidden units represented as \( \{h^{(1)}, h^{(2)}, h^{(3)}\} \) and the model parameter \( \theta = \{W^{(1)}, W^{(2)}, W^{(3)}\} \). This represents symmetric interactions between the visible units and hidden units. Restricted Boltzmann Machine model is recovered by assigning \( W^{(2)} \) and \( W^{(3)} \) as 0. Initially the weights in hidden units are asymmetric as the weights are twice as the top layer. Later to maintain the symmetry, weights are modified.

![Deep Boltzmann Machine](image)

**Fig. 4.** a) Deep Boltzmann machine has undirected connection. b) DBM is pretrained with RBM stack. The weights of first and last layers of RBM is modified as \( W^{(1)} \) and \( W^{(3)} \) to maintain the symmetric pattern [44].

Yu et al[43] proposed the Convolutional Restricted Boltzmann Machine to extract the gait features to identify the individuals. CRBM was experimented on CASIA B gait dataset with 11 different angles and 3 variations in outfit obtained the accuracy of 96.5%.
5.2. Deep Stacked Autoencoder

**Fig. 5.** a) single autoencoder with same dimensionality of input layer \((x_1, x_2, \ldots, x_n)\) and output layer, lower dimensionality in feature extraction layer. b) Stacked autoencoder enables the supervised classification in the final output layer [45].

Deep Stacked Autoencoder has more layers to extract the discriminative and representative features. It is formed by stacking the autoencoder network one above the other and implemented multilayer perceptron. A single autoencoder network has encoding layer for encoding the input and decoder layer for reconstructing the input. In decoding layer, number of neurons and dimension of input are considered to be in the same count. It recovers the input \(x\) with high accuracy using the two stage approximation model. The code \(h\) used for recovering is

\[
\hat{x} = f_{dec}(f_{enc}(x))
\]

Where \(f_{enc}\) is encoding function and \(f_{dec}\) is the decoding function computed by their respective coding layer. The main function of stacked autoencoder to train the unsupervised input layer by layer. Figure 5 shows how the single autoencoder turns to stacked autoencoder to perform the classification. The final output layer produces the supervised classification by fine tuning the pre-trained network with Back Propagation[45]. Sun et al [46] proposed the deep network by combining the convolutinal neural network and stacked autoencoder to recognize the human with the gait subtle features. Micro-motion dataset is a radar echo dataset collected using the Radio Frequency device is used to implement the proposed network and achieved 96% as highest average precision.

5.3. Recurrent Neural Network

Recurrent Neural Network (RNN) is designed to solve classification problems. RNN deals with sequential data with more interdependencies especially well-suited for clinical data. In RNN, input sequence is termed as \(X_1, X_2, \ldots, X_n\) and output sequence is termed as \(Y_1, Y_2, \ldots, Y_n\). Output Vector \(Y_i\) depends on the past values of \(X\) and it is given by

\[
Y_i = f_i(X_1, \ldots, X_n)
\]

Hidden states are utilized to learn parameters from the sequential data given in the previous and current recurrence. The hidden state, \(h_i\) and recurrent layer output, \(y_i\) is calculated as

\[
h_i = \tanh(W_{hh}h_{i-1} + W_{xh}x_i + b_h)
\]
Where W denote weight, b denote biases and i denote the frame index. To solve the vanishing gradient problem, Long Short Term Memory (LSTM) and Gated Recurrent units (GRU) is proposed. These networks maintain the long term inter relations and nonlinear dynamics effectively. To control the number of parameters in the network, same weight is determined for all layers. Hidden state is updated with forget gate value $f_i$, input gate value $g_i$ and output gate value $o_i$. Learning variables are given as follows:

$$f_i = \sigma(W_{xf}x_i + W_{hf}h_{i-1} + b_f)$$

$$g_i = \sigma(W_{xi}x_i + W_{hi}h_{i-1} + b_i)$$

$$C_i = f_t \cdot C_{i-1} + i_t \cdot \tanh(W_{xc}x_i + W_{hc}h_{i-1} + b_c)$$

$$o_i = \sigma(W_{xi}x_i + W_{ho}h_{i-1} + b_o)$$

$$h_i = i_t \cdot \tanh(C_i)$$

Where W denote weights and b denote biases. Prado et al [65] used RNN to classify the cerebral palsy disease in adults and children. Data were collected from 28 adults and 7 children with cerebral palsy using the pressure inertial foot sensor. They have proposed RNN with instrumented foot sensor which does not require any preprocessing technique, thereby processing time is reduced to less than 1 second.

5.4. Convolutional Neural Network

Convolutional neural network (CNN) is multilayer perceptron in a regularized version effectively implemented on Graphics Processing Unit (GPU) or General Purpose computing on Graphics Processing Unit (GPGPU). GPU which is mainly developed for video game application is used to increase the speed of deep Neural Network processing. Convolutional Neural network has convolution layer, Pooling Layer, Activation function, Loss function, regularization, optimization process. Figure 6 demonstrates the general outline of convolutional neural network. First parameter of the CNN is the input data in the form of image or signal. Second parameter is the filter. Fully connected layer gives the output in the form of vector from which feature map is obtained. In the form of multidimensional array, each input and filter are saved independently in the network. Local information of the image is identified by convolutional layer.

**Fig. 6. Schematic diagram of convolutional Neural network**

Convolution layer is categorized into titled convolution, transposed convolution, dilated convolution, network in network and inception module. Pooling layer is categorized into average pooling, max pooling, Lp pooling, mixed pooling, stochastic pooling, spectral pooling, spatial pyramid pooling and multi-scale order less pooling. Activation function used are Rectified Linear Unit (ReLU), Leaky ReLU, parametric ReLU, Randomized ReLU, Exponential Linear Unit, Maxout and Probout. Methods used for convolution neural network are as follows: LeNet 5 is a 8 Layered architecture...
used for image recognition. AlexNet uses GPU to train the high resolution large dataset with 1000 different classes. Alexnet does faster processing compared to LeNet5 due to its non saturating neurons. Overfitting problem is reducing by using dropout method. Zeiler and Fergus proposed ZFNet as multi-layered deconvolutional neural network with 8 layers. ZfNet uses GPU to classify large image dataset and dropout to reduce the parameters. Network – In-Network (NIN) is proposed to implement per-pixel nonlinearity and to reduce the number of parameters. NIN uses 1 x 1 convolutional network with ReLU activation. Global average pooling is used to produce logit vectors. Region based Convolutional Neural Network (R-CNN) detects the object by extracting the features with region proposals. Forward computations is required to choose the thousand number of region proposals from a single image increase the processing time. To overcome this, Fast RCNN is proposed to perform the forward computation as a whole. Entire image is given as an image for processing each region proposals at a time unlike RCNN where region proposals are given as an input. Faster R-CNN uses region proposal network, thereby reducing the number of generated region proposal and maintaining the accuracy of object detection. Mask R-CNN uses fully convolutional layer for locating the objects at pixel level. Comparatively Mask R-CNN enhanced object detection accuracy rather than Faster R-CNN. GoogLeNet uses 22 convolutional layer for transfer learning. Optimization is increased when image classification is done via GoogLeNet CNN model. Visual Geometry Group (VGG) is widely used in gait recognition when compared to other CNN model since it provides high accuracy. VGG performs training faster since it has reduced number of parameters.

Table 3 provides the overview of recent work in gait recognition using deep architectures.

<table>
<thead>
<tr>
<th>YEAR</th>
<th>REFERENCE</th>
<th>APPROACH</th>
<th>GAIT FEATURES</th>
<th>DATASET</th>
<th>CLASSIFIER</th>
<th>PURPOSE</th>
<th>ACCURACY</th>
</tr>
</thead>
<tbody>
<tr>
<td>2006</td>
<td>[40] Yu et al</td>
<td>Model free</td>
<td>GEI</td>
<td>CASIA B</td>
<td>Multilayer autoencoder</td>
<td>Identification</td>
<td>-</td>
</tr>
<tr>
<td>2015</td>
<td>[28] Chao Yan et al</td>
<td>Model free</td>
<td>GEI</td>
<td>CASIA B</td>
<td>CNN</td>
<td>Identification</td>
<td>96%</td>
</tr>
<tr>
<td>2015</td>
<td>[29] Perlin et al</td>
<td>-</td>
<td>Cloth and gender</td>
<td>H3D dataset, ViPer dataset, HATdb</td>
<td>CNN</td>
<td>Pattern for clothes and gender</td>
<td>-</td>
</tr>
<tr>
<td>2016</td>
<td>[48] Hannink et al</td>
<td>Model Based</td>
<td>acclerator, gyroscope parameter s</td>
<td>GAITRite</td>
<td>CNN with modifications</td>
<td>Identification</td>
<td>-</td>
</tr>
<tr>
<td>2017</td>
<td>[37] Alotaibi et al</td>
<td>Model free</td>
<td>CASIA B</td>
<td>CNN, keras, python</td>
<td>Identification</td>
<td>Mis-classification rate &lt;0.15%</td>
<td></td>
</tr>
<tr>
<td>2017</td>
<td>[41] Sokolova and Konushin</td>
<td>Model free</td>
<td>Optical Flow</td>
<td>TUM-GAIT dataset</td>
<td>CNN - VGG</td>
<td>Identification</td>
<td>16 hours , 97%</td>
</tr>
<tr>
<td>2017</td>
<td>[42] Dehzangi et al</td>
<td>Model Based</td>
<td>Gait measurements</td>
<td>Sensor collected data</td>
<td>Deep CNN</td>
<td>Identification</td>
<td>97.60%</td>
</tr>
<tr>
<td>Year</td>
<td>Author(s)</td>
<td>Model Type</td>
<td>Features</td>
<td>Dataset</td>
<td>Classifier</td>
<td>Identification Rate</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>-----------</td>
<td>------------</td>
<td>----------</td>
<td>---------</td>
<td>------------</td>
<td>---------------------</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[31] Gadaleta et al</td>
<td>Model Based</td>
<td>accelerometer, gyroscope parameters</td>
<td>GaitDataset McGill CS</td>
<td>CNN</td>
<td>94%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[32] Batchuluun et al</td>
<td>Model free</td>
<td>GEI</td>
<td>DBMHI-DB1</td>
<td>CNN - VGG</td>
<td>Authenticaton Identification 100%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[47] Liu et al</td>
<td>Model free</td>
<td>GEI</td>
<td>CASIA B</td>
<td>CNN - VGG</td>
<td>Identification 89.62%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[53] Yuan et al</td>
<td>Model Based</td>
<td>accelerometer, gyroscope parameters</td>
<td>GITHUB</td>
<td>CNN</td>
<td>Authentication Identification 87%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[54] Batchuluun et al</td>
<td>Model free</td>
<td>camera image</td>
<td>Sensor collected data DBMHI-DB1, DBMHI-DB2</td>
<td>LSTM CNN</td>
<td>Authentication Identification 97%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[55] Sun et al</td>
<td>Model Based</td>
<td>spectogram</td>
<td>Self collected dataset from radar and microwave anechoic chamber</td>
<td>CNN, Stacked Auto encoder</td>
<td>Authentication Identification 96.36%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[43] Yu et al</td>
<td>Model free</td>
<td>GEI</td>
<td>CASIA B</td>
<td>CNN - BOLTZMAN MACHINE</td>
<td>Identification 96.50%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[57] Liu et al</td>
<td>Model free</td>
<td>GEI</td>
<td>OU-ISIR</td>
<td>CNN, SIAMESE NEURAL NETWORK</td>
<td>Identification 88%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[49] Tong et al</td>
<td>Model free</td>
<td>GEI</td>
<td>CASIA-B, OU-ISIR, and CMU MoBo</td>
<td>SPATIAL TEMPORAL DEEP NN</td>
<td>Identification 95.67%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[58] Takemura et al</td>
<td>Model free</td>
<td>GEI</td>
<td>OU-ISIR</td>
<td>CNN</td>
<td>Identification 95%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[59] Yao et al</td>
<td>Model free</td>
<td>Skeleton Gait image</td>
<td>CASIA B</td>
<td>CNN</td>
<td>Identification 92.09%</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>[64] Verlekar et al</td>
<td>Model free</td>
<td>GEI</td>
<td>DAI, DAI2 INIT</td>
<td>CNN – LSTM</td>
<td>Clinical gait disease 95%</td>
<td></td>
</tr>
<tr>
<td>Year</td>
<td>Authors</td>
<td>Model Type</td>
<td>Sensor Type</td>
<td>Dataset Description</td>
<td>Classification Method</td>
<td>Accuracy</td>
<td></td>
</tr>
<tr>
<td>------</td>
<td>---------</td>
<td>---------------------</td>
<td>---------------------------------</td>
<td>----------------------------------------------</td>
<td>-----------------------</td>
<td>--------------</td>
<td></td>
</tr>
<tr>
<td>2018</td>
<td>Wang et al</td>
<td>Model-based</td>
<td>Threadmill sensor</td>
<td>Recorded with Body weight support threadmill</td>
<td>CNN-AlexNet</td>
<td>99.98%</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>Maryam et al</td>
<td>Model-free</td>
<td>GEI</td>
<td>OULP and Casia-B</td>
<td>CNN</td>
<td>-</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>Wang et al</td>
<td>Model-free</td>
<td>GEI</td>
<td>CASIA B, OU-ISIR</td>
<td>CNN-2 BRANCH</td>
<td>94%</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>Hawas et al</td>
<td>Model-free</td>
<td>GEI</td>
<td>CASIA B</td>
<td>CNN</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>2019</td>
<td>Wang et al</td>
<td>Model-free</td>
<td>GEI</td>
<td>CASIA A, B, OU-ISIR</td>
<td>CNN</td>
<td>95%</td>
<td></td>
</tr>
<tr>
<td>2020</td>
<td>Shao et al</td>
<td>Model-free</td>
<td>GEI</td>
<td>CASIA B, LeNet5</td>
<td>Identification with Covariates</td>
<td>98.3% (normal) 89.2% (Backpack) 95.8% (Wearing coat)</td>
<td></td>
</tr>
</tbody>
</table>

**Human Gait Recognition:**
Hannink et al [48] used convolutional neural network with ReLU activation function to extract the gait feature from wearable sensor dataset. Tong et al [49] used spatial-temporal convolutional neural network to recognize the gait in multi-view angle with 92.54% accuracy. Liu et al [47] combined the convolutional neural network and support vector machine to predict the gender of an individual using the CASIA B dataset with 87.94%. The convolution model VGGNet-16 is used to train the network with GEI input while SVM classify the descriptors provided by optimizing the weight parameters. Shao et al [63] proposed improved LeNet5 network combined with Residual Network to accelerate the convergence speed and to improve the recognition rate. CASIA B dataset is trained and tested with the improved LeNet model of CNN. They focused to work on the covariates like Viewing angle, backpack, jacket and speed. The accuracy of gait recognition with backpack, wearing coat and normal are 89.2%, 95.8% and 98.3% respectively. Maryam et al [27] used FCN (Fully Convolutional Neural Network) to reconstruct the incomplete Gait Energy Image (GEI) into a complete GEI. Experiments were carried out in OU-ISIR Biometric – OULP database and CASIA B database to obtain the complete GEI from 0.1 part of a gait cycle. This method helps in matching view-variant incomplete gait image.

**Clinical Gait Analysis:**
Verlekar [64] proposed a fine tuned VGG-19 Deep architecture to classify gait disease type. The pathological dataset with diplegia, Hemiplegia, neuropathy, parkinson and normal individual gait dataset is used to experiment the proposed model. The classification of normal and impaired gait type is performed with high accuracy. Wang et al [64] used Alexnet CNN model to perform the gesture recognition system with 99.98% accuracy. Five different gestures such as moving forward, moving backward, moving left, moving right and stop are recognized with Body Weight Support Treadmill (BWST) Platform. This recognition system helps in lower limb rehabilitation. This system is controlled by a trainer to monitor the patient walking sequence in BWST platform.
6. DISCUSSION
A comprehensive survey of gait recognition and analysis methods based on wearable sensors and vision based methods, various features and deep architecture techniques were presented. A review of gait databases with different techniques used to capture gait features were given. Figure 7 depicts the recurrently used gait features in various fields and obtained the best accuracy with the recent techniques of deep architectures.

![Statistical figures based on the referred papers](a) gait features and usage b) Gait analysis applied in the various fields c) Different combinations of Deep architectures used in gait recognition.

**Future Research Directions.**
The scope of future research provides an insight into remote health monitoring, surveillance and robotics. Immense opportunities are open to enhance and innovate in the current technological
environment in the following fields. Application of Multivariate statistical analysis and deep learning techniques is becoming increasing popular in the area of clinical gait analysis, thereby providing new insights into human gait control. According to Global Burden of Disease (GDB), musculoskeletal disorder is estimated to witness the second highest growth rate across the globe. In recent years, number of people with abnormal walking posture is increasing. Research were carried out to differentiate the healthy walking posture and abnormal walking posture. Sports professionals utilized gait analysis to monitor the movement of leg in terms of angle, position etc. The measurement and assessment of the players improved their performance as well as decreased the risk of injury. Model free method is preferred quickly to interpret about the movement of the players. Trajectory movements of model based human gait helps to design the walking nature of robots. Human gait recognition system is becoming widely spread due to its nature of identifying individual without their intervention with the system. Still there exist some limitations such as impact of environmental changes, user location and orientation, sensing multi-user activity and security considerations. To perform the longitudinal study over the wide range of covariates, there is a need for a large multi-modality, multi-covariate standard dataset. Our future work is to carry out the experiments with the recent network architectures and novel approaches with the huge dataset of different view angles, clothing and walking style promoting the better automatic gait recognition and gait analysis system.

7. CONCLUSION
The most interesting research in biometrics is automatic gait recognition when compared to other human unique features. Though there are many approaches to overcome the variations in gait recognition, still there are challenges to recognize a person. Challenges are distinctive gait datasets, degree of stability in identifying, sensing modality, covariates and spoofing effects, and exploring new algorithms. From the developmental perspective, human gait recognition has gained maturity in adopting recent methodologies to provide the high accuracy and the same was analyzed and studied in the view of vision based system. In this paper, the basic knowledge about the human gait is identified and explored. The comparative analysis with the existing techniques of gait recognition were discussed. In this work, the survey of recent deep architecture model on human gait identification, authentication and clinical applications were discussed. Future research directions and challenges were identified.

8. REFERENCES


[42] Lukas Vareka and pavel Mautner “Stacked Autoencoders for the P300 Component Detection”, Neuroinformatics Research Group, Department of Computer Science and Engineering, Faculty of Applied Sciences, University of West Bohemia, Pilsen, Czechia, (2017).


[64] Ver


