An Overview: Deep Learning For GAIT Recognition

Kandarpa Uday Narasimha Murthy[1], Mr. Rishi Kushwaha[2]
[1, 2] School of Engineering
[1, 2] SSSUTMS, Sehore, M.P(India)

Abstract- *After a quick discussion of the key human gait metrics, the state-of-the-art in deep learning for human gait analysis is covered in detail. According to the sensing technology, the gait data collection methods are divided into three categories: video sequences, wearable sensors, floor sensors, and publicly accessible datasets. The performance of each group's known artificial neural network architectures for deep learning is reviewed, with special attention paid to the spatiotemporal nature of gait data and the justification for multi-sensor, multi-modality fusion. It is demonstrated that deep learning convolutional neural networks often outperform shallow learning models by the majority of key measures. This is attributed to the gait data's stated characteristics.*

Keywords- Deep learning, floor sensor, gait, neural network, sensor fusion, video sequence, wearable sensor.

I. INTRODUCTION

GAIT refers to the displacement of the center of gravity during locomotion. In humans, it is achieved through the synchronized movement of the lower limbs and the trunk, resulting in a move from one position to the other [1]. Every human being has a particular behavioural characteristic that is influenced by mutually exclusive factors including weight, gender, and age. Gait analysis has a long history, which documents a gradual growth from descriptive research to more complex techniques. Aristotle was the first to study animal and human motion about 350 BC. [2]. However, it was the works of Newton, Galileo, and Leonardo da Vinci that first provided useful explanations of how people move. Galileo's apprentice and the founder of biomechanics, Borelli [3], supplied a significant boost to scientific gait analysis methods by determining the human body's center of gravity and how walkers maintain their balance. [4]. The Weber brothers defined the gait cycle in 1836 and described gait as a periodic movement based on the forward leg motion that resembled a pendulum [5]. In order to demonstrate that all four of the horse's hooves were off the ground while trotting, Muybridge utilized 12 cameras to record racehorse motion in 1878. He also adopted a similar strategy to take a number of images of people moving [6]. The first substantial quantitative use of gait analysis was in 1895 [4] when Braune and Fisher used a photographic technique to determine a human body’s velocity, acceleration, and dimensional trajectory to estimate the forces involved during the gait cycle. In 1930s, Bernstein studied the dynamic locomotion of 150 subjects to determine the center of gravity of each limb segment of the subjects using a photographic technique [7].

Ground Reaction Force (GRF) was introduced in human gait understanding in 1924 when Cavanagh and Lafortune [8] designed a force plate to measure the magnitude and the direction of GRF. The platform was improved by Elftman in 1938 using a high-speed cinematic camera to capture a pointer movement resulting from the force applied to the platform [9]. A substantial amount of knowledge was contributed to the human locomotion analysis in the 1950s, with the motivation to treat World War II veterans [10].

Over the past two decades, the rapid expansion of the capabilities of sensor systems, including advances in scientific processing, has allowed us to extract more lavish data from increasing means of detection. In this context, progress in evaluating various human locomotion parameters based on the ever-increasing amount and quality of data is aided by advances in new gait detection tools. Of course, this also highlighted the difficulties arising from the requirement to achieve multi-source and multi-sensor fusion from various big data. Furthermore, it is unclear whether the complex personas of gait adequately reflect the simple and widely used indices that are typically mediated by systems for rapid and robust diagnosis, recognition, and grouping. . Deep learning models, on the other hand, emerged as a result of advances in machine learning technology. These models can provide faster and more accurate results from databases that continue to grow in size and scope, and can be applied to complex data with minimal preprocessing. It offers new possibilities for detection, fusion and classification from a wide variety of multi-source and multi-sensor data. Among them, spatio-temporal parameters of walking are currently attracting attention because of their potential for various applications such as walking. health care [11], [12], sport [13], [14], and identification of individuals for security [15], [16].

Gait analysis is immature and there is no gold standard for collection or data processing. The rest of this essay organizes the most commonly used modalities.

We study human gait based on sensor principles and the amount and type of sensor data generated into three groups.

Video Sequences (VS), Wearable Sensors (WS), and Ground Sensors (FS). We show that the detection principle used for this grouping also influences the choice of deep learning processing method.

The VS solution is based on action detection using spatio-temporal information. WS systems typically include inertial sensors to detect the velocity, acceleration, and orientation of the human body during physical human activity. FS typically monitors GRF induced by contact with the ground during the gait cycle. Data obtained from these modalities are analyzed and classified using advanced supervised learning techniques based on appropriate assumptions. This review is underpinned by an extensive literature search but only the most recent works, combining gait recognition with deep learning algorithms, are presented in more detail.

II. LITERATURE REVIEW

To define the contribution of deep mastering in human gait analysis, it is indispensable to apprehend how people stroll and the functions of gait giving upward jab to the set of strategies utilized for analysis.

A. Gait Parameters

Gait can be perceived as a transformation of a Genius undertaking to muscle contraction patterns ensuing in a walking sequence. It is a chain of instructions generated in the intelligence and transmit- ted thru the spinal cord to prompt the decrease neural center, which will hence result in muscle contraction patterns assisted by using sensory comments from joints, muscle groups and different receptors to control the movements. This will end result in the ft often contacting the floor floor to cross the trunk and lower limbs in a coordinated way, turning in a alternate in the physique center-of-mass function.

Gait is a sequence of periodic events characterized as repetitive cycles for each foot [4]. Each cycle is divided into two phases (see figure 1):

a) Stance Phase (approximately 60% of the gait cycle, with the foot in contact with the ground). This phase is subdivided into four intervals (**A, B, C, D**).
b) Swing Phase (approximately 40% of the gait cycle with the foot swinging and not in contact with the ground). This phase is subdivided into three intervals (**E, F, G**).

A. Heel strike or Initial contact: It starts the moment the foot touches the ground, and it is the initial double-limb support interval. In the case of the right foot leading, the double support starts with left foot being on the ground when the right foot heel makes initial contact and finishes when the left foot leaves the ground with the left toe-off prepared to swing. At the end of this interval, the body weight is completely shifted onto the stance (leading) limb.
B. Loading response or Foot flat: This is a single support interval following the double support interval. The trunk is at its lowest position, the knee is flexed, and a plantarflexion occurs at the ankle.
C. Mid-stance: This is a single support interval between opposite toe-off and heel-off. The trunk is in its highest point

Fig. 1. Important gait events and intervals in a normal gait cycle.

and slowing its forward speed. The body center-of-mass is aligned with the forefoot (ball of the foot).

D. Terminal stance or Heel-off: The heel rises in preparation for opposite swing. The trunk is sinking from its highest point, the knee has extant peak near the time of heel rise and ankle has dorsiflexion after heel rise.
E. Pre-swing: This is the second double-limb support interval. The opposite initial contact occurs, and the hip is beginning to flex, the knee is flexing, and the ankle is at plantarflexion. The toe is in last contact before the swing, finishing the push-off started in interval **D.**
F. Initial swing and Mid-swing: This interval begins with the toe-off into single support and starting to swing. The body weight is shifted to the opposite forefoot. In this instant, the knee joint gets the maximum flexion. The hip is flexing and the limb advances in preparation for a stride.
G. Terminal swing: This is the last interval of gait cycle and the end of the swing phase. The interval begins at maximum knee flexion and ends with maximum

extension of the swinging limb forward. The hip continues flexion and the knee extends in regard to gravity, the ankle continues dorsiflexion to end neutral, ready for the heel strike.

With regard to the above gait events, the following parame- ters of human gait are usually analyzed in clinical settings [17] for healthcare tasks, using various sensing and data processing methods:

- Cadence or rhythm (number of steps per unit time)
- Stride length
- Velocity
- Direction of leg segments
- Step angle
- Swing time for each foot
- Step width
- Support time
- Ground Reaction Force (GRF)
- Electrical activity produced by muscles
- Momentum and forces
- Body posture

It is well worth bringing up at this point, that while the listed parameters have clear observational value, it is hard to declare that any of these, or their combination, would characterize the maximum variability of the uncooked statistics due to a fitness condition. This issue has a direct impact on the capability to detect, with the lowest threshold less expensive through the uncooked records quality, a significant deviation from the norm.

B. *Applications of Gait Analysis*

The field of research in human gait is broad, with many specific applications. In clinical applications, as gait abnormality impacts a excessive share of the population, gait is studied to diagnose neurodegenerative diseases such as Parkinson’s disease (PD), myelopathies, spinal amyotrophy, more than one sclerosis, cerebellar ataxia, intelligence tumors, cranioencephalic trauma, positive sorts of dementia, neuromuscular diseases etc. [17]. In fact, the floor response force of humans all through the gait cycle has been used to observe PD in [18]. The study shows that stance time, swing time, stride time and foot strike profiles can be used to distinguish PD patients from healthy controls. In addition, the spatiotemporal parameters of gait have been studied [19] to investigate lower limb prosthesis users.

In protection applications, gait evaluation as a biometric has demonstrated its success to distinguish and perceive people, with minimal cooperation required from the subject. The aim is to perceive humans from a faraway primarily based on their on-foot habit. Typically, people gait is captured by CCTV cameras as suggested in [20], [21]. In [22], [23], the floor response pressure has been determined to be enormous in figuring out topics primarily based on their footstep alerts and stepping behavior.

Injuries in many instances appear at some stage in sports activities pastime and some techniques to consider athletes’ recovery are based totally on gait,e.g. by means of analyzing forces exerted on each muscle through electromyography in [24]. The kinematic parameters of gait are used to analyze a variety of indoor and out of doors activities, such as sports coaching and medical rehabilitation of sufferers using a wearable sensor [13]. Even exceptional gait characteristics evaluation techniques are used to assess athletes’ potential to return to recreation after surgical procedure due to tear in the anterior cruciate ligament which motives knee instability [25]. Further, the gait dual-task paradigm for comprehensive athlete contrast following a sports-related concussion are reviewed in [14].

It is fascinating to word that gait analysis is utilized to clas- sify a person’s gender based totally on their gait [26]. Furthermore, tries to pick out a person’s emotional state, such as pride, happiness, fear and anger, have been based totally on gait [27].

C. *Deep Learning for Gait Analysis*

Supervised machine learning is a branch of artificial intelligence (AI) and a particular variety of machine learning. Algorithms or mathematical models are built and educated with a given set of inputs and desired outputs. A getting to know algorithm trains the mannequin based on two gaining knowledge of styles, shallow learning or deep learning, to produce an educated “machine” that incorporates out the desired task. The fashions are tested by exploring the information structure primarily based on the learned mapping feature to assign speculation type which is managed by way of the person to consider the model performance [28]. Shallow getting to know depends on handcrafted features discovered in a predefined relationship between the inputs and the output, such as linear regression, logistic regression, choice tree, Support Vector Machine (SVM), random forest, naïve Bayes, and k-nearest neighbor.

Deep structured gaining knowledge of or hierarchical getting to know is inspired by means of the biological neural networks’ shape and function. It is based totally at the start on the idea of multi-layer Artificial Neural Network (ANN) with the goal to analyze statistics representations automatically; thus, deep studying will become the method of preference the place the classification features, if known at all, are complex, with no straight forward quantitative relation to the raw data. Typically, the time period 'deep' refers to the quantity of layers in the variety of viable networks structures: Deep Belief Networks (DBN), Feedforward Deep Networks (FDN), Boltzmann Machine (BM), Generative Adversarial Networks (GAN), Convolutional Neural Networks (CNN), Recurrent Neural Networks (RNN), and Long-Short Term Memory (LSTM) a distinct form of RNN. A comprehensive presentation of the theory of ANNs and deep gaining knowledge of is no longer inside the scope of this Review, and the reader is referred to set up sources [29]. Further, we focal point on models with sensible value for gait purposes such as CNN and LSTM [30].The CNN model is suitable for processing 1D, 2D or 3D data that has a known grid-like topology [31]. The network has the ability to learn a high level of abstraction and features from large datasets by applying a convolution operation to the input data. Commonly, the network consists of convolution layers, pooling layers and normalization layers, with a set of filters and weights shared among these layers.

The convolutional layers output a feature map harvested automatically from the raw input data. The pooling layers are utilized to reduce the size of representation and make the convolution layer output more robust [29], [30]. The CNN model uses commonly two types of pooling layers: max pooling and average pooling. All convolution layers and pooling layers have activation functions (e.g. Sigmoid, Tanh, ReLU, Leaky ReLU), to calculate the weight of neuron and add a bias, deciding whether to fire the neuron or not [32].

LSTM networks are favorable for processing time series data, where the order is of importance, such as gait data sequences. In essence, they exploit recurrence, by using information from a previous forward pass over the network.

The computational complexities of deep learning are not specific to gait applications. The goal of using ANNs in gait analysis is to develop a model to extract gait features and perform well on unseen real-world gait data with high prediction accuracy. Commonly, for appropriate training and testing, the model is trained and validated on 70% of the data and tested on the remaining 30%. In supervised train- ing, the procedure is launched by initializing the weights randomly, processing the inputs and comparing the resultant output against the desired output. During training, the weights and biases are adjusted in every iteration, until the error is minimized, and validation is used to estimate the model performance during training. Lastly, the model is tested with unseen data, allowing to identify over-training.

The widely used accuracy measure for ANN gait analysis is the confusion matrix [33]. It is a table to visualize the number of predictions classified correctly and wrongly for each class. The table consists of true positive, true negative, false positive, and false negative classification occurrences. One of the advantages of the confusion matrix display is that it is straightforward to identify the decision confusions, thus possibly concluding on the quality of the model and data involved.

III. GAIT MODALITIES

The evolution of research in gait analysis suggests that, in order to capture the distinctiveness of gait, the various sens- ing modalities attempt to access biomechanical measures per- taining to the body's physical dimensions, body part masses, or the time-varying muscle-generated forces applied during the gait cycle. In the past decades, a number of modalities have proven their ability to capture gait characteristics and anomalies; however, the historically established methods used

Fig. 2. From right to left: video sequence, silhouette images and EGI image [20].

and skeleton model-free categories. (The above sub-division reference is to skeleton models, not machine learning models.) The model-based approach is in essence fitting video sequences of gait to multi-segment skeleton models, as pro- posed in [34], [35]. This method is computationally expensive, because of fitting skeletal segment models on sensor data, as well as the need to use the model-derived parameters to extract features. The extracted features are classified using shallow learning methods.

The model-free approach is based on extracting gait from VS using feature engineering, as proposed in [36], [37]. Here, deep learning is utilized to automatically extract gait features from VS, which maximizes the use of data variability and eliminates the dependence on handcrafting. Most of the avail- able model-free processed data is represented by Gait Energy Image (GEI), maps of optical flow and silhouettes [38] or Chrono-Gait Images (CGI) [39], [40]. These representations, extracted from VS, can capture both spatial and temporal information. As an example representation, GEI is defined mathematically as:

to analyze gait heavily rely on handcrafted features. With such an approach, salient features of the problem may be lost in the

$$GEI(x,y) = 1/s \sum_{t=1}^{s} F(x,y), \qquad (1)$$

process of feature engineering, and the classification result can be data dependent. This can be mitigated by utilizing deep learning for its capability of automatic feature extraction, delivering high statistical confidence by learning rich features of gait patterns from sensor data. Sensing modalities for gait data capture used in conjunction with deep learning can be divided into three main groups: video sequence (VS), wearable sensors (WS), and floor sensors (FS); further, each of these is described in more detail. In addition, the different types of algorithms typically applied to analyze gait data are presented and their ability to adjust to the characteristics of a modality or and/or scenario is elucidated.

A. *Video Sequence*

Gait recognition based on VS has been driven by the advances in general machine learning and image processing methods. The most common aim is to distinguish the identity of a person from a distance. A typical VS system consists of several cameras with optics suitable for capturing the gait cycle. Common VS data sources are suitably positioned CCTV cameras. The information gathered in the form of sequential video frames is subjected to image processing techniques, such as threshold filtering, edge detection, pixel count, background segmentation, counting of light and dark pixels, and converting images to black and white [17]. Gait recognition based on VS in literature is sub-divided into skeleton model-based

where *s* is the total number of frames to represent one gait cycle, and $F^t(x, y)$ is the binary silhouette of the subject at time *t*. Figure 2 shows schematically the extraction of GEI from the video sequence.

1) Video Sequence Databases: Once the VS representation algorithm is implemented, the machine learning model must be trained, validated and tested to assess its performance. The widely used benchmark is to train and test the algorithm with the following datasets (in chronological order of availability): CMU Motion of Body (MoBo) [41], USF Gait Based Human ID Challenge [42], CASIA [43], OU-ISIR treadmill [44], OU-ISIR [45] and TUM-GAID [46].

The Carnegie Mellon University Robotics Institute Motion of Body (MoBo) dataset [41] encompasses 25 subjects per- forming four different walking patterns on a treadmill, namely slow walk, fast walk, incline walk and walking with a ball. The subjects' gait is captured by six high-resolution cameras, distributed around the treadmill.

The University of South Florida Gait Based Human ID Challenge dataset [42] captures 122 subjects walking outside with shoes and clothes variations, as well as under different carrying load conditions. Gait is captured from a single viewing angle.

The Chinese Academy of Sciences Institute of Automation Gait Database CASIA [43] is divided into A, B, C, and D datasets. The CASIA A dataset contains 20 people; for each person, it contains 12 image sequences, four sequences for each of 3 angles (0, 45 and 90 degrees) to the image plane. The CASIA B dataset consists of 124 subjects' gait sequences captured from 11 views. The subjects performed normal walking, wearing a coat while walking, and carrying a bag while walking. The CASIA C dataset was captured by an infrared (thermal) camera from 153 subjects performing normal walking, slow walking, fast walking, and normal walking with a bag. The video sequence was taken from one angle at night time. The CASIA D dataset contains the video sequence and footprint images scans of 88 subjects with a wide age distribution. The video sequence is captured from a single angle and with no variations in clothing and carrying conditions.

The Osaka University Institute of Scientific and Industrial Research treadmill dataset, OU-ISIR treadmill [44], contains 200 subjects' gait captured on a treadmill by 25 cameras from different angles, 34 subjects with walking at different speeds and 68 subjects with 32 clothing variations. The dataset is dis- tributed in the form of silhouette sequences of subjects while walking on a treadmill. The same group's database on normal surface walking (not involving a treadmill), OU-ISIR [45] dataset, contains 4,007 (2135 males and 1872 females) with ages from 1 to 94 years. The dataset consists of silhouette sequences of the subject's gait captured by two cameras.

The Technical University of Munich Gait from Audio, Image and Depth database, TUM-GAID [46], contains 305 subjects' gait captured by video recording cameras at a single angle, while subjects walk indoors in both directions. Six walking conditions are captured for each subject from the side view namely four normal walks: one with coating, shoes and one without (left and right), and two normal walks with carrying a backpack variation (left and right). 32 subjects of the cohort are recorded in two sessions (January and April), adding clothes variation.

2) CNN Architectures: Table I summarizes the results yielded by gait recognition VS models, comparing deep convolutional ANNs with automatic feature extraction to shallow learning algorithms, where features are handcrafted. Deep learning models can be split into two groups, a single deep ANN and multiple deep ANNs joined in the last layer. The network inputs are single or a pair of processed silhouettes sequences. The latter case is mostly used for verifying individual's identity, with a view of 'probe and gallery' gait features. The 'probe' is an identified or verified subject, and the 'gallery' consists of templates as a browsing data set, where the probe is searched and matched to the closest instance in the gallery. These are examined in more detail below, for gait identification or verification.

a) Single deep ANNs: The single ANN input is a video sequence of images, on which the top *softmax* layer will perform classification based on the desired output for the given input. The *softmax* score outputs 1 for the true-match subject and 0 for false-match subjects. During validation, the loss is computed using cross-entropy between the *softmax* outputs and the corresponding desired output (the ground truth). Single CNN with a single input architecture has been investigated by a number of groups, with some examples outlined below.

TABLE I
RESULTS FOR GAIT RECOGNITION FROM VS

(A): Different scenes; (B): Different views
wQVTM: with Quality Measures View Transformation Model
LF+iHMM: Local binary pattern Flow + individual Hidden Markov Model
TCM: Transformation Consistency Measures
STIP: Space-Time Interest Points

Accuracy (%)	Model	Reference
70.51 (A)	wQVTM	Muramatsu et al. [51]
71.76 (A)	LF+iHMM	Hu et al. [52]
72.80 (A)	TCM+	Muramatsu et al. [53]
79.66 (A)	STIP	Kusakunniran et al. [37]
80.50 (A)	CNN	Zhang et al. [55]
85.51 (B)	CNN	Alotaibi et al. [15]
85.70 (A)	CNN	Alotaibi et al. [15]
90 (B)	CNN	Wu et al. [58]
90.45 (A)	CNN	Shiraga et al. [48]
92.50 (B)	RBM	Hossain et al. [36]
94.8 (B)	Parallel CNN	Wu et al. [58]
98.8 (A)	CNN	Wu et al. [58]
95.0 (A)	CNN	Yan et al. [20]
95.04 (A)	CNN	Li et al. [61]
97.35 (A)	3D-CNN	Wolf et al. [49]
98.8 (B)	Parallel +Triplet CNNs	Takemura et al. [54]

Yeoh *et al.* [47] used a CNN model trained on a single input as GEI. For testing, the *softmax* classifier in the last layer based on Euclidean distance is replaced by a Support Vector Machine (SVM) classifier to compute one-vs-all (probe vs gallery). The model, evaluated on OU-ISIR Treadmill dataset, yielded competitive performance in clothing-invariant for the identification of people.

Yan *et al.* [20] proposed a CNN model with Multilayer Perceptron (MLP) classifier. The input is a single GEI for automatic extraction of gait features. The CASIA-B dataset is used for evaluating the methods. The model is trained using multitask learning to predict multiple human attributes. 95.88% accuracy for each task is achieved; however, it was realized that the changes of scenes or view could be general- ized better by training on more data.

Shiraga *et al.* [48] designed GEINet, which is a CNN with two sequential groups. The network input is a single GEI image (from OU-ISIR database) in the training stage. In the testing stage, the dissimilarity between a probe GEI and gallery GEI pair is computed using the distance between them at the fully connected layer. The model performs well on cross-view for gait verification and identification.

Wolf *et al.* [49] proposed a 3D CNN with a 3D spatio-temporal tensor as input, consisting of a grey-scale image for the first channel and optical flow for the second and third channels. The model is trained and tested using the CASIA-B dataset, MoBo database and UFS database. The approach was evaluated on variations in walking speed, clothing and the view angle. Based on this architecture, Castro *et al.* [50] used a spatiotemporal 3D tensor of the optical flow as the input of the CNN. The network was trained and tested using the TUM-GAID database with gait scenarios, clothing and carrying variations for each subject. Although the network accuracy was significantly improved using the optical flow rather than using silhouette-based input. However, it is difficult to generalize on which feature extraction method outperformed

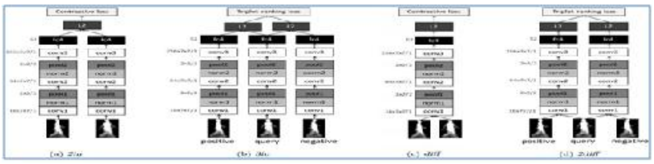

Fig. 3. [54]: High-level difference architectures for small view-angle differences: a) Siamese CCN with probe and gallery input; b) Triplet CNN with positive probe, negative probe and gallery; Low-level difference variants of a) and b) for substantial view-angle differences: c) single CNN with probe and gallery;
d) Siamese CCN.

the other, since [49] and [50] are evaluated on different datasets. Nevertheless, it is clear that the optical flow feature can present robust gait spatiotemporal information for use in a CNN architecture.

b) Dual deep ANNs: The input into a dual network consists of two different images, as probe and gallery under similar conditions; however, different gait scenarios, viewing angle, as well as clothes and carrying conditions, may be involved. This architecture is effective in gait verification since the networks have the same weight and structure, which allow the extraction of gait features automatically in the same manner. The outputs are matched using contrastive loss to find the Euclidean distance. The latter can be compared to a threshold to identify matching pairs or to label an imposter if a match cannot be found. Below is an outline of architectures applied for CNNs with two inputs. Figure 3 [54] presents some dual architectures used for verification and identification.

Zhang *et al.* [55] designed a shared parameters ‘Siamese twin’ CNN, each twin comprising a convolutional layer, a max-pooling layer and three fully connected layers to extract gait features automatically. The two twin outputs are connected to the contrastive loss layer. A pair of similar or dissimilar GEI images from the OU-ISIR database are used as an input to the Siamese network. In the training stage, the weights are shared simultaneously to optimize the network, and the model is fine-tuned by back-propagating with a contrastive loss. The gallery member with the nearest training sample is identified by testing to allow the feature metric computation of a discriminative loss function. The latter drives the similarity metric [56] to be small for pairs representing the same subject, and large for different subjects. Considering the changes of cross-view in real-world human identification scenarios, the model performs well in gait verification.

Wu *et al.* [57] proposed a CNN to extract gait features directly from the raw silhouettes’ sequence for cross-view gait recognition. Gait sequences from the CASIA-B dataset are used to train and test the network. In the testing stage, the Euclidean distance is measured for similarity using the probe and gallery method, achieving an accuracy of 94.1%. Furthermore, in [58] several CNN that take two inputs as probe and gallery have been shown to outperform other approaches, including twin CNNs [55], [57]. Two GEI images are used for gait verification based on cross-view gait recognition. The dataset to train and test the proposed networks are the CASIA-B dataset, OU-ISIR database and USF database. The proposed methods outperformed the previous state-of-the-art methods by a significant margin in the three datasets.

For cross-view gait recognition, Takemura *et al.* [54] consid- ered different architectures for verification and identification. This is based on the assumption that the absolute similarity scores are important for the verification task, while the relative similarity scores between a probe and the galleries are impor- tant for the identification task. For verification, a Siamese CNN with shared parameters is proposed (see figure 3a) to discrimi- nate whether two inputs originate from the same subject or not, based on the contrastive loss value. For identification, three parallel CNNs are deployed as a triplet network (see figure 3b). The triplet input is three GEIs: a query (the probe subject), a positive (from the same subject) and a negative (from a different gallery member). A triplet ranking loss is defined as the difference between two feature vector distances: the distance between positive and query and the distance between negative and query. The parameters of the triplet CNN are trained so that the dissimilarity between a probe and the same subject is relatively lower than that between a probe and different subjects. To accommodate possible substantial differences in the GEIs by viewing angle, low-level difference structures are introduced, as they are more directly affected by

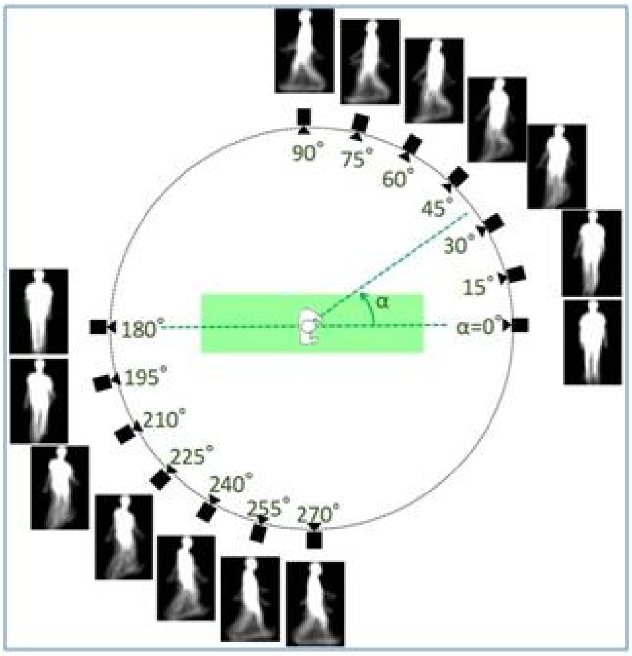

Fig. 4. Gait GEI images at 14 viewing angles [54].

appearance differences due to taking the difference between a matching pair closer to the input level (see figure 3c) and figure 3d). Cross-view gait recognition is demonstrated on OU-ISIR and OU-ISIR Multi-View Large Population datasets, with 10,307 subjects’ video sequences captured from 14 angles (see figure 4); however, the existing methods are difficult to evaluate on this dataset, and OU-ISIR LP is utilized to confirm the hypothesis regarding the network architecture.

3) Transfer Learning: Transfer learning is a comparatively new concept in ANNs and is the next strongest driver, after supervised learning, of the commercial success of machine learning [59]. Essentially, it is applying knowledge gained to solve a problem to a multiplicity of related problems. ‘Pre-trained’ models are beneficial as a starting point on specific ANN solutions, given the vast computing and time resources required to develop detailed physical models on these problems. Compared to starting from scratch, Transfer learning allows a substantial jump in the starting point for the delivery of a related ANN model [60].

Li *et al.* [61] used supervised pre-training of a VGG-D CNN (Visual Geometry Group) model and evaluated the efficacy of learned features on gait recognition tasks. The network consists of 16 convolutional layers and 3 fully connected layers with a nearest neighbor classifier. The silhouette images from the OU-ISIR dataset are used to train and test the network without fine tuning to capture gait spatiotemporal aspects. The probe and gallery method is used to identify people in a cross-view setting, significantly outperforming prior state- of-the art methods for both verification and identification.

Alotaibi and Mahmood [15] determined empirically the appropriate CNN architecture for automatic gait feature extraction from GEI images using the CASIA-B dataset. They applied two transfer learning methods to the network
pre-trained with 24 subjects. ‘Fine-tuned CNN’ involved adding one more subject (new total of 25 subjects) and dropping the weights of the *softmax* layer followed by re-training of the entire model; ‘re-learn *softmax* only’ involved ‘freezing’ the weights of the convolutional layers and the weights of the *softmax* layer were re-learned. While the computational time for pre-training was 124.82 s, adding a single subject by fine-tuned CNN took 42.41 s and only 22.12 s by *softmax* re-learning.

B. *Wearable Sensors*

WS are an obvious means to acquire human gait due to their convenience, efficiency and lower price. Unlike other gait capturing systems, WS impose upon the user to coop-erate wearing the device in a non-invasive way to provide gait signals. The advances in electronic devices and signal processing techniques have extended the applications of WS sensors to produce a measurement of human body orientation, position and specific force in space and time. The inertial measurement unit (IMU) is a type of WS system that has been extensively used due to its small size, cost, light weight, and good precision characteristics. A typical IMU provides the most widely used combination of sensing modalities to capture human activities, including gait. It comprises of an accelerometer, a gyroscope and often a magnetometer, which gives the heading direction. Additional components such as batteries, microprocessors and communication modules are arranged to jointly operate an IMU system.

Gyroscope sensors measure the angular velocity as the rate of change of the sensor’s orientation, while accelerometer sensors measure the acceleration of the body resulting from the acting forces in the opposite direction. A combination of these sensors can create a comprehensive report on the human body orientation, gravitational forces, velocity and acceleration [5].

Furthermore, it has been found convenient to use the gyroscope and accelerometer, usually integrated in a smart-phone, benefiting from predictable availability and positioning, as well as eliminating the need for additional hardware. Mobile users’ authentication is an acceptable approach when other gait authentication is not deployable. In the health- care domain, IMU-equipped smartphones allow inexpensive prediction of falls due to neurological disorders or freezing of gait in patients [62]. The computing power on-board of a smartphone can be used as a standalone system to perform all tasks required for decision making and com-municating with healthcare providers in any life-threatening situation.

The analysis of WS signals is a challenging task con-sidering the large number of observations recorded per unit time. This is due to the spatiotemporal nature of the gait cycle and the difficulty to relate in a straightforward manner WS signals to a known gait characteristic. Manual feature extraction is the classical way for gait analysis using WS, and it is time-consuming and depends on knowledge of the context in which the signals are acquired. Since perfor- mance is key in real world applications, deep learning has emerged as a promising data processing method by extracting

TABLE II
RESULTS FOR GAIT RECOGNITION FROM WS

Reference	Model	Accuracy (%)	system	feature
Zebin et al. [64]	CNN	**97.01**	Single IMU sensors	Activity + gait (A)
Ordóñez et al. [65]	DeepConvLSTM	95.8	Five IMU sensors	Activity + gait (A)
Gadeleta et al. [66]	CNN+SVM	**FP, FN <1**	Single IMU sensors	Gait cycle (A)
Gadaleta et al. [67]	CNN+SVM	FP, FN <0.15	Smartphone	Gait cycle (A)
Zhao et al. [68]	CNN	61.0	Smartphone	Gait cycle (A)
Dehzangi et al. [70]	CNN	**97.06**	Five IMU sensors	Gait cycle (A)
Lorenzi et al. [71]	ANNs	-	Single IMU sensors	Gait irregular steps PD (B)
Camps et al. [72]	CNN	90.6	Single IMU sensors	Gait cycle (B)
Kim et al. [73]	CNN	91.8	Smartphone	Gait cycle FOG (B)
Xia et al. [74]	CNN	90.60	Three IMU sensors	Gait cycle FOG (B)
Rad et al. [76]	CNN	0.65 ± 0.01	Three IMU sensors	Gait cycle FOG (B)
Murad et al. [79]	RNN	94.1	Three IMU sensors	Gait cycle FOG (B)
Ravi et al. [77]	Deep ANNs	95.8	Three IMU sensors	Gait cycle FOG (B)
Hannink et al. [80]	CNN	**0.97 CC**	Two IMU sensors	Gait stride length (B)
Alcha et al. [82]	ConvLSTM	94	Single IMU sensors	Gait cycle (B)
Hu et al. [83]	LSTM	**97.0**	Single IMU sensors	Gait on uneven and flat surface (B)

A: Normal gait; B: Abnormal gait; CC Spearman correlation

automatically reliable discriminative features of human gait, outperforming the approaches based on handcrafted features. Table II summarizes the results yielded by gait recognition models based on WS using various deep ANN models.

The sensor position on the body and the number of sen- sors comprising the system are an essential factor for the quality of the harvested data. In a systematic review analy- sis, Panebianco *et al.* [63] assessed accuracy and repeatability using 17 algorithms for their ability to monitor temporal parameters of human gait from 5 IMUs: one on the back, two pairs on the shanks and two pairs on the feet. For estimates of stance time, algorithms based on the acceleration of the shank and foot perform better than those based on the lower back; however, the sensor position did not affect the step estimation. For toe-off and heel strike detection, algorithms estimating angular velocity performed better overall, with notable dependence on the sensor positioning. Analysis has concerned mostly with the distinction between normal and abnormal gait, as follows below.

1) Normal Gait Analysis: Analysis of normal gait para- meters using WS has immensely attracted the interest of researchers and clinicians. The following are different methods and techniques that have been proposed and implemented for various applications.

Zebin *et al.* [64] proposed a system comprising 5 IMU sensors, worn on the lower back, thighs and shanks, for activity recognition including gait. A CNN based model is used to extract the features automatically from time-series raw data and achieve higher accuracy compared to the handcrafted features with shallow learning. In another work, 7 IMUs posi- tioned on the chest, arms and legs along with the12 accelerom- eters close to the limb joints, were used by Ordóñez and Roggen [65]. A DeepConvLSTM model is trained in a fully- supervised manner on human activities including gait. The DeepConvLSTM model outperforms previous results on the same dataset. However, increasing the number of sensors exacerbated the extraction of gait features compared to the use of WS attached to the pelvis and lower limb only [64].

For gait authentication, Gadaleta *et al.* [66] used a CNN model (see figure 5) to extract gait features from a single WS placed on the shank for each subject. Data from 15 subjects' gait is used in the training stage and 9 in the testing stage. In the latter, the network weights are frozen, and the CNN model is used to extract features, further the features are feed to SVM for classification. Thus, increasing the training dataset was suggested for improving the model performance. In a later work by Gadaleta and Rossi [67], the proposed CNN model is used to extract gait feature vector from a single subject automatically, the gait feature are used to train a single-class SVM. The system can distinguish between an impostor and the user whose gait is used for training. The IMU signals acquired from smartphones are tested on a user against 14 impostors, yielding false positive and false negative rates less than 0.15%. Zhao and Zhou [68] proposed a CNN model for gait labeling and authentication. The input to the network for automatic gait features' extraction is an Angle- Embedded Gait Dynamic Image (AE-GDI), which is a transformation of a WS data series. This allowed comparison with the state-of-the- art performance on VS (OU-ISIR) and WS (MCGILL [69]) datasets.

Similar to [64], Dehzangi *et al.* [70] placed 5 WS at various body locations. WS signals obtained from the sensors at chest, right wrist, knee and ankle, as well as the lower back of the subject, allows the study of CNN performance on time- frequency image transformation of raw signals. A total of 10 subjects' gait data were used to train and test the network; accounting for the multi-sensor character of the data, early and late fusion methods were applied, achieving state- of-the-art in subject identification. The deep learning approach to sensor fusion is addressed in more detail in Section VI.

2) Abnormal Gait Recognition: Deviations from normalgait are extensively studied by WS, the main targets being to classify neurodegenerative conditions, or to prevent falls in older adults. While the assumptions underlying various algo- rithms differ, in practical applications it often appears more convenient to use a single WS for capturing a discriminative

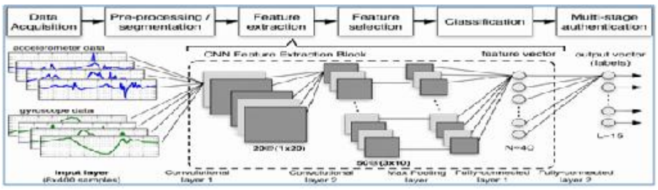

Fig. 5. Convolutional neural network to extract and classify gait features from wearable inertial measurement unit with accelerometer and gyroscope sensors [66].

gait feature. The sensor system embodiments used for abnor- mal gait analysis can be grouped into dedicated IMU systems and smartphones. Lorenzi et al. [71] used a single IMU unit positioned on the head, to collect gait patterns during the gait cycle, aiming to distinguish normal gait from the freezing of gait and irregular steps in Parkinson's disease (PD), using dynamic time warping to select the input features to the ANN. Deep learning recommended itself as an improved approach to recognizing the abnormality in human gait, in terms of classification accuracy and computational requirements. Camps *et al.* [72] used a waist-positioned IMU and an 8-layers CCN to achieve an accuracy of 90.6% to detect freezing of gait (FOG) detection in PD patients. The optimal architecture implemented with two convolution layers and 20 convolution filters. The gait of 32 patients was recorded by a smartphone accelerometer and gyroscope casually placed in the subject's trouser pocket. The CNN detected the FOG events in Fourier space with 91.8% accuracy, which is slightlyhigher than the CNNs methods proposed in [72].

In a recent study, Xia *et al.* [74] proposed a CNN to extract gait features from three accelerometers positioned above the hip, knee, and ankle. Against the aim to dis- tinguish FOG events from normal gait, evaluation on the Daphnet FOG dataset [75] from 10 subjects yielded an accu- racy of 90.60%. Several other deep ANNs [76], [77] have been trained and tested for human activity recognition from raw spatiotemporal datasets, including the FOG dataset used in [75]. Rad *et al.* [76] and Hammerla *et al.* [78] used a CNN performing well in human activity recognition; however, the performance on the FOG dataset was weaker. Murad and Pyun [79] improved the FOG recognition accuracy to 94.1% with their proposed deep RNN trained on the Daphnet FOG dataset. Ravì *et al.* [77] argued that deep learning models do not perform well when a small number of activity segments are available and proposed feature fusion, where shallow features are fused with features derived by deep learning in the fully connected and the *softmax* layers. With Daphnet FOG data, this method yielded for 'freeze' and 'no freeze' precision of 67.89% and 97.40%, as well as recall of 59.52% and 98.15%, respectively.

As an alternative use of deep learning, stride length esti- mates are derived in clinical settings to indicate, an early or further progression stage of neurological disorders. In the work reported by Hannink *et al.* [80], stride length is estimated automatically using WS and deep CNNs. The WS set consists of a 3D-accelerometer and a 3D-gyroscope attached below each ankle joint. The aim of this approach is to extract spatiotemporal gait parameters to aid the physician in scor- ing gait impairment objectively. The CNN performance was evaluated on the eGAIT dataset [81], using 10-fold cross validation on three different stride types. It was observed that the performance was dependent on stride definition and the better results were achieved for mid-stance to mid-stance intervals. Importantly, the CNN analysis of WS data was not affected by the use of a four-wheeled walking aid, where the data processing became problematic with the GAITRite walkway sensor system (see Section IV. C.).

Gait analysis using WS has been extensively studied for the detection of falls in older adults. Most of the reported work is based on handcrafted features deep learning is appeared as an improved approach in terms of increased classification accuracy and reduced computational load. Aicha *et al.* [82] reported work on CNN, LSTM, and ConvLSTM models used to extract gait features from raw accelerometer signals positioned on the lower back. The model trained and tested on 296 participants' gait to predict fall risk as the main task and user identity as an auxiliary task. The models' performance with features extracted using deep learning was observed to be marginally better compared to handcrafted features.

Hu *et al.* [83] attempted to capture the higher risk of falling while walking on uneven surfaces as compared to the flat surfaces walk. Essential here is the ability of subjects, as a function of age, to produce the stability required to avoid a fall. A single IMU unit positioned on the trunk delivered raw signals from 35 users: 17 older adults (age: 71.5 4.2 years) and 18 young adults (age: 27.0 4.7 years) used as input to the LSTM network. Automatically extracted spatiotemporal gait parameters are used to classify age-related differences in walking on flat or uneven surfaces.

IV. CONCLUSIONS

The character of gait statistics poses the trouble of figuring out facets appropriate for gait classifications, suited in a wide variety of software areas. The three gait-sensing modalities included in this Review have produced records which is most amenable to the use of deep learning, to tackle the automated extraction of such features. Deep studying CNNs usually outperform shallow getting-to-know fashions in the most vital metrics. Furthermore, multi-sensor and multi-modality fusion outcomes in higher accuracy and robustness. This is executed by using the on-hand flexibility in information representations, ANN architectures, and the preference for mannequin hyper-parameters. Gait evaluation benefits from strategies brought and examined for different purposes of deep learning. However, it requires specific interest due to its spatiotemporal character, the selections for ubiquitous gait sensing, and the privacy issues they raise, as properly as the fee of attaining research, improvement, and commercialization objectives. Deep gaining knowledge of multi-sensor, multi-modality gait records provides new preferences in the robust power in the direction of customized healthcare, as nicely as in the direction of extra strong and un-intrusive biometrics for security and security. These are some of the challenges of the day, however, the state-of-the artwork indicates a promising step in achieving similarly into the future, alternatively than simply the contemporary horizon.

REFEEERENCS

[1] W. Schiehlen and D. García-Vallejo, “Walking dynamics from mech- anism models to parameter optimization,” in *Proc. Symp. Hum. Body Dyn.*, vol. 2, 2011, pp. 199–211.

[2] Aristotle. *On the Gait of Animals, 350BC*. Accessed: Oct. 8, 2018.

[3] M. H. Pope, “Giovanni Alfonso Borelli—The father of biomechan- ics,” *Spine*, vol. 30, no. 20, pp. 2350–2355, 2005.

[4] M. W. Whittle, *Whittle’s Gait Analysis*, 5th ed. Amsterdam, The Netherlands: Elsevier, 2012, p. 30.

[5] W. Weber and E Weber, “Üeber die mechanik der menschlichen gehwerkzeuge, nebst der beschreibung eines versuchs über das herausfallen des schenkelkopfs aus der pfanne im luftverdünnten raume,” *Annal. Phys. Chem.*, vol. 116, no. 1, pp. 1–13, 1837.

[6] P. Connor and A. Ross, “Biometric recognition by gait: A survey of modalities and features,” *Comput. Vis. Image Understand.*, vol. 167, pp. 1–27, Feb. 2018.

[7] V. L. S. Profeta and M. T. Turvey, “Bernstein’s levels of movement con- struction: A contemporary perspective,” *Hum. Movement Sci.*, vol. 57, pp. 111–133, Feb. 2017.

[8] P. R. Cavanagh and M. A. Lafortune, “Ground reaction forces in distance running,” *J. Biomech.*, vol. 13, no. 5, pp. 397–406, 1980.

[9] H. Elftman and J. T. Manter, “The axis of the human foot,” *Science*, vol. 80, no. 2082, p. 484, 1934.

[10] D. H. Sutherland, “The evolution of clinical gait analysis Part III—Kinetics and energy assessment,” *Gait Posture*, vol. 21, no. 4, pp. 447–461, 2005.

[11] M. J. Major, P. Raghavan, and S. Gard, “Assessing a low-cost accelerometer-based technique to estimate spatial gait parameters of lower-limb prosthesis users,” *Prosthetics Orthotics Int.*, vol. 40, no. 5, pp. 643–648, 2015.

[12] R. Camicioli, D. Howieson, S. Lehman, and J. Kaye, “Talk- ing while walking: The effect of a dual task in aging and Alzheimer’s disease,” *Neurology*, vol. 48, no. 4, pp. 955–958, 1997.

[13] D. Gouwanda and S. M. N. A. Senanayake, “Emerging trends of body- mounted sensors in sports and human gait analysis,” in *Proc. 4th Kuala Lumpur Int. Conf. Biomed. Eng. (IFMBE)*. Springer, vol. 2, 2008, pp. 715–718.

[14] H. Lee, S. J. Sullivan, and A. G. Schneiders, “The use of the dual-task paradigm in detecting gait performance deficits follow- ing a sports-related concussion: A systematic review and meta- analysis,” *J. Sci. Med. Sport*, vol. 16, no. 1, pp. 2–7, 2013.

[15] M. Alotaibi and A. Mahmood, “Improved gait recognition based on specialized deep convolutional neural network,” *Comput. Vis. Image Understand.*, vol. 164, pp. 103–110, Nov. 2017.

[16] G. Qian, J. Zahang, and A. Kidane, “People identification using floor pressure sensing and analysis,” *IEEE Sensors J.*, vol. 10, no. 9, pp. 1447–1460, Sep. 2010.

[17] A. Muro-De-La-Herran, B. Garcia-Zapirain, and A. Mendez-Zorrilla, “Gait analysis methods: An overview of wearable and non-wearable systems, highlighting clinical applications,” *Sensors*, vol. 14, no. 2, pp. 3362–3394, 2014.

[18] E. Abdulhay, N. Arunkumar, K. Narasimhan, E. Vellaiappan, andV. Venkatraman, “Gait and tremor investigation using machine learning techniques for the diagnosis of Parkinson disease,” *Future Gener. Comput. Syst.*, vol. 83, pp. 366–373, Jun. 2018.

[19] G. Bastas, J. J. Fleck, R. A. Peters, and K. E. Zelik, “IMU-based gait analysis in lower limb prosthesis users: Comparison of step demarcation algorithms,” *Gait Posture*, vol. 64, pp. 30–37, Jul. 2018.

[20] C. Yan, B. Zhang, and F. Coenen, “Multi-attributes gait identifi- cation by convolutional neural networks,” in *Proc. 8th Int. Congr. Image Signal Process. (CISP)*, Shenyang, China, 2016, pp. 642–647.

[21] F. Battistone and A. Petrosino, “TGLSTM: A time based graph deep learning approach to gait recognition,” *Pattern Recognit. Lett.*,

[22] J. Suutala, S. Pirttikangas, J. Riekki, and J. Röning, “Reject-optional LVQ-based two-level classifier to improve reliability in footstep identi- fication,” in *Proc. 2nd Int. Conf. Pervasive Comput.*, 2004, pp. 818–821.

[23] A. Mostayed, S. Kim, M. M. G. Mazumder, and S. J. Park, “Foot step based person identification using histogram similarity and wavelet decomposition,” in *Proc. Int. Conf. Inf. Secur. Assurance (ISA)*, Apr. 2008, pp. 307–311.

[24] S. C. White and D. A. Winter, “Predicting muscle forces in gait from EMG signals and musculotendon kinematics,” *J. Electromyography Kinesiology*, vol. 2, no. 4, pp. 217–231, 1992.

[25] S. L. Di Stasi, D. Logerstedt, E. S. Gardinier, and L. S. Mackler, “Gait patterns differ between ACL-reconstructed athletes who pass return-to-sport criteria and those who fail,” *Amer. J. Sports Med.*, vol. 41, no. 6, pp. 1310–1318, 2013.

[26] E. R. H. P. Isaac, S. Elias, S. Rajagopalan, and K. S. Easwarakumar, “Multiview gait-based gender classification through pose-based voting,” *Pattern Recognit. Lett.*,

[27] I. Birch, T. Birch, and D. Bray, “The identification of emotions from gait,” *Sci. Jus.*, vol. 56, no. 5, pp. 351–356, 2016.

[28] O. Simeone, “A very brief introduction to machine learning with applications to communication systems,” *IEEE Trans. Cogn. Com- mun. Netw.*, vol. 4, no. 4, pp. 648–664, Dec. 2018.

[29] Y. LeCun, Y. Bengio, and G. Hinton, “Deep learning,” *Nature*, vol. 521, pp. 436–444, May 2015.

[30] I. Goodfellow, Y. Bengio, and A. Courville, “Deep learning,” in *Deep Learning*. Cambridge, MA, USA: MIT Press, 2016, ch. 2, pp. 326–415.

[31] N. Ketkar, “Convolution neural network and recurrent neural networks,” in *Deep Learning With Python*. New York, NY, USA: Springer, 2017, ch. 5, pp. 63–79.

[32] N. Yi, C. Li, X. Feng, and M. Shi, “Research and improvement of convolutional neural network,” in *Proc. IEEE/ACIS 17th Int. Conf. Comput. Inf. Sci. (ICIS)*, Jun. 2018, pp. 637–640.

[33] S. Ruskaa, W. Hämäläinen, S. Kajava, M. Mughal, P. Matilainen, and J. Mononen, “Evaluation of the confusion matrix method in the validation of an automated system for measuring feeding behaviour of cattle,” *Behavioural Processes*, vol. 148, pp. 56–62, Mar. 2018.

[34] J. Preis, M. Kessel, M. Werner, and C. Linnhoff-Popien, “Gait recog- nition with Kinect,” in *Proc. 1st Workshop Kinect Pervasive Com- put.*, 2012, pp. 1–4. [Online]. Available: https://www.martinwerner. de/pdf/GAITKINECT.pdf

[35] B. Gálai and C. Benedek, “Feature selection for lidar-based gait recognition,” in *Proc. Int. Workshop Comput. Intell. Multi- media Understand. (IWCIM)*, Oct. 2015, pp. 1–5.

[36] E. Hossain and G. Chetty, “Multimodal feature learning for gait biometric based human identity recognition,” in *Neural Information Processing—ICONIP* (Lecture Notes in Computer Science), vol. 8227,M. Lee, A. Hirose, Z. G. Hou, and R. M. Kil, Eds. Berlin, Germany: Springer, 2013, pp. 721–728.

[37] W. Kusakunniran, “Recognizing gaits on spatio-temporal feature domain,” *IEEE Trans. Inf. Forensics Security*, vol. 9, no. 9, pp. 1416–1423, Sep. 2014.

[38] M. Jeevan, N. Jain, M. Hanmandlu, and G. Chetty, “Gait recognition based on gait pal and pal entropy image,” in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2013, pp. 4195–4199.

[39] C. Wang, J. Zhan, J. Pu, X. Yuan, and L. Wang, “Chrono-gait image: A novel temporal template for gait recognition,” in *Computer Vision—ECCV* (Lecture Notes in Computer Science), vol. 6311. Berlin, Germany: Springer, 2010, pp. 257–270.

[40] C. Wang, J. Zhang, L. Wang, J. Pu, and X. Yuan, “Human identification using temporal information preserving gait template,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 34, no. 11, pp. 2164–2176, Nov. 2012.

[41] R. Gross and J. Shi, “The CMU motion of body (MoBo) database,” Tech. Rep. CMU-RI-TR-01-18, Robotics Institute, Carnegie Mellon Univ. Pittsburgh, PA, USA, 2001. Accessed: Dec. 8, 2018.

[42] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, andK. W. Bowyer, “The humanID gait challenge problem: Data sets, performance, and analysis,” *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, pp. 162–177, Feb. 2005.

[43] *CASIA Gait Database*. Accessed: Dec. 8, 2018. [Online].

[44] Y. Makihara *et al.*, “The OU-ISIR gait database comprising the treadmill dataset,” *IPSJ Trans. Comput. Vis. Appl.*, vol. 4, pp. 53–62, Apr. 2012.

[45] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi, “The OU-ISIR gait database comprising the large population dataset and performance eval- uation of gait recognition,” *IEEE Trans. Inf. Forensics Security*, vol. 7, no. 5, pp. 1511–1521, Oct. 2012.

[46] M. Hofmann, J. Geiger, S. Bachmann, B. Schuller, and G. Rigoll, “The TUM gait from audio, image and depth (GAID) database: Multimodal recognition of subjects and traits,” *J. Vis. Commun. Image Represent.*, vol. 25, no. 1, pp. 195–206, Jan. 2014.

[47] T. Yeoh, H. E. Aguirre, and K. Tanaka, “Clothing-invariant gait recognition using convolutional neural network,” in *Proc. Int. Symp. ISPACS*, Oct. 2016, pp. 1–5.

[48] K. Shiraga, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, “GEINET: View-invariant gait recognition using a convolutional neural network,” in *Proc. Int. Conf. Biometrics (ICB)*, Jun. 2016, pp. 1–8.

[49] T. Wolf, M. Babaee, 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.

[50] F. M. Castro, M. J. M. Jiménez, N. Guil and N. P. D. L. Blanca “Automatic learning of gait signatures for people identification,” in *Advances in Computational Intelligence*, I. Rojas *et al.*, Eds. IWANN 2017, Part II, *Lecture Notes in Computer Science*, vol. 10306. Cham, Switzerland: Springer, 2017, pp. 257–270.

[51] D. Muramatsu, Y. Makihara, and Y. Yagi, “View transformation model incorporating quality measures for cross-view gait recognition,” *IEEE Trans. Cybern.*, vol. 46, no. 7, pp. 1602–1615, Jul. 2016.

[52] M. Hu, Y. Wang, Z. Zhang, D. Zhang, and J. J. Little, “Incre- mental learning for video-based gait recognition with LBP flow,” *IEEE Trans. Cybern.*, vol. 43, no. 1, pp. 77–89, Feb. 2013.

[53] D. Muramatsu, Y. Makihara, and Y. Yagi, “Cross-view gait recogni- tion by fusion of multiple transformation consistency measures,” *IET Biometrics*, vol. 4, no. 2, pp. 62–73, 2015.

[54] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, "On input/output architectures for convolutional neural network-based cross-view gait recognition," *IEEE Trans. Circuits Syst. Video Technol.*, to be published.

[55] C. Zhang, W. Liu, H. Ma, and H. Fu, "Siamese neural network based gait recognition for human identification," in *Proc. IEEE Int. Conf. Acoust., Speech Sig. Process. (ICASSP)*, Mar. 2016, pp. 2832–2836.

[56] S. Chopra, R. Hadsell, and Y. LeCun, "Learning a similarity metric discriminatively, with application to face verification," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2005, pp. 539–546.

[57] Z. Wu, Y. Huang, and L. Wang, "Learning representative deep features for image set analysis," *IEEE Trans. Multimedia*, vol. 17, no. 11, pp. 1960–1968, Nov. 2015.

[58] Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, "A comprehensive study on cross-view gait based human identification with deep CNNs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 2, pp. 209–226, Feb. 2016.

[59] K. Weiss, T. M. Khoshgoftaar, and D. Wang, "A survey of trans- fer learning," *J. Big Data*, vol. 3, no. 1, pp. 1–40, Dec. 2016.

[60] S. J. Pan and Q. Yang, "A survey on transfer learning," *IEEE Trans. Knowl. Data Eng.*, vol. 22, no. 10, pp. 1345–1359, Oct. 2010.

[61] C. Li, X. Min, S. Sun, W. Lin, and Z. Tang, "DeepGait: A learn- ing deep convolutional representation for view-invariant gait recog- nition using joint Bayesian," *App. Sci.*, vol. 7, no. 3, p. 15, 2017.

[62] A. Schrag, M. Choudhury, D. Kaski, and D. A. Gallagher, "Why do patients with Parkinson's disease fall? A cross-sectional analysis of possible causes of falls," *NPJ Parkinson's Disease*, vol. 1, Jun. 2015, Art. no. 15011.

[63] G. P. Panebianco, M. C. Bisi, R. Stagni, and S. Fantozzi, "Analysis of the performance of 17 algorithms from a systematic review: Influence of sensor position, analysed variable and computational approach in gait timing estimation from IMU measurements," *Gait Post.*, vol. 66, pp. 76–82, Oct. 2018.

[64] T. Zebin, P. J. Scully, and K. B. Ozanyan, "Human activity recog- nition with inertial sensors using a deep learning approach," in *Proc. IEEE SENSORS*, Oct./Nov. 2016, pp. 1–3. doi

[65] F. J. Ordóñez and D. Roggen, "Deep convolutional and LSTM recur- rent neural networks for multimodal wearable activity recognition," *Sensors*, vol. 16, no. 1, p. 115, 2016.

[66] 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.

[67] M. Gadaleta and M. Rossi, "IDNet: Smartphone-based gait recogni- tion with convolutional neural networks," *Pattern Recognit.*, vol. 74, pp. 25–37, Feb. 2018.

[68] Y. Zhao and S. Zhou, "Wearable device-based gait recognition using angle embedded gait dynamic images and a convolutional neural network," *Sensors*, vol. 17, no. 3, p. 478, 2017.

[69] J. Frank, S. Mannor, J. Pineau, and D. Precup, "Time series analysis using geometric template matching," *IEEE Trans. Pat- tern Anal. Mach. Intell.*, vol. 35, no. 3, pp. 740–754, Mar. 2013.

[70] O. Dehzangi, M. Taherisadr, and R. ChangalVala, "IMU-based gait recognition using convolutional neural networks and multi-sensor fusion," *Sensors*, vol. 17, no. 12, p. 2735, 2017.

[71] P. Lorenzi *et al.*, "Smart sensors for the recognition of specific human motion disorders in Parkinson's disease," in *Proc. 6th Int. Workshop Adv. Sensors Interfaces (IWASI)*, Jun. 2015, pp. 131–136.

[72] J. Camps *et al.*, "Deep learning for freezing of gait detection in Parkinson's disease patients in their homes using a waist-worn iner- tial measurement unit," *Knowl.-Based Syst.*, vol. 139, pp. 119–131, Jan. 2018.

[73] H. B. Kim *et al.*, "Validation of freezing-of-gait monitoring using smartphone," *Telemed. E-Health*, vol. 24, no. 11, pp. 899–907, 2018.

[74] Y. Xia, J. Zhang, Q. Ye, N. Cheng, Y. Lu, and D. Zhang, "Evalua- tion of deep convolutional neural networks for detection of freezing of gait in Parkinson's disease patients," *Biomed. Signal Process. Control*, vol. 46, pp. 221–230, Sep. 2018.

[75] M. Bachlin *et al.*, "Wearable assistant for Parkinson's disease patients with the freezing of gait symptom," *IEEE Trans. Inf. Technol. Biomed.*, vol. 14, no. 2, pp. 436–446, Mar. 2010.

[76] N. M. Rad, T. Van Laarhoven, C. Furlanello, and E. Marchiori, "Novelty detection using deep normative modeling for IMU-based abnormal movement monitoring in Parkinson's disease and autism spectrum disorders," *Sensors*, vol. 18, no. 10, p. 3533, 2018.

[77] 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.

[78] N. Y. Hammerla, S. Halloran, and T. Plotz, "Deep, convolutional, and recurrent models for human activity recognition using wearables," in *Proc. 25th Int. Joint Conf. Artif. Intell. (IJCAI)*, S. Kambhampati, Ed., 2016, pp. 1533–1540.

[79] A. Murad and J.-Y. Pyun, “Deep recurrent neural networks for human activity recognition,” *Sensors*, vol. 17, no. 11, p. 2556, 2017.
[80] J. Hannink *et al.*, “Mobile stride length estimation with deep convolutional neural networks,” in *Proc. IEEE J. Biomed. Health Inform.*, vol. 22, no. 2, pp. 354–362, Mar. 2018.
[81] A. Rampp, J. Barth, S. Schülein, K.-G. Gaßmann, J. Klucken, andB. M. Eskofier, “Inertial sensor-based stride parameter calculation from gait sequences in geriatric patients,” *IEEE Trans. Biomed. Eng.*, vol. 62, no. 4, pp. 1089–1097, Apr. 2015.
[82] A. N. Aicha, G. Englebienne, K. S. Van Schooten, M. Pijnappels, and B. Kröse, “Deep learning to predict falls in older adults based on daily-life trunk accelerometry,” *Sensors*, vol. 18, no. 5, p. 1654, 2018.
[83] B. Hu, P. C. Dixon, J. V. Jacobs, J. T. Dennerlein, andJ. M. Schiffman, “Machine learning algorithms based on signals from a single wearable inertial sensor can detect surface- and age-related differences in walking,” *J. Biomech.*, vol. 71, pp. 37–42, Apr. 2018.