

A comprehensive survey on deep-learning based gait recognition for humans in the COVID-19 pandemic

Md Shohel Sayeed, Ibrahim Bin Yusof, Mohd Fikri Azli bin Abdullah, Md Ahsanul Bari,
Pa Pa Min

Centre for Intelligent Cloud Computing, Faculty of Information Science and Technology, Multimedia University, Melaka, Malaysia

Article Info

Article history:

Received Sep 3, 2022

Revised Nov 29, 2022

Accepted Dec 26, 2022

Keywords:

Convolutional neural network

COVID-19 pandemic

Deep learning

Gait analysis

Gait energy image

Gait recognition

Recurrent neural networks

ABSTRACT

Human gait recognition is a biometric technique that has been utilized for security purposes for the last decade. Gait recognition is an appealing biometric modality that aims to identify individuals based on the way they walk. The outbreak of the novel coronavirus (COVID-19), has spread across the world. The number of people infected with COVID-19 is rising rapidly throughout the world. Even though some vaccines for this pandemic have been developed to minimize the effects of COVID-19, deep learning-based gait recognition techniques have shown themselves to be an effective tool for identifying the individuals wearing face mask in COVID-19 pandemic. These techniques play an important part in reducing the rate of COVID-19 spreading throughout the world in the context of the COVID-19 pandemic. Deep learning methods are currently dominating the state-of-the-art in gait recognition and have fostered real-world applications. The main objective of this paper is to provide a comprehensive overview of recent advancements in gait recognition with deep learning, including datasets, test protocols, state-of-the-art solutions, challenges, and future research directions. The purpose of this discussion is to identify current challenges that need to be addressed as well as to suggest some directions for future research that could be explored.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Md Shohel Sayeed

Centre for Intelligent Cloud Computing, Faculty of Information Science and Technology

Multimedia University

75450 Melaka, Malaysia

Email: shohel.sayeed@mmu.edu.my

1. INTRODUCTION

In recent years have shown the rise of a new coronavirus known as COVID-19, which was caused by the severe acute respiratory syndrome coronavirus 2 (SARS-CoV-2). This virus has caused a pandemic around the world [1]. Both number of people who have been infected as well as the mortality rate is rising rapidly. Since the time when this article is being written, it has been reported that more than 108,000,000 persons have been infected with COVID-19, the number of death cases is around 2,400,000, and the number of patients who have recovered is approximately 80,000,000 [2]. The widespread transmission of COVID-19 has resulted in the quarantining of a significant section of the world's population and the destruction of a broad variety of industrial sectors, both of which have contributed to the current state of global economic instability. Fever, dry cough, myalgia, dyspnea, and headache are some of the most usual symptoms of the new coronavirus [3].

In recent decades, a subfield of biometric identification known as gait recognition has emerged. Using the intrinsic patterns of movement in individuals to detect them, this subfield focuses on detecting them based on their physical characteristics, such as their size and relationships with the trunk and the limbs, as well as space-time information about their intrinsic movement patterns [4]. Gait recognition is a subfield of biometric

identification. Such a strategy would be of great use in situations involving surveillance systems or foggy settings. For example, biometric monitoring, which often involves the use of distinguishing characteristics for biometric identification methods such as fingerprints and facial features are either difficult or impossible to differentiate.

In addition, techniques based on gait recognition offer certain advantages over other biometric identification models since it is difficult to hack such designs. The challenge stems mostly from the features that are inherent to the idea itself, namely the identification of the notion based on the silhouette and its movement, both of which are notoriously difficult to reproduce. This does not really hold true for other techniques, such as when a person may manipulate the system by, for example, hiding their face. In addition, it is also possible to detect gait patterns for models that do not require high resolution photos or specialized devices in order to be correctly identified, such as fingerprints and iris scans, which do not need high-resolution photos. In addition, techniques of gait recognition do not require the individual being analyzed to cooperate with the identification system [5]. This contrasts with other approaches which require the person to be identified to work together with the system.

Similarly, gait information can easily be obtained from a distance, which is a huge advantage over other techniques when it comes to collecting gait data, particularly in situations where the identification of the person being analyzed does not require the assistance of the person being analyzed, such as in criminal investigations [6]. Also, these methods can be used to get the data without having to use expensive, specialized hardware. This can often lower the overall cost. The process of getting them out has become easier as surveillance technologies and smartphones with accelerometers have become more common. This makes it easy to get data signals out of these devices.

Identification of people based on how they walk, and move is not at all an easy process, even though the instruments discussed above are quite straightforward. The intricacy of the job causes the standard gait recognition approaches, which entail data pre-processing and manually collected characteristics for subsequent recognition, to have a lot of limits and issues [7]. For instance, they need to be able to perceive the whole gait as well as huge intra-class variances, occlusions, and shadows, and they need to be able to find the body segments. These are just some of the requirements. In recent years, a trend in machine learning known as deep learning has emerged as a revolutionary tool for solving problems related to image and sound processing, computer vision, and speech processing. This tool has been shown to perform significantly better than virtually any baseline that had been established up until that point. The new prevailing model eliminates the need for manual extraction of representative characteristics from seasoned professionals. Aside from that, it gives the best results for gait detection, getting rid of the problems that were there before and opening the door for more research.

Several studies have been published in recent years summarizing the main achievements in biometrics and gait recognition. Biometric identification has made significant progress over the past few decades in general, as demonstrated by [8]-[10]. On the other hand, discussed by [11] the novelties regarding brain biometrics. In both [12], [13] provided a comprehensive survey that covered general topics in regards to gait recognition. Meanwhile, a study published by [14] examined the recognition of gaits based on vision, while proposed a similar concept for smart visual surveillance [15]. Furthermore, a summary of the work with regard to the gait-based re-identification of individuals was published by [16] in 2019. In contrast, it has been investigated by [17] how wearable sensors could be used to detect gait movements. Despite this, none of the approaches to gait recognition developed over the past few years have focused on the main advances regarding deep learning-based approaches that have made an impact on this field in recent years.

As more research is being conducted on biometrics and gait recognition in recent years, a growing number of authors have contributed papers that give a summary of the most significant advancements made in each subject [18]. According to the survey, the majority of the works used convolutional neural networks (CNNs) [19] as well as long short-term memory (LSTM) [20], [21] some novel and relevant architectures are also being left behind, such as autoencoders [22], capsule networks (CapsNet) [23], deep belief networks [24], and generative adversarial networks (GANs) [25]. As a result, the following are the primary goals of this study which shows the most current and major publications with the most effective tactics and methods based on deep learning for gait recognition aim.

To give the reader with a thorough and illustrated theory on this topic investigation of its biometric recognition antecedents and disclosure of the most prevalent gait feature extraction methods and architectures utilized to address the linked problems include boundaries, as well as datasets that may be utilized for gait recognition, including images, classifications, and descriptions as well. Among the publications that were considered for examination, we focused our attention on gait recognition employing deep learning approaches, such as convolutional neural networks, LSTM, and so on. Deep learning techniques were substituted by the structures themselves. Filter number three a result of the work's invention, comparable methodologies and designs were published in publications that weren't discarded. The research that was used in this study was published no more than five years ago [26].

2. METHODS

This section introduces the concept of the theoretical framework of the deep learning algorithms used for gait identification, as well as the gait recognition issue. As a part of this section, we will provide an overview of some of the most popular deep-learning architectures that have been used in the past to recognize gaits, as well as some options that may be available in the future. Furthermore, several approaches to recognizing individuals by their biological characteristics have been proposed and discussed as well.

2.1. The use of deep learning approaches considered in the recognition of gaits

Over the past few years, traditional machine learning strategies and gait recognition algorithms have produced relatively satisfactory results in terms of recognizing patterns in walking, it is generally true that these methods are limited to features that are handcrafted and do not have the ability to discover intrinsic patterns within data. This is the case even though such strategies have provided relatively satisfactory results in the past. In this regard, deep learning-based algorithms have emerged as an elegant option for solving image/video and sequence difficulties, among other challenges; in addition, they have shown themselves as a potent tool for gait detection. Hence, the purpose of this section is to provide an overview of some of the most popular deep learning architectures that have been used in the past to recognise gaits. These include recurrent neural networks (RNN), CNN, autoencoders, capsule networks, generative adversarial networks and deep belief networks (DBN) [27].

2.1.1. Convolutional neural networks

Since the beginning of the 2010s, CNNs have gained an extraordinary degree of popularity [28]. They have been indispensable in the process of resolving issues that are associated with image processing, such as image classification and segmentation. The fundamental building blocks of a CNN are called convolutional neurons, and they typically consist of either 3 X 3 or 5 X 5 kernels. These kernels are tasked with the responsibility of executing convolutional operations on the input data. In a nutshell, the conclusion of this procedure results in the production of a new group of matrices, which are then included in the model as successive layers. In the field of signal processing, the term "convolution" refers to the act of multiplying two signals in order to produce a third signal. The concept is shown in Figure 1.

Therefore, a CNN may be seen as a stack of convolutional kernels that generates the input for consecutive layers. It is important to take note that the procedure may include the inclusion of certain intermediary pooling layers, and that a fully linked layer is coupled at the top of the architecture for the sake of categorization. The learning process is carried out via an algorithm called backpropagation, which works in conjunction with a gradient descent computation. The architecture of a CNN is shown in Figure 2, which was influenced by LeNet-5.

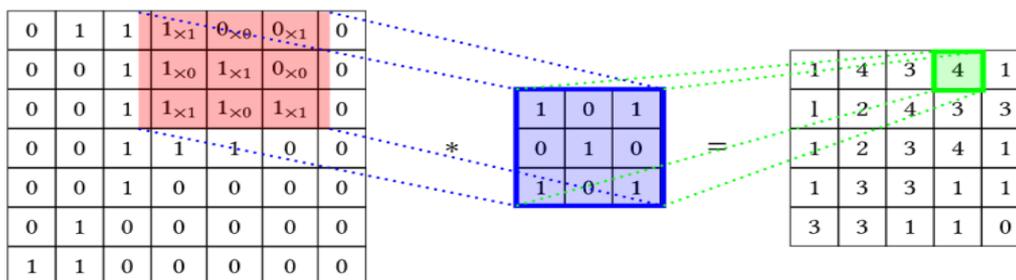


Figure 1. An example of 3X3 convolution

2.1.2. Capsule networks

CNNs are particularly effective at interpreting visual characteristics; yet, they tend to become confused by the spatial connections between many oddities due to their complexity. In other words, a CNN that has been trained is typically, the system is capable of recognizing dogs if it is able to find a dog's body, face, and tail, even if the dog is arranged in a different sequence or if its parts are located in different areas of an image. This is true even if the parts of the dog are located in different parts of the image. However, in order to solve this issue, the authors of capsule networks propose using a hierarchical strategy [29]. In a nutshell, the model has a structure that consists of two layers. The first component is referred to as a convolutional encoder, and it is responsible for the identification of low-level characteristics. There is also the second kind of linear decoder which is fully connected and employs an algorithm called routing by agreement to address such low-

level information to the appropriate spot in a hierarchy with a higher level of information [30]. This type of decoder is also known as a fully connected linear encoder. Cabinets are thus more resistant to the effects of object orientation. In addition, they may do better when it comes to distinguishing several things or objects that overlap in a scene.

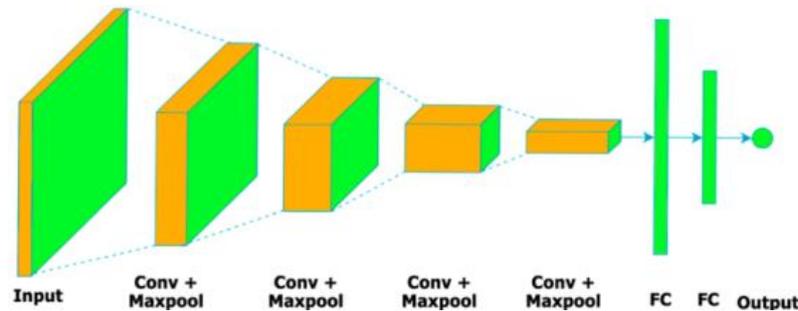


Figure 2. Example of a standard CNN architecture

2.1.3. Recurrent neural networks

The subject of gait identification has been approached in a number of different studies as a series of pictures characterizing an individual's movement. RNNs are a typical way that may be used to recognize such a strategy when employing deep learning. These networks calculate each neuron's activation by taking into consideration the information from the input data as well as the output of other neurons in a recurrent manner. As an option, it may be possible to integrate this architecture with a CNN in order to get additional information about the input pictures as a whole so that the inference can be carried out accordingly. Owing to the fact that defining a person's gait often involves a significant quantity of sequential information, a certain group of RNNs known as gated RNN are better ideal for the job due to their capacity to cope with extended sequences. When discussing this topic, it is possible to make reference to two primary designs, namely the gated recurrent unit (GRU) and the long-short term memory [31].

2.1.4. Long-short term memory

The lengthy-short term memory was initially built in with the primary goal of improving outcomes on long datasets. This was done with the intention of saving time. In a nutshell, LSTMs function in a manner that is similar to that of the conventional RNN; specifically, the output of a particular neuron is determined by the recurrent information gleaned from the results of neurons that came before it. The primary distinction may be seen in the architecture of the LSTM cell, which is characterized by more intricate and intricate interactions. A structure like this one is made up of three primary gates that are responsible for controlling the flow of information. These gates are as follows:

- The forget gate: In order to decide how much information should be stored, this gate is called "the forgetting gate," since it reduces the amount of information that is stored. The data from the prior state and the present state are fed into a sigmoid function, which produces values between 0 and 1 depending on the input. The closer you are to 1, the more knowledge you will retain. A new value is computed by the input gate in order to bring the present concealed state up to date. The input gate takes into account two primary values: i) the relevance of the previously concealed state is determined by a sigmoid function; and ii) the original value is sent to a hyperbolic tangent (tanh) function, which is responsible for squishing this value somewhere between 1 and 1. The current concealed state is defined by the product of these two variables multiplied together.
- The output gate: After determining the value of the forget and the input gates, the output gate is what determines the value of the cell's output. The procedure is carried out in the following manner: i) the values from the forget gate and the input gate are added together and then sent to a function; ii) the prior state of the cell is sent to a sigmoid function in order to determine its current state; and iii) Sigmoid function's output will be multiplied by the output from the subsequent function in order to produce the output of the cell.
- Gated recurrent unity: As a recurrent neural network, the gated recurrent unit was originally envisioned as an optimal means of improving the results of neural machine translations by enhancing outcomes from neural machine learning. A GRU is very similar to an LSTM in that it has internal gates that regulate data flow through it. The primary distinction between the two types of models is the number of gates that are available in each; for example, a GRU has only two gates, which are known as the forget and the output

gates, where as an LSTM has three gates. Studies have shown that the GRU can achieve similar results to those achieved by the LSTM, despite the fact that the GRU uses fewer gates than the LSTM does. In this manner, the GRU offers the advantage of receiving a reduced computational burden and a faster performance when both training and inference tasks are considered.

2.1.5. Autoencoders

Autoencoders are a kind of generated neural network that are often used for the purposes of data reduction and picture denoising. The model may be broken down into two primary stages, which are as follows:

- Encoder: This component is in charge of encoding the information that is provided as input into a typically more condensed feature space.
- Decoder: This component is responsible for the reconstruction of encoded data in an unsupervised manner. Figure 3 is a depiction of the model's architectural layout.

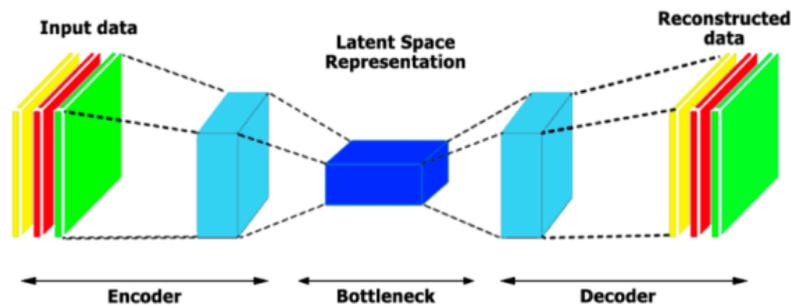


Figure 3. Example of a standard autoencoder architecture

2.1.6. Deep belief networks

Stochastic neural networks suited for generative tasks are called deep belief networks. These networks are structured in such a way that each layer represents a greedy-fashioned trained restricted boltzmann machine (RBM). In a nutshell, RBM is a graphical model that is made up of a visible set of units and a latent set of units, which are referred to as the visible layer and the hidden layer, respectively. These units are connected to one another by a weight matrix, and there is no connection between neurons that are located on the same layer. Attaching a set of input data to the model's visible layer and searching for a representation of that data in the model's hidden units are the two steps that make up the learning technique.

In addition, the training phase is carried out by reducing the amount of energy used by the system. This is often accomplished by using a Markov chain technique in conjunction with Gibbs sampling for the sake of optimization. In order to perform classification tasks efficiently, the most common approach is to couple a SoftMax layer on top of a DBN architecture and, after greedily pre-training all of the RBMs, fine-tune the weights using backpropagation, for example, it might be necessary to adjust the weight matrices and to fit the labels correctly for proper identification. This is considered to be the most effective method.

2.1.7. Generative adversarial networks

In recent years, generative adversarial networks have gained a lot of popularity thanks to their remarkable capacity to produce realistic synthetic pictures. Two separate networks are used in this model. One network is a generator, which learns how the distribution of the data is distributed and generates synthetic samples to be analyzed, and the second network is a discriminator, which works to determine whether a particular instance was created naturally or artificially. The generator is responsible for learning the distribution of the data and producing synthetic samples. The generator and the discriminator engage in an antagonistic relationship in which they compete with one another, with the generator attempting to create samples that are convincing enough to mislead the discriminator. On the other hand, the discriminator becomes better and better at identifying these phone photos. The illustration of the model may be seen in Figure 4.

2.1.8. Summary of the deep learning techniques

There is a summary in Table 1 which provides an overview of the deep learning techniques considered in this paper. The Table 1 also provides how the deep learning techniques are frequently used in the context of recognizing gait patterns. The same work may at some point be used by a variety of architectures, resulting in its appearance in the table more than once.

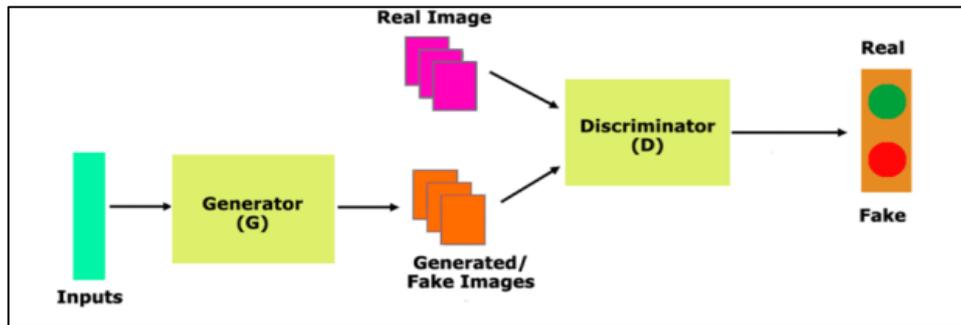


Figure 4. Example a standard GAN architecture

Table 1. Gait recognition approaches based on deep learning, organized according to the type of neural network

“Technique	Task
Convolutional neural	Networks (CNN) of a video.
Auto encoder	Works by compressing and decompressing features from the input.
Capsule	Improve the semantic organization of the outputs from a CNN.
Deep belief networks	Encode features and patterns into compressed representations.
Generative adversarial networks	A training method that relies on the differentiation of an original input and a generated counterpart from a model, such as a CNN.
Recurrent neural networks	Comprise both GRUs and LSTMs, which are composed of several gates to control the flow of information and are employed to deal with temporal information.”

2.2. Gait recognition for human

A number of approaches of recognizing persons based on their biological traits have been discussed up to this point. There is no doubt that these methods are reliable and safe, as evidenced by their success in governance systems of the banking sector and the public sector, there are two primary drawbacks that need to be emphasized: i) the provision of personal biometric information by passive means is a requirement of these systems. In order for the recognition of the individual, the required information must be provided or registered by the individual; and ii) It is important to note that some of these systems require specialized equipment in order to operate. In spite of these limitations, the methods are considered to be reliable and safe, as evidenced by their success in these areas.

Gait recognition models might be an alternate solution to address such problems, particularly when taking video-based techniques into consideration. It is also important to note that, in most cases. The process of acquiring the biometric information relies on only a camera without any special features or any non-evasive sensors and due to the passive nature of the process of collecting such information, a great deal of power and privacy are compromised, these methods do not experience the issues that have been outlined above.

As a result, the individual who was observed is no longer needed to help with the identifying process or offer any information at this point. Nevertheless, video-based recognition is not the only option; despite the fact that it is the approach that makes the most sense, it is not the only one. Identifying gait may also be accomplished by the use of different features, such as the footprint, which takes into account the pressure and size of the region. The expense involved in implementing the technology that is essential to get such functionality is undeniably one of the models' most notable drawbacks [32].

In addition, the approaches that have been developed for gait identification may be separated into two primary categories: those that are template-based and those that are not. The aim of template-based methods is to determine whether the trunk is moving or whether the legs are moving; more specifically, these methods are based on the dynamics of the movement through space, or they may be based on a spatio-temporal perspective [33]. There are a number of techniques that can be employed to extract information from walking paths, gait information images, and gait energy images, among which canonical correlation analysis (CCA) and joint sparsity models (JSM) can be utilized. On the other hand, approaches that are not dependent on templates consider the shape and its qualities as the most important traits; this means that the identification of a person is accomplished via the use of measures that represent the shape.

The following pieces of equipment may be found in the scene: sidewall cameras, then a multi-spectrum photographic sensor. During the stroll, the connection between space and time is measured using sensors that monitor velocity and gyroscopes that are housed on the roof. The gait cycle is finished off with a sensor that looks like a carpet and is embedded into the floor. This sensor is used to measure the pressure that the foot exerts on the ground. Figure 5 depicts the many types of sensors and devices that may be used in the

gait information collecting process. These are shown in the preceding sentence. The image depicts an environment similar to a laboratory that has been designed with several embedded devices for the purpose of gait data acquisition.



Figure 5. An example of gait data collecting setting

Among these devices, there are cameras, the most commonly used device for this purpose because it is quite accessible and allows for the acquisition of data over a greater distance, as well as multispectral photographic sensors, which are also very effective, but due to their high cost, they are less widely accessible than other types of cameras. On the roof, the picture also consists of gyroscopes and velocity sensors, both of which are responsible for characterizing the acquisition of spatio-temporal context. In the last phase, a step-pressure carpet-like sensor is used to measure the walking rhythm [34].

This sensor is very useful for detecting gait irregularities, such as those caused by Parkinson's disease and other neurodegenerative conditions. Deep learning-based techniques have brought about a paradigm change in the area of gait detection, which has led to the achievement of significant outcomes in a number of applications. Therefore, the next part provides an introduction to an in-depth discussion of numerous studies that were reviewed within the framework of systems for gait identification that are based on deep learning [35].

2.3. Discussion

An overview of the studies that were surveyed and provided in this part can be found in Table 2. The Table 2 provides details of the gait recognition using deep learning approaches based on the research articles published year, model used, input type, dataset used, obtained results and measure interms of identificate rate/accuracy/prediction error/ root-mean-square error (RMSE). In this section, we are going to explore in more depth how deep learning approaches can be used to recognise gait patterns.

Convolutional neural networks are, in general, the most common option when contemplating deep learning solutions, particularly with image- and video-based difficulties, such as gait recognition. This tendency is to be anticipated given that CNNs have achieved exceptional results in a variety of applications and have won the majority of benchmark competitions over the last several years. Despite this, the other designs each make a significant contribution to the area by performing more effectively than the others in a variety of particular jobs. For example, capsule networks have the ability to extract partial gait representations in a hierarchical form. This enables them to provide superior outcomes in scenarios in which persons or items are displayed in the scene with different orientations or overlapped. In a similar vein, recurrent neural networks are indispensable when it comes to dealing with sequential data such as movies, and as a result, they offer themselves as an indispensable instrument for the gait identification technique described here.

Other data sources, such as accelerometers, gyroscopes, sensor-based features in general, as well as handcrafted features, have been shown in many studies to be able to produce reliable results, as well, despite the fact that the vast majority of deep learning-based approaches to gait recognition use images or videos as their source data. The majority of these research focus on unsupervised methods of deep learning throughout the process. Examples of such methods are autoencoders and DBNs, both of which tend to be more expressive across various data sets. On the other hand, CNNs are essential when working with raw picture and video files. These

unsupervised approaches are able to extract information on the data distribution and change it, often in a lower-dimensional space. As a result, they are able to provide gait recognition with more representative characteristics.

Table 2. Approaches to the recognition of gait based on deep learning

Ref.	Year	Model	Input Type	Dataset	Result	Measure
[36]	2016	CNN	GEI	OU-ISIR	94.6%	Identification rate
[37]	2017	CNN	Cross-view	CASIA B	90.8%	Accuracy
[38]	2017	CNN + Joint Bayesian	Sensors	OU-ISR	97.6%	Identification rate
[39]	2017	CNN	Optical flow	TUM-GAID and CASIA B	97.52%	Accuracy
[40]	2017	CNN Siamese networks	Cross-view	OU-ISIR	98.8%	Accuracy
[41]	2017	CNN + Nearest Neighbor	Optical Flow	TUM-GAID, CASIA B, and OU-ISIR	99.8%	Identification rate
[42]	2017	Autoencoders + PCA	GEI	CASIA B and SZU RGB	97.58%	Identification rate
[43]	2018	CNN + LSTM	Accelerometer and Gyroscope	whuGAIT and OU-ISIR	99.75%	Accuracy
[44]	2018	GAN	PEI	PEI OU-ISIR, CASIA B and USF	94.7%	Accuracy
[45]	2018	DBN	Sensors	Data collected by the authors	93%	Accuracy
[46]	2019	Capsule	LBC and MMF	OU-ISIR	74.4%	Accuracy
[47]	2019	CNN	GEI + Data Augmentation	CASIA B	98%	Accuracy
[48]	2019	Autoencoders + LSTN	Cross- and Frontal-view	CASIA B, USF, and FVG	99.1%	Accuracy
[49]	2019	Autoencoders + LSTM	GEI	OU-ISIR and CASIA B	96.15%	Accuracy
[50]	2019	LSTM	Sensors	Data collected by the authors	0.02	Prediction error
[51]	2019	GAN	GEI	CASIA A and CASIA B	82%	Recognition result
[52]	2020	LSTM	GEI	CASIA B and OU-ISIR	99.1%	Recognition rate
[53]	2020	CNN	Cross-view	OU-MVLP, OU-LP, and CASIA B	98.93%	Identification rate
[54]	2020	GAN	Cross-view	OU-MVLP and CASIA B	93.2%	Identification rate
[55]	2020	Capsule	Multi-scale representations	CASIA-B and OU-MVLP	84.5%	Identification rate
[56]	2020	DBN	Sensors	Data collected by the authors	2.61	RMSE
[57]	2021	LSTM	IMU	whuGAIT and OU-ISIR	94.15%	Accuracy
[58]	2021	Capsule	Sensors	Several	99.69%	Accuracy"

It is worth noting that generative adversarial networks are used to describe a specific scenario where gait systems can gain knowledge of a variety of features, such as orientations, clothing, the number of people present, and so on, since they can create synthetic data to train the models. This allows gait systems to learn a wider range of features than was previously possible. In addition to this, they are helpful for analysing frauds in gait-based systems for the purpose of ensuring the system's security by creating phoney photographs for the purpose of testing. Gait energy image is one of the most common ways that are used to describe gait image data. This approach represents the sequence of a basic energy image cycle by utilising the weighted average method.

In addition, the sequences that make up a journey cycle are subjected to processing so that the binary silhouette may be. As a result, GEI preserves both the static and dynamic properties of human walking while greatly reducing the amount of computer effort required for picture processing. After doing a thorough investigation of the approach, one is able to identify a handful of features that are distinctive of the model which are as follows:

- A unique representation of the human walk is a key part of the presentation, which does not soften the context of vector images in any way.
- The motion of a human is captured in a single image while keeping temporal information at the same time.
- The GEI's are quite sensitive to silhouette noise when it comes to an individual photograph. In a similar vein, cross-view-based gait detection is a common method that is used to cope with many visual viewpoints. Due to the fact that the input type requires a number of completely controlled cameras as well as cooperating surroundings, it can only be used to real-world circumstances. In addition to this, before completing any combination, it visually normalizes the gait characteristics, which enables the model to understand the correlations between the various visual motions in the scene. An example of this method can be seen in the cross-view gait recognition system that was developed at Newcastle University in the United Kingdom. This system, known as DiGGAN, achieved state-of-the-art results on the largest Multiview gait dataset in the world (which included more than 10,000 individuals), and it did so by employing cross-view approaches [59].

In spite of the success that such approaches have achieved, there are still certain areas in which they are lacking, and it is these areas that other, less common ways try to address. For example, the Period Energy Image includes a multichannel gait template that was suggested to address the issue of view angle fluctuation caused by cross-view and GEI loss of temporal information [60]. Optical flow is still another strategy that has been shown in various studies. By means of this method, it can be determined exactly what items are seen within the scene, how they are arranged spatially, and what changes have taken place over the course of time [61]. In conclusion, for deep learning-based gait identification methods, a broad variety of hand-crafted features or sensor-based input data types are also used. Inertial measurement units, accelerometers, gyroscopes, sensor output, and other similar techniques are examples of the types of methods that fall under this category.

Regarding the current state of the art in gait recognition, there are primarily two different techniques that one may take: one is literature-based, and the other is representation-based. The literature suggests that two-dimensional convolutional neural networks (2D-CNNs) are the most popular form of deep neural network (DNN) for gait recognition by means of deep learning. In fact, approximately 50% of all solutions are just based on two-dimensional convolutional neural network architectures for classification. Following the most popular categories are those that include solutions using 3D-CNN and GAN, with each category accounting for 10% of the total published material. In addition, deep autoencoder (DAE), RNN, DBN, and graph convolutional networks get less considerations than other types of DNNs. These percentages equate to 5%, 3%, 2%, and 1%, respectively. However, hybrid methods account for 26% of the solutions. Within this category, CNN and RNN combinations are the most popularly used approach, accounting for approximately 15% of presence. A combination of DAE with GAN and RNN is responsible for 10% of the methods, followed by RNN-CapsNet methods that are responsible for 2% of solutions. When it comes to solutions that are state-of-the-art and are based on representation, silhouettes are the ones that are used the most often for gait identification; this accounts for more than 85% of all solutions. Skeletons, despite the fact that they represent just 10% of the potential options, have been studied less often in proportion to silhouettes. This is despite the fact that the approach is an encouraging one. There were also some approaches, which account for around 5% of the literature that is now accessible, that investigate the representations of the skeleton as well as the silhouette. It should be noted that these methods either use unraveled representation learning or score of fusion algorithms; depending on the case.

In order to comprehend gait data, one has to acknowledge that it is complex, arising from the interaction of many factors, such as occlusion/obstruction, camera points of view, pose of individuals, the sequence of events, motion of body parts, or light sources present in the data, whilst others may not necessarily be present in the data [62], is one of the primary challenges associated with deep learning-based gait recognition. These kinds of elements might interact in a convoluted manner, which makes the work of gait detection more difficult. At the moment, numerous approaches are available in various fields connected to pattern recognition, for example, face recognition, mood estimation, and stance estimation, and there is an increasing number of approaches being tested. These experts devote their time and energy to gaining an understanding of complex settings and isolating representations that categorise the different components that contribute to an explanation inside a high-dimensional space. However, the vast majority of gait identification systems that are now available rely on deep learning, and these methods have not yet been examined. As a result, they are unable to directly separate the underlying structure of the gait data into major discrete variables. There is still opportunity for advancement in the use of ambiguous context techniques in various gait recognition systems, in spite of the recent success that has been made in this area [63].

3. DATASETS

The training and assessment processes of machine learning models are dependent on a dataset that includes the topic of the task. This is the case regardless of the learning paradigm that is being used, which might be supervised, unsupervised, or any other paradigm. In addition, the use of such datasets enables the determination of how successful a technique is in the resolution of a particular issue, in addition to the comparison of that way to other potential solutions.

When it comes to gait recognition, the number of unique datasets that are available to do the job is quite limited. This is true whether one considers public or private solutions. The creation of new models of artificial intelligence that are capable of identifying humans based on how they walk, or move has been hampered by a paucity of training data of this kind. Gait biometry requires a decent quantity of movement records for a subject, which implies recording and making many videos for each individual. This is a challenge since it is time-consuming to record and generate all of these movies for each person. In addition to this, movies of this kind often have an inherent high dimensionality, which results in the expansion of the dataset and necessitates a large amount of storage space.

It is necessary to get consent from each participant before extracting biometric data and making it publicly available. It's possible that litigation will result from the eventual establishment of a dataset that doesn't have the official agreement of each person. There are a number of datasets that are frequently used for gait recognition-related activities that are included in the following section.

3.1. CMU MoBo dataset

The CMU MoBo dataset has a very limited quantity of data, which consists of various films for gait recognition taken from a total of 20 different individuals. The segmentation method is simplified as a result of the inclusion of a silhouette mask set as well as bounding box information inside the dataset. It is possible to download the dataset without making reservations or having to sign any forms of agreement; all that is required is a connection to the file transfer protocol (FTP) server at the University of Calgary. This is one of the most significant benefits offered by the dataset [64].

3.2. TUM GAID dataset

The TUM gait from audio, image, and depth (GAID) database is comprised of red, green and blue (RGB) pictures, audio, and depth data, which were initially captured from 305 individuals in three distinct variations. As the name implies, this database was built using GAID. In addition, 32 individuals were re-recorded in order to have some variety, bringing the total number of recordings to 3,370. The scientists claim that their dataset is the only one of its kind that enables recognition by merging video, audio, and depth information. The number of shows, the circumstances of carrying (with or without a bag), and the amount of time are all variables. The first record was created in the winter of 2012 (in the month of January), and it had 176 recordings that were made while wearing a jacket and snow boots. The second performance was place in April 2012 with 161 individuals participating, and since it was warmer at the time, they wore noticeably different clothing. Out of this total, 32 persons were documented at both points in time, which allows for variance in both fabric and time. The presence of a clear-cut assessment procedure is a significant plus point for using this dataset [65].

3.3. HID-UMD dataset

The human identification at a distance (HID)-UMD dataset is comprised of many movies of individuals walking that were collected from four distinct perspectives, as well as their associated binary masks for the foreground segmentation [66]. Its principal objective is to provide assistance to researchers in the development of novel gait and face biometrics. Figure 6 presents several picture samples seen from a variety of perspectives. The size of the whole dataset is roughly 2.2 gigabytes.



Figure 6. Several picture samples seen from a variety of perspectives

Furthermore, the dataset constitutes an aggregate consisting of two datasets described in the following way: Dataset 1 is comprised of walking sequences of 25 different persons in four separate poses such as i) frontal view of a person walking towards the camera; ii) frontal view of a person walking away; iii) frontal-parallel view of a person walking toward left; and iv) frontal-parallel view of a person walking toward right

[67]. The films included in Dataset 2 were taken by 55 different people as they walked down a T-shaped walkway. The sequences were captured using two cameras that were positioned in lines that were perpendicular to one another [68].

3.4. CASIA

The CASIA gait database is a collection of four datasets intended specifically for gait recognition applications [69]. It is made available by the Institute of Automation of the Chinese Academy of Sciences (CASIA), and its details are as follows:

- CASIA A: This dataset was developed in December 2001 and was formerly known as the NLPR Gait Database. It consists of 20 different people, and each of these people has 12 videos, which means that there are four videos for each of the three directions, which are parallel to the image plane, 45 degrees to the image plane, and 90 degrees to the image plane. In addition, the length of each visual sequence is unique to itself and varies depending on the individual's walking speed [70]. Figure 7 presents the example frames from CASIA A dataset viewed from each angle, with Figure 7(a) corresponding to the parallel view, Figure 7(b) corresponding to 90 degrees, and Figure 7(c) being an example of recognition algorithms viewed from 45 degrees.

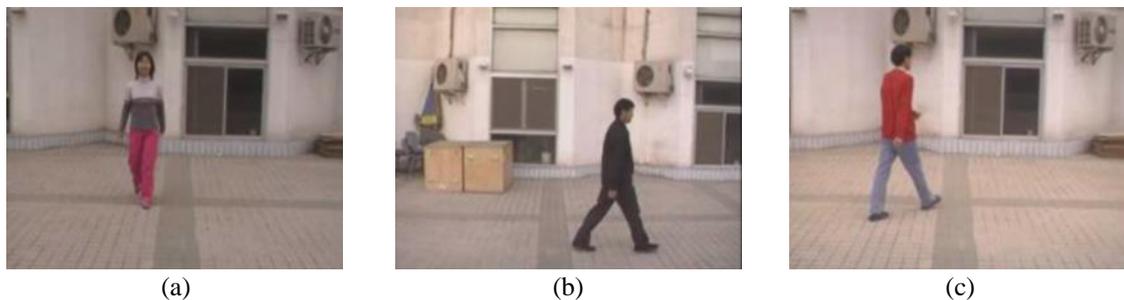


Figure 7. Example frames from CASIA A, (a) corresponding to the parallel view, (b) corresponding to 90 degrees, and (c) being an example of recognition algorithms viewed from 45 degrees

- CASIA B: The CASIA B was developed in 2005 and involves shooting 124 distinct people from a total of 11 different perspectives. Each scenario was performed three times, each time with a different alteration, such as a change in the dress or the pace at which the characters walked. In addition to that, the dataset contains a group of silhouettes that are supplied for each sequence. These silhouettes are used for foreground segmentation. Both Figure 8 and Figure 9 illustrate different angles and outfit choices [71].
- CASIA C: Acquired in 2005, the CASIA C dataset includes 153 subjects who were filmed by infrared cameras (thermal spectrum) while walking in one of four different variations: normal walking, slow walking, fast walking, or normal walking while carrying a backpack. Each of these variations was filmed by the infrared cameras [72]. Examples are shown in Figure 8.

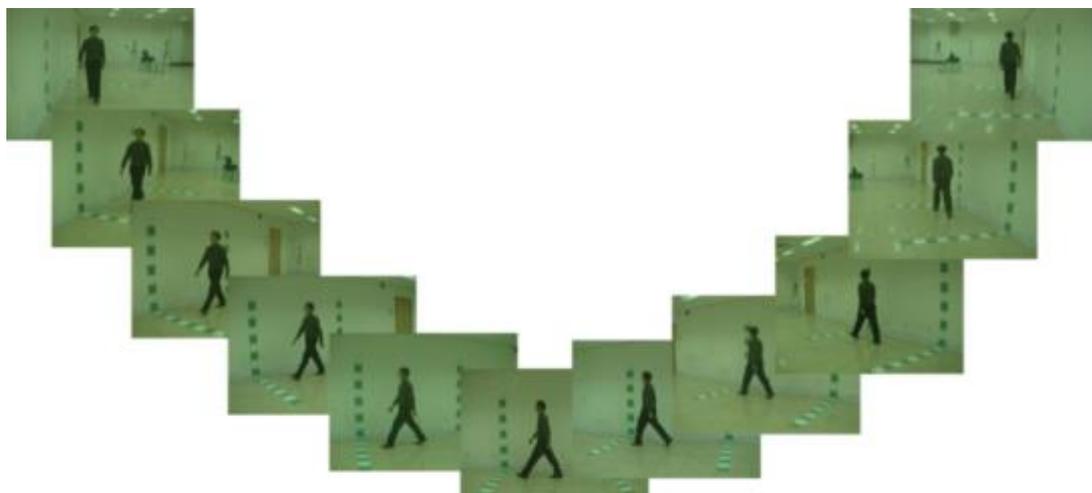


Figure 8. Example frames from CASIA B dataset

CASIA D: The CASIA gait-footprint dataset consists of photographs in addition to information on the cumulative foot pressure. The gait posture photographs total 3, 496, and there are 2, 658 cumulative images in the collection [73]. Figure 8 depicts the example frames from CASIA B dataset. Each image corresponds to one of the 11 angles that were considered [74].

The movement may be noticed in people's choice of attire, as well as in personal goods like bags and other accessories. The designs for the data collection are shown in Figure 9. For each person under each viewing angle, 10 videos are provided, 2 videos with a long coat shown in Figure 9(a), 2 videos with a bag shown in Figure 9(b), and 6 videos in normal styles shown in Figure 9(c). The user is not need to pay anything in order to get the silhouettes for datasets A, B, or C. these downloads are available to them. 5 When it comes to the method of data collection, potential workers are required to fill out a form and then wait for the authorization of the dataset's owner. Figure 10 depicts the example frames from CASIA C dataset. The images were obtained through an infrared camera at night and with variations in the manner of walking. Figure 10(a) shows walking without bag and Figure 10(b) walking with bag.

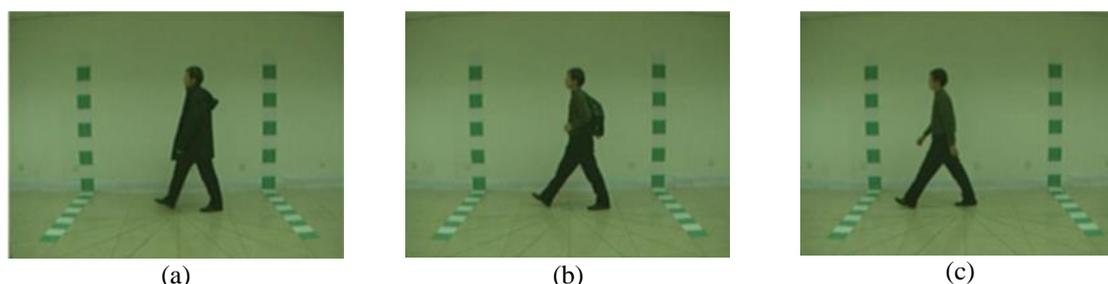


Figure 9. Example frames from CASIA B dataset, each image corresponds to one of the 11 comprised angles: (a) 2 videos with a long coat, (b) 2 videos with a bag, and (c) 6 videos in normal styles

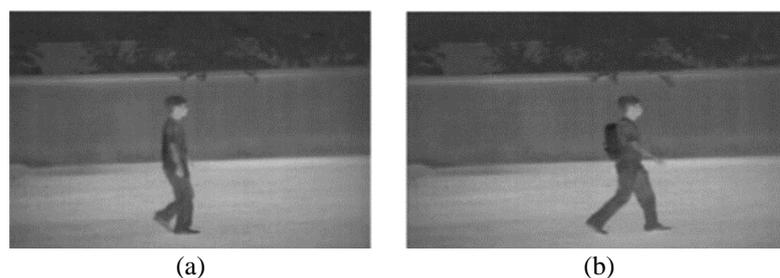


Figure 10. Example frames from CASIA C where: (a) shows walking without bag and (b) walking with bag

3.5. The OU-ISIR biometric database and some of its applications

Since 2007, researchers at Osaka University's Institute of Scientific and Industrial Research (ISIR) have been collecting the largest dataset for gait identification ever assembled anywhere in the world. A total of eight separate groups are working together on this project. They are illustrated below in alphabetical order.

Treadmill dataset: This collection is made up of sequences of individuals walking on electronic treadmills while being filmed by a total of 25 cameras at a rate of 60 frames per second with a resolution of 640 by 480. There are four distinct sections within it such as;

- a) Treadmill dataset A - Speed variation: This subset includes 34 people in a lateral view and has speeds that range from 2 to 10 km/h with 10 km/h intervals [75].
- b) Treadmill dataset B - Clothes variation: This dataset consists of 68 persons in lateral vision, each wearing one of 32 different outfits. The method of acquiring the photographs in synchronous fashion with the data on the footsteps [76].
- c) Treadmill dataset C - View: An expansive database that includes 168 individuals ranging in age from 4 to 75 years old. In addition to this, the subset is made up of 25 views that were obtained using a multi-view synchronous gait system [77]. Figure 10 depicts the example frames from CASIA C dataset.
- d) Treadmill dataset D - Gait fluctuation: This collection is made up of gait silhouette sequences with 185 participants. They were examined from a lateral viewpoint, and the subject's speed changed during the sequences [78]. Figure 11 depicts the example frames from CASIA D dataset.

Based on the high-speed and low-speed variations, the data were split into two groups, each consisting of 100 participants (with an overlap of 15 individuals).

Large population dataset: the large population dataset has been collected through various outreach activity events. A total of 4,016 subjects were filmed from a total of four different camera angles to produce the video. The video is filmed in a frame rate of 30 frames per second and has a resolution of 640 by 480 pixels [79].

Speed transition dataset: There are two subsets of the speed transition dataset [80], which can be found in the following order for your convenience:

- a) The scenes in Dataset A consist of persons walking at a constant speed of 4 kilometers per hour on a treadmill or on the ground. There are 179 of these scenarios. In this collection, the backdrop was covered up with wallpaper to eliminate any distracting elements.
- b) The sequences in Dataset B were collected from 25 individuals who walked at speeds ranging from 1 to 5 kilometers per hour while using a treadmill. Everyone is videotaped not once, but twice. The acceleration and deceleration are each carried out over the course of three seconds, and the middle sequences, each lasting one second, were taken from both processes.

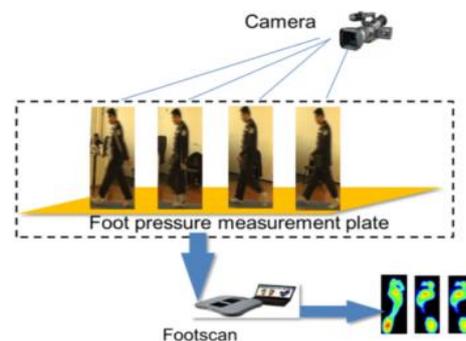


Figure 11. Example frames from CASIA D

Multi-view large population dataset: This dataset is made up of 10,307 samples, with 5,114 representing males and the remaining 5,193 representing women. The participants' ages span from 2 to 87 years old, and the dataset was designed for motion detection algorithms that use cross-vision. The photographs were captured using a camera that had a resolution of 1,280 by 980 pixels, a frame rate of 25 frames per second, and 14 distinct camera angles. The apparatus that was used for the capture was positioned at a lateral distance of 8 meters and at a height of 5 meters, respectively [81].

Large population dataset with bag: this dataset has been generated focused on the recognition of gait in relation to individuals carrying objects, with the aim of taking biometric information into account, and identifying the position of the part of the body that is being transported, along with the distance from the body. Specifically, the dataset examines people who are walking while carrying bags. The large population dataset with bag includes photographs taken by a camera placed at roughly 8 meters and at a height of 5 meters. The ages of the persons in this dataset range from 2 to 95 years old. The sequences were captured at a resolution of 1,280 by 980 pixels at a frame rate of 25 frames per second. Each person was videotaped three times, and although the first person, in this case AI, may or may not be holding an item, the second and third individuals do not bring anything into the frame. In the event that anything has to be carried, a total of four regions the front, the rear, the lower side, and the upper side are designated as carrying areas. Each movie is also provided with its own binary mask, which may be used to remove the backdrop if desired [82].

Large population dataset with age: the objective of this study was to evaluate gait recognition in relation to people's age and gender. The dataset includes the motion data of 62,846 people moving along a predetermined route while being photographed by cameras with a resolution of 640 by 480 pixels and a frame rate of 30 per second. All of the films include binary masks that were acquired after the removal of the backdrop, and the ages of the persons in the sequences range from 2 to 90 years [83].

Inertial sensor dataset: inertial sensor dataset is one of the largest data sets containing information about gait based on inertial sensors. The device was designed to test and evaluate the methods of identifying an individual by movements through the use of motion sensors and accelerometers and was designed for research and evaluation purposes as well. The database is made up of images collected from 744 subjects with ages ranging from 2 to 78 years old (389 males and 355 females) [84].

Actions that are comparable inertial dataset: the parallel courses of events the 460 individuals that make up the inertial dataset range in age from 8 to 78, and the gender distribution is quite close to being even

across the board. Their walking characteristics were measured alongside the data reported. In addition to that, this dataset also includes information on the floor, which includes invalid, flat, stair ascending, stair climbing, ramp climbing, and ramp descending [85].

3.6. The dataset from the University of South Florida

The dataset from the University of South Florida (USF) includes 1,870 sequences totaling 122 people. These subjects wore one of two distinct kinds of shoes. The dataset also takes into account people who are carrying a briefcase and those who are not carrying a briefcase, as well as different surface conditions like grass and concrete, and varied camera perspectives, such as left or right viewpoints. The films were shot in various locations outside at two distinct times to capture different moments in time [86].

3.7. Dataset from the Southampton

The southampton human ID at a distance (SOTON) database is a project that was undertaken by the University of Southampton. This database is comprised of three primary components. These three primary components are illustrated below:

- SOTON small database: consists of 12 participants that walked around an indoor track at varied speeds, wearing a variety of shoes and clothing, and either carrying or not carrying bags [87].
- SOTON large database: this database includes 114 participants who walked both indoors and outdoors, as well as on a treadmill and a track within the laboratory. Images were captured from six distinct camera vantage points and included in a compilation that had more than 5,000 individual segments [88].
- SOTON temporal: data for this study was collected using the multi-biometric tunnel, which is equipped with 12 synchronized cameras to record the stride of individuals throughout the course of the timeframe. The collection is made up of several dynamic scenes, each with its own backdrop, lighting, walking surface, and camera placement. There is a total of 25 people included in the dataset, with their ages spanning from 20 to 55 years old. There are 17 male subjects and 8 female subjects. Take note that none of the scenes feature anybody wearing shoes [89].

3.8. AVA dataset with multiple views for use in gait recognition (AVAMVG)

There is a database known as the AVA multiview dataset for gait recognition (AVAMVG) that was created specifically for algorithms that recognize gaits based on 3D measurements. This collection contains gait photographs from 20 actors displaying diverse trajectories, and it was created for use with gait recognition algorithms. After obtaining the sequences using cameras that had been precisely calibrated for the job, they were subjected to a post-processing stage in which 3D image reconstruction methods were used. In addition to that, a relevant binary silhouette that may be used for segmentation is shown for each sequence. In conclusion, the database stores 200 videos with six channels that may be seen simultaneously. These movies can also be used as 1,200 videos with a single view (that is, 6 times 200) [90]. Figure 12 illustrates various samples.



Figure 12. Example of the multiview dataset, this image portrays of people walking in different directions, from multiple points of view

3.9. Kyushu University 4D gait database (KY4D)

The Kyushu University 4D Gait Database (KY4D) is made up of successive 3D models and picture sequences of 42 persons walking over a total of six different trajectories, four of which are straight and two of which are curved [91]. The dataset A (Straight) consists of sequential 3D models and images of people walking along straight paths. Dataset B (Curved) is made up of image sequences of individuals walking along curved trajectories. Figure 13 depicts an example of the studio generated from the KY4D Gait Database. Figure 13(a) is an illustration of the trajectory, which is shown as a red arrow in the figure. Figure 13(b) illustrates numerous 3D reconstructed models.

Figure 13(a) and Figure 13(b) depicts the example of the studio generated from the KY4D Gait Database where Figure 13(a) provides a description of the trajectory indicated by the arrow and Figure 13(b) Shows sequential 3D models of someone walking straight. Database B (Curve): This database contains picture sequences of humans walking along curved paths; it was named after the shape of the trajectories (a). The radius r may be either 1.5 metres or 3.0 metres in length. This database is a representation of various 3D reconstructed models taking the curve route into consideration. Figure 14 depicts an example of the studio generated from the KY4D Gait Database B. Figure 14(a) and Figure 14(b) depicts the example of the studio generated from the KY4D gait database.

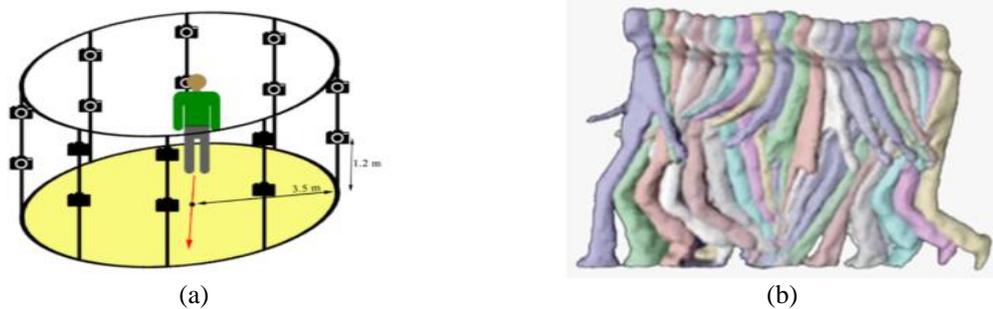


Figure 13. An example of the studio generated from the KY4D Gait Database where: (a) is an illustration of the trajectory, which is shown as a red arrow in the figure and (b) illustrates numerous 3D reconstructed models

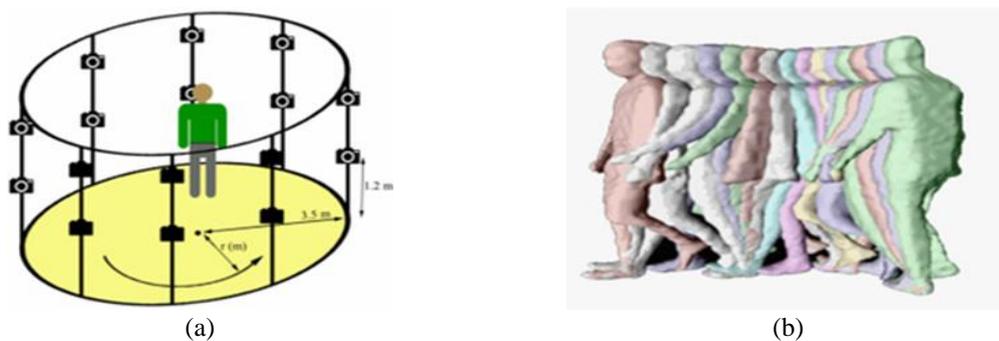


Figure 14. An example of the studio generated from the (a) KY4D Gait Database B and (b) depicts the example of the studio generated from the KY4D gait database

Kentucky infrared (IR) shadow gait database: This database is made up of time-series shadow photographs of 54 different people [92]. Figure 15 depicts the experimental setting and actual scene of the data generation where Figure 15(a) provides a description of the variation of radius and Figure 15(b) illustrates sequential 3D models showing the trajectory and path that a person follows as they walk along it. Everyone walks in a straight line, as shown by the arrow showing the direction of walking in Figure 15(a). A camera and two infrared lamps are shown in Figure 15(b), which illustrates the data collection process for the shadow database. The infrared lights were hung at an angle, and the camera was hung from the ceiling in a direction that was perpendicular to the floor. Figure 16 shows an example of the results that may be obtained from these captures.

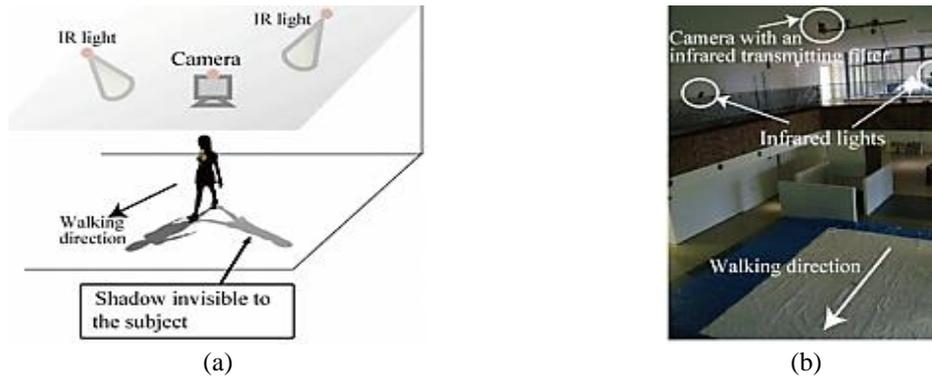


Figure 15. Experimental setting actual scene where (a) provides a description of the variation of radius and (b) illustrates sequential 3D models showing the trajectory and path that a person follows as they walk along it



Figure 16. An example movie from IR shadow images

3.10. WhuGAIT datasets

The findings presented in the research may be reproduced by using the whuGAIT Datasets which were made publicly accessible by Wuhan University in 2018 together with the source code and pre-trained models. In contrast to prior data collections, whuGAIT contains information obtained from 3D accelerometers and 3-axis gyroscopes acquired from 118 individuals, 20 of which were collected over the course of three days and 98 of which were taken in a single day [93]. According to the purpose of the analysis, the dataset has been segmented into the following six distinct subsets:

- Dataset #1: The first dataset is made up of 33,104 samples that were used for training and 3,740 that were used for testing. These samples were from 118 different people and were segmented in a two-step process.
- Dataset #2: A very similar dataset to that of Dataset #1, this dataset is also a two-stage segmentation dataset. 49,275 samples are included in the training portion of the study, and 4,936 samples are included in the testing portion, and it was derived from the data gathered over the course of three days by 20 individuals.
- Dataset #3: three contains a subset that has been segmented into temporal size windows, with each sample taking up 2.56 seconds. There are 26,283 examples in this set that are used for training, and there are 2,991 examples that are used for testing.
- Dataset #4: Similarly, to Dataset #3, this subset of data is divided into time frames of 2.56 seconds; however, it uses the data obtained from 20 different people over the course of three days. As part of this subset of records, 35,373 records are used in training purposes and 3,941 records are used in testing purposes.
- Dataset #5: For reasons of authentication, this particular subset is being used. It is made up of 74,142 examples contributed by 118 people, with the information retrieved from 98 of those persons being used for training and the remaining 20 being used for validation. There are several factors that negatively affect the authentication process, including the fact that both samples were obtained from either one or two distinct individuals. The cases include gyroscopic data and two-step acceleration measurements.
- Dataset #6: Unlike Dataset #5, this subset is also based on the same structure that was used in Dataset #5, but rather than using horizontal alignment, it makes use of vertical alignment.

4. DATASETS IN THEIR APPROPRIATE USE CONTEXTS

In this part, the datasets that are most often used for gait recognition and that have been utilised in the studies that were considered for this study were provided. In light of this, throughout the course of the sequence, we will also provide a general review of several topics, including use settings, acquisition environments, and

spectrum. The relevant information for each dataset is shown in Table 3. The variables were broken down into 12 primary categories, which included perspective, pace, objects, shoe, clothes, time, surface, silhouette, gait variation, walking on a treadmill, walking over ground, and foot pressure [94]. In addition, the construction habitats of the dataset. These environments are classified as either static inside or static outside, active indoors or active outdoors, or active both indoors and outdoors.

Table 3. Datasets in usage contexts

Context/Database	3.1	3.2	3.3	3.4	3.5	3.6	3.7	3.8	3.9	3.10
Viewpoint	x		x	x	x	x	x	x	x	x
Pace	x			x	x	x	x			x
Objects	x	x		x			x			
Shoe	x	x	x	x	x	x	x	x		
Clothing	x	x	x	x	x		x	x		
Time				x	x		x		x	x
Surface	x	x	x	x		x	x		x	x
Silhouette	x		x	x				x		
Gait Fluctuation				x						
Treadmill walking	x				x		x			
Overground walking	x	x	x	x		x		x	x	x
Foot pressure				x						

5. CONCLUSIONS AND FUTURE DIRECTIONS

In this article, a comprehensive review of the methodologies used in crucial stages as well as current developments in gait detection and identification approaches for humans affected by the COVID-19 pandemic. There are three key concerns about gait recognition, including representations of gait and features dimensionality reduction and gait classification are both terms that are often used. In addition to highlighting the latest advances in deep learning-based gait recognition technology for human operation, this survey provides a comprehensive review of the most important works on gait identification that were published in the previous several years. In addition to this, it offered an in-depth overview of the nine datasets that were used most for the work, along with the extraction procedure for each dataset and the limits of each dataset. When it comes to these datasets, it is easy to see that attempts are being made toward a detection that is unaffected by gender, angle in the scene, clothing and shoes worn, the carrying of bags and other goods, and several other factors that are often encountered. As part of this paper, we examine the primary issues that have been addressed by the scientific community and the efforts that have been made in order to develop approaches that can help ensure access to monitored resources during the COVID-19 pandemic via gait identification for humans conducted by the scientific community. In addition to this, it investigates the benefits of a gait-based method in comparison to regularly used biometric characteristics. For example, gait systems do not need the users to consent to being recognized, while biometric features must. In this context, the aim of the review is to emphasize the positive features of gait biometrics as compared to other biometric approaches, and make sure to emphasize the most recent methodologies as well as state-of-the-art structures that have been developed over the years. It is important to keep in mind some of the most important aspects of video-based models and databases that deserve your attention in order to maximize their effectiveness.

For the sake of training or validation, the datasets that are offered do not include more than one person in each video. Additionally, the surroundings in which the datasets were recorded are completely under the control of the researchers. In some of the sets, the angles are different, yet there are no changes to the items or even the colors in the backdrop. This is despite the fact that the angles are different. On some of the other films, the subjects are shown walking on electronic treadmills, which suggests an even higher level of control over both their motions and their surroundings.

As a consequence of this, the models that were provided and mentioned in the study are probably going to be prone to showing deficiency and misbehavior when they are put to the operation of the system in the real world, in public places, for example, it is possible to identify persons of interest who are strolling along within the general public space. It is anticipated that the number of works that are related to the following topics will increase in the coming years due to the emerging tendencies:

- a) Transformer networks: This method targets data streams such as video and performs very well in the context of gait detection [95].
- b) Recognizing a person's gender and age: It is one strategy that has been more popular over the last several years owing to the significance it has in a variety of applications [96].
- c) Monitoring of hazardous environments: The use of gait features is ideal for the purpose of monitoring hazardous environments because these types of environments typically have limited illumination. In order to obtain better biometric identification characteristics, it is necessary to obtain more descriptive data [97]. In the context of hazardous environments, gait features are perfectly suited for the purpose of monitoring.

- d) The presence of more than one person in the scene: the majority of the research on gait recognition has focused on identifying a single person in controlled environments when there are several people in the scene at the same time.

However, situations that really occur in real life often demand solutions that can withstand uncontrolled conditions in which more than one person is present. In addition, we found that there was a growing need for gait identification in humans who were wearing various garments or were carrying things or wearing face mask in COVID-19 pandemic. In addition to hybrid techniques that were comprised of gait biometric traits, such as the face as well as the ear.

ACKNOWLEDGEMENTS

This work was supported by Fisabilillah R&D Grant Scheme (FRDGS).

REFERENCES

- [1] M. N. A. Al-Hamadani, "COVID-19 prediction and detection using deep learning techniques," *International Journal of Computer Information Systems and Industrial Management Applications*, pp. 168-181, 2021, doi: 10.13140/RG.2.2.28613.58085.
- [2] M. M. Islam, F. Karray, R. Alhaji, and J. Zeng, "A review on deep learning techniques for the diagnosis of novel Coronavirus (COVID-19)," *IEEE Access*, pp. 30551–30572 2021, doi: 10.1109/ACCESS.2021.3058537.
- [3] U. Saeed, S. Y. Shah, J. Ahmad, M. A. Imran, Q. H. Abbasi, and S. A. Shah, "Machine learning empowered COVID-19 patient monitoring using non-contact sensing: An extensive review," *J. Pharm. Anal.*, vol. 12, no. 2, pp. 193–204, 2022, doi: 10.1016/j.jpha.2021.12.006.
- [4] S. Dargan and M. Kumar, "A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities," *Expert Syst. Applic.*, vol. 143, 2020, doi: 10.1016/j.eswa.2019.113114.
- [5] M. J. Nordin and A. Saadoon, "A survey of gait recognition based on skeleton model for human identification," *Research J. Applied Sci. Eng. Technol.*, vol. 12, no. 7, pp. 756–763, April 2016, doi: 10.19026/rjaset.12.2751.
- [6] P. Connor and R. Arun, "Biometric recognition by gait: a survey of modalities and features," *Comput. Vis. and Imag. Unders.*, vol. 167, pp. 1–27, 2018, doi: 10.1016/j.cviu.2018.01.007.
- [7] E. J. Harris, I. H. Khoo, and E. Demircan, "A survey of human gait-based artificial intelligence applications," *Frontiers in Robotics and AI*, vol. 8, pp. 1–28, 2022, doi: 10.3389/frobt.2021.749274.
- [8] P. H. Pisani, *et al.*, "Adaptive biometric systems: review and perspectives," *ACM Comput. Surv.*, vol. 52, no. 5, 1–38, 2019, doi: 10.1145/3344255.
- [9] S. Khan, S. Parkinson, L. Grant, N. Liu, and S. Mcguire, "Biometric systems utilising health data from wearable devices: Applications and future challenges in computer security," *ACM Comput. Surv.*, vol. 53, no. 4, pp. 1–29, 2020, doi: 10.1145/3400030.
- [10] A. Sundararajan, A. I. Sarwat, and A. Pons, "A survey on modality characteristics, performance evaluation metrics, and security for traditional and wearable biometric systems," *ACM Comput. Surv.*, vol. 52, no. 2, pp. 1–36, 2019, doi: 10.1145/3309550.
- [11] Q. Gui, V. Maria, R. Blondet, S. Laszlo, and Z. Jin, "A survey on brain biometrics," *ACM Comput. Surv.*, vol. 51, no. 6, pp. 1–38, 2019, doi: 10.1145/3230632.
- [12] C. Wan, L. Wang, and V. V. Phoha, "A survey on gait recognition," *ACM Comput. Surv.*, vol. 51, no. 5, pp. 1–35. 2018, doi: 10.1145/3230633.
- [13] Z. Zhang, L. Tran, F. Liu, and X. Liu, "On learning disentangled representations for gait recognition." *IEEE Trans. Pattern Anal. Mach. Intell.* vol. 44, no. 1, pp. 345-360, 2022, doi: 10.1109/TPAMI.2020.2998790.
- [14] J. P. Singh, S. Jain, S. Arora, and U. P. Singh, "Vision-based gait recognition: A survey," *IEEE Access.*, vol. 6, pp. 70497–70527, 2018, doi: 10.1109/ACCESS.2018.2879896.
- [15] I. Bouchrika, "A survey of using biometrics for smart visual surveillance: Gait recognition," in *Surveillance in Action*, Springer, pp. 3-23, 2018, doi: 10.1007/978-3-319-68533-5_1.
- [16] A. Nambiar, A. Bernardino, and J. C. Nascimento, "Gait-based person re-identification: a survey," *ACM Comput. Surv.*, vol. 52, no. 2, pp.1–34, 2019, doi: 10.1145/3243043.
- [17] M. D. Marsico and A. Mecca, "A survey on gait recognition via wearable sensors," *ACM Comput. Surv.*, vol. 52, no. 4, pp. 1–39, 2019, doi: 10.1145/3340293.
- [18] M. S. Seyfioğlu, S. Z. Gürbüz, A. M. Özbayoğlu, and M. Yüksel, "Deep learning of micro-doppler features for aided and unaided gait recognition," in *Proc. IEEE Radar Conf.*, 2017, pp. 1125–1130, doi: 10.1109/RADAR.2017.7944373.
- [19] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, pp. 2278–2324, 1998, doi: 10.1109/5.726791.
- [20] S. Hochreiter and J. Schmidhuber, "Long short-term memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [21] V. C. Lobo, L. A. Passos, and J. P. Papa, "Evolving long short-term memory networks," in *Proc. Int Conf. Comput. Scien. (ICCS'20)*. IEEE, 2020.
- [22] P. Vincent, H. Larochelle, I. Lajoie, Y. Bengio, P.-A. Manzagol, and L. Bottou, "Stacked denoising autoencoders: learning useful representations in a deep network with a local denoising criterion," *J. Mach. Learn. Res.*, vol. 11, no. 12, 2010.
- [23] S. Sabour, N. Frosst, and G. E. Hinton, "Dynamic routing between capsules," in *Proc. Int. Conf. Adv. Neu. Info. Proce. Syst.*, pp. 3856–3866, 2017.
- [24] G. E. Hinton, S. Osindero, and Y.W. The, "A fast learning algorithm for deep belief nets," *Neural Comput.*, vol. 18, no. 7, pp. 1527–1554, 2006, doi: 10.1162/neco.2006.18.7.1527.
- [25] I. J. Goodfellow, *et al.*, "Generative adversarial networks," *arXiv preprint arXiv:1406.2661*, 2014.
- [26] F. Battistone and A. Petrosino, "TGLSTM: A time based graph deep learning approach to gait recognition," *Pattern Recognit. Letters*, vol. 126, pp. 132–138, 2019, doi: 10.1016/j.patrec.2018.05.004.
- [27] V. Kukreja, D. Kumar, and A. Kaur, "Deep learning in human gait recognition: an overview," in *proc. Int. Conf. Adv. Comput. and Innov. Tech. in Eng., ICACITE 2021*, 2021, pp. 9–13, doi: 10.1109/ICACITE51222.2021.940461.
- [28] M. Alotaibi and A. Mahmood, "Improved gait recognition based on specialized deep convolutional neural networks," *Comput. Vis. Image Understanding*, vol. 164, pp. 103–110, 2017, doi: 10.1016/j.cviu.2017.10.004.

- [29] B. Rahi, "View-invariant gait person re-identification with spatial and temporal attention," Ph.D dissertation, Brunel Univ., London, 2021.
- [30] A. S. Alharthi, S. U. Yunas, and K. B. Ozanyan, "Deep learning for monitoring of human gait: a review," *IEEE Sensors J.*, vol. 19, no. 21, pp. 9575-9591, 2019, doi: 10.1109/JSEN.2019.2928777.
- [31] S. Sayeed, P. P. Min, and A. Bari, "Deep learning based gait recognition using convolutional neural network in the COVID-19 pandemic," *Emerging Sci. J.*, vol. 6, no. 5, pp. 1086-1099, 2022, doi: 10.28991/ESJ-2022-06-05-012.
- [32] J. Wang, M. She, S. Nahavandi, and A. Kouzani, "A review of vision-based gait recognition methods for human identification," in *proc. Digit. Imag. Comput.: Tech. Appl.*, pp. 320-327, 2010, doi: 10.1109/DICTA.2010.62.
- [33] W. Liu, C. Zhang, H. Ma, and S. Li, "Learning efficient spatialtemporal gait features with deep learning for human identification," *Neuroinformatics*, vol. 16, no. 3-4, pp. 457-471, February 2018, doi: 10.1007/s12021-018-9362-4.
- [34] F. G. D. Santos, et al., "Gait recognition based on deep learning: a survey," *ACM Comput. Surv.*, vol. 55, no. 2, 2022, doi: 10.1145/3490235.
- [35] F. Deligianni, Y. Guo, and G. Yang, "From emotions to mood disorders: A survey on gait analysis methodology," *IEEE J. Biomed. Health Inform.*, vol. 23, no. 6, pp. 2302-2316, 2019, doi: 10.1109/JBHI.2019.2938111.
- [36] 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, IEEE*, 1-8, 2016, doi: 10.1109/ICB.2016.7550060.
- [37] Z. Wu, Y. Huang, L. Wang, X. Wang, and T. Tan, "A comprehensive study on crossview gait based human identification with deep CNNs," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 2, pp. 209-226, 2017, doi: 10.1109/TPAMI.2016.2545669.
- [38] C. Li, X. Min, S. Sun, W. Lin, and Z. Tang, "DeepGait: A learning deep convolutional representation for viewinvariant gait recognition using joint Bayesian," *Appl. Sci.* vol. 7, no. 3, 2017, doi: 10.3390/app7030210.
- [39] A. Sokolova and A. Konushin. "Gait recognition based on convolutional neural networks," *Int. Arch. Photogramm. Rem. Sens. Spatial Inf. Sci.*, vol. 42, 2017, doi: 10.5194/isprs-archives-XLII-2-W4-207-2017.
- [40] 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. Circ. Syst. Vid. Technol.*, vol. 29, no. 9, pp. 2708-2719, 2017, doi: 10.1109/TCSVT.2017.2760835.
- [41] A. Sokolova and A. Konushin. "Pose-based deep gait recognition," *CoRR abs/1710.06512*, 2017.
- [42] S. Yu, H. Chen, Q. Wang, L. Shen, and Y. Huang, "Invariant feature extraction for gait recognition using only one uniform model," *Neurocomputing*, vol. 239, pp. 81-93, 2017, doi: 10.1016/j.neucom.2017.02.006.
- [43] Q. Zou, Y. Wang, Y. Zhao, Q. Wang, C. Shen, and Q. Li, "Deep learning based gait recognition using smartphones in the wild," *arXiv preprint arXiv:1811.00338*, 2018.
- [44] Y. He, J. Zhang, H. Shan, and L. Wang, "Multi-task GANs for view-specific feature learning in gait recognition," *IEEE Trans. Info. Foren. Secu.*, vol. 14, no. 1, pp. 102-113, 2018, doi: 10.1109/TIFS.2018.2844819.
- [45] C. Fernandes, et al., "Artificial neural networks classification of patients with Parkinsonism based on gait," in *Proc. IEEE Int. Conf. Bioinfo. Biomed.*, IEEE, 2018, pp. 2024-2030, doi: 10.1109/BIBM.2018.8621466.
- [46] Z. Xu, W. Lu, Q. Zhang, Y. Yeung, and X. Chen. "Gait recognition based on capsule network," *J. Vis. Commun. Image Represent.*, vol. 59, pp. 159-167, 2019, doi: 10.1016/j.jvcir.2019.01.023.
- [47] K. Wang, L. Liu, Y. Lee, X. Ding, and J. Lin, "Nonstandard periodic gait energy image for gait recognition and data augmentation," in *proc. Chinese Conf. Pattern Recog. Comput. Vis.*, Springer, pp. 197-208, 2019, doi: 10.1007/978-3-030-31723-2_17.
- [48] Z. Zhang, et al., "Gait recognition via disentangled representation learning," *arXiv preprint arXiv:1904.04925*, 2019, doi: 10.1109/CVPR.2019.00484.
- [49] M. Babae, L. Li, and G. Rigoll, "Person identification from partial gait cycle using fully convolutional neural networks," *Neurocomputing*, vol. 338, pp. 116-125, 2019, doi: 10.1016/j.neucom.2019.01.091.
- [50] S. Potluri, S. Ravuri, C. Diedrich, and L. Schega, "learning based gait abnormality detection using wearable sensor system," in *proc. 41st Annual Int. Conf. IEEE Eng. Medic. Bio. Society. IEEE*, pp. 3613-3619, 2019, doi: 10.1109/EMBC.2019.8856454.
- [51] M. Jia, H. Yang, D. Huang, and Y. Wang, "Attacking gait recognition systems via silhouette guided GANs," in *Proc. 27th ACM Int. Conf. Multimedia*, 2019, pp. 638-646, doi: 10.1145/3343031.3351018.
- [52] X. Wang and W. Q. Yan, "Human gait recognition based on frame-by-frame gait energy images and convolutional long short-term memory," *Int. J. Neural Syst.*, vol. 30, no. 1, 2020, doi: 10.1142/S0129065719500278.
- [53] C. Xu, Y. Makihara, X. Li, Y. Yagi, and J. Lu, "Cross-view gait recognition using pairwise spatial transformer networks," *IEEE Trans. Circ. Syst. Vid. Technol.*, vol. 31, no. 1, pp. 260-274, 2020, doi: 10.1109/TCSVT.2020.2975671.
- [54] B. Z. Hu, Y. Guan, Y. Gao, Y. Long, N. Lane, and T. Ploetz, "Robust cross-view gait recognition with evidence: A discriminant gait GAN (DiGGAN) approach," *arXiv preprint arXiv:1811.10493*, 2018.
- [55] S. M. Alireza, S. Ghorbani, N. F. Troje, and A. Etemad, "Gait recognition using multiscale partial representation transformation with capsules," in *Proc. Int. Conf. Pattern Recog. IEEE*, 2021, pp. 8045-8052.
- [56] D. Xiong, D. Zhang, X. Zhao, and Y. Zhao, "Continuous human gait tracking using sEMG signals," in *Proc. Annual Int. Conf. IEEE Eng. Medicine & Biology Society (EMBC '20). IEEE*, 2020, pp. 3094-3097, doi: 10.1109/EMBC44109.2020.9175238.
- [57] L. Tran, T. Hoang, T. Nguyen, H. Kim, and D. Choi, "Multi-model long short-term memory network for gait recognition using window-based data segment," *IEEE Access*, pp. 23826-23839, 2021, doi: 10.1109/ACCESS.2021.3056880.
- [58] A. Zhao, J. Dong, J. Li, L. Qi, and H. Zhou, "Associated spatio-temporal capsule network for gait recognition," *arXiv preprint arXiv:2101.02458*, 2021, doi: 10.1109/TMM.2021.3060280.
- [59] N. Subramanian, O. Elharrouss, S. Al-Maadeed, and M. Chowdhury, "A review of deep learning-based detection methods for COVID-19," *Comput. in Biology and Medicine*, vol. 143, 2022, doi: 10.1016/j.compbiomed.2022.105233.
- [60] J. Luo and T. Tjahjadi, "Multi-set canonical correlation analysis for 3D abnormal gait behaviour recognition based on virtual sample generation," *IEEE Access*, vol. 8, pp. 32485-32501, 2020, doi: 10.1109/ACCESS.2020.2973898.
- [61] C. Shen, S. Yu, J. Wang, G. Q. Huang, and L. Wang, "A comprehensive survey on deep gait recognition: algorithms, datasets and challenges," pp. 1-20, 2022, <http://arxiv.org/abs/2206.13732>.
- [62] R. K. Mishra, G. Y. S. Reddy, and H. Pathak, "The understanding of deep learning: a comprehensive review," *Math. Problems in Eng.*, vol. 2021, pp. 1-15, 2021, doi: 10.1155/2021/5548884.
- [63] J. Wang, Y. Chen, S. Hao, X. Peng, and L. Hu, "Deep learning for sensor-based activity recognition: A survey," *Pattern Recog. Letters*, vol. 119, pp. 3-11, March 2019, doi: 10.1016/j.patrec.2018.02.010.
- [64] R. Gross and J. Shi, "The CMU motion of body (MoBo) database," *Technical Report CMU-RI-TR-01-18*. Carnegie Mellon Univ., 2021.
- [65] 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, 2014, doi: 10.1016/j.jvcir.2013.02.006.
- [66] M. A. Khan et al., "Human gait analysis for osteoarthritis prediction: a framework of deep learning and kernel extreme learning machine," *Complex Intell. Syst.*, 2021, doi: 10.1007/s40747-020-00244-2.

- [67] Human Identification at a Distance, Retrived November 12, 2022, from <http://www.umiacs.umd.edu/labs/pirl/hid/data.html>
- [68] S. Sivapalan, "Human identification from video using advanced gait recognition techniques," PhD Thesis, Queensland University of Technology, Image and Video Research Laboratory, Science and Engineering Faculty, September 2014.
- [69] S. Zheng, J. Zhang, K. Huang, R. He, and T. Tan, "Robust view transformation model for gait recognition," in *Proc. IEEE Int. Conf. on Image Processing. IEEE*, 2011, 2073–2076, doi: 10.1109/ICIP.2011.6115889.
- [70] Gait Dataset, Retrived November 12, 2022, from http://www.cbsr.ia.ac.cn/users/szheng/?page_id=71.
- [71] S. Yu, D. Tan, and T. Tan, "A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition," *Pattern Recog.*, vol. 4, pp. 441–444, 2006, doi: 10.1109/ICPR.2006.67.
- [72] CASIA Gait Database, Retrived November 12, 2022, from <http://www.cbsr.ia.ac.cn/english/Gait%20Databases.asp>.
- [73] S. Zheng, K. Huang, T. Tan, and D. Tao, "A cascade fusion scheme for gait and cumulative foot pressure image recognition," *Pattern Recog.*, vol. 45, no. 10, pp. 3603–3610, 2012, doi: 10.1016/j.patcog.2012.03.008.
- [74] H. Chao, K. Wang, Y. He, J. Zhang, and J. Feng, "Gaitset: Crossview gait recognition through utilizing gait as a deep set," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 44, no. 7, pp. 3467–3478, 2022.
- [75] A. Tsuji, Y. Makihara, and Y. Yagi, "Silhouette transformation based on walking speed for gait identification," in *Proc. IEEE Comput. Society Conf. Comput. Vis. and Pattern Recog. IEEE*, pp. 717–722, 2010, doi: 10.1109/CVPR.2010.5540144.
- [76] M. A. Hossain, Y. Makihara, J. Wang, and Y. Yagi, "Clothing-invariant gait identification using part-based clothing categorization and adaptive weight control," *Pattern Recog.*, vol. 43, no. 6, pp. 2281–2291, 2010, doi: 10.1016/j.patcog.2009.12.020.
- [77] Y. Makihara, H. Mannami, and Y. Yagi, "Gait analysis of gender and age using a large-scale multi-view gait database," in *Proc. Asian Conf. on Comput. Vis., Springer*, pp. 440–451, 2010, doi: 10.1007/978-3-642-19309-5_34.
- [78] A. Mori, Y. Makihara, and Y. Yagi, "Gait recognition using period-based phase synchronization for low frame-rate videos," in *Proc. Int. Conf. Pattern Recog., IEEE*, pp. 2194–2197, 2010, doi: 10.1109/ICPR.2010.537.
- [79] H. Iwama, M. Okumura, Y. Makihara, and Y. Yagi, "The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition," *IEEE Trans. Inf. Forens. Secur.*, vol. 7, no. 5, 1511–1521, 2021, doi: 10.1109/TIFS.2012.2204253.
- [80] A. Mansur, Y. Makihara, R. Aqmar, and Y. Yagi, "Gait recognition under speed transition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recog.*, pp. 2521–2528, 2014, doi: 10.1109/CVPR.2014.323.
- [81] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, "Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition," *IPSSJ Trans. Comput. Vis. Applic.*, vol. 10, no. 1, 2018, doi: 10.1186/s41074-018-0039-6.
- [82] M. Z. Uddin *et al.*, "The OU-ISIR large population gait database with real-life carried object and its performance evaluation," *IPSSJ Trans. Comput. Vis. Applic.*, vol. 10, no. 1, pp. 1–11, 2018, doi: 10.1186/s41074-018-0041-z.
- [83] C. Xu, Y. Makihara, G. Ogi, X. Li, Y. Yagi, and J. Lu, "The OU-ISIR gait database comprising the large population dataset with age and performance evaluation of age estimation," *IPSSJ Trans. Comput. Vis. Applic.*, vol. 9, no. 24, pp. 1–14, 2017, doi: 10.1186/s41074-017-0035-2.
- [84] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "The largest inertial sensor-based gait database and performance evaluation of gait-based personal authentication," *Pattern Recog.*, vol. 47, no. 1, pp. 228–237, 2014, doi: 10.1016/j.patcog.2013.06.028.
- [85] T. T. Ngo, Y. Makihara, H. Nagahara, Y. Mukaigawa, and Y. Yagi, "Similar gait action recognition using an inertial sensor," *Pattern Recog.*, vol. 48, no. 4, pp. 1289–1301, 2015, doi: 10.1016/j.patcog.2014.10.012.
- [86] S. Sarkar, P. J. Phillips, Z. Liu, I. R. Vega, P. Grother, and K. W. Bowyer, "The humanID gait challenge problem: Data sets, performance, and analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 27, no. 2, 162–177, 2005, doi: 10.1109/TPAMI.2005.39.
- [87] Southampton Human ID at a Distance database, Retrived November 12, 2022, from <https://web-archive.southampton.ac.uk/www.gait.ecs.soton.ac.uk/database/index.html>.
- [88] Large population database, Retrived November 12, 2022, from https://web-archive.southampton.ac.uk/www.gait.ecs.soton.ac.uk/database/large_db.php3.html.
- [89] S. Matovski, M. S. Nixon, S. Mahmoodi, and J. N. Carter, "The effect of time on the performance of gait biometrics," in *Proc. IEEE Int. Conf. on Biometrics: Theory, Appl. and Syst. (BTAS'10). IEEE*, 2010, pp. 1–6, doi: 10.1109/BTAS.2010.5634547.
- [90] D. L. Fernández, F. J. M. Cuevas, Á. C. Poyato, M. J. M. Jiménez, and R. M. Salinas, "The AVA multi-view dataset for gait recognition. in activity monitoring by multiple distributed sensing," *Pier Luigi Mazzeo, Paolo Spagnolo, and Thomas B. Moeslund (Eds.). Springer International Publishing*, pp. 26–39, 2014, doi: 10.1007/978-3-319-13323-2-3.
- [91] Y. Iwashita, K. Ogawara, and R. Kurazume, "Identification of people walking along curved trajectories," *Pattern Recog. Lett.* 2014, doi: 10.1016/j.patrec.2014.04.004.
- [92] Y. Iwashita, R. Kurazume, and A. Stoica, "Gait identification using invisible shadows: Robustness to appearance changes," in *Proc. Int. Conf. on Emerging Sec. Techno. (EST'14)*, UK, 2014, doi: 10.1109/EST.2014.18.
- [93] Q. Zou, Y. Wang, Y. Zhao, Q. Wang, C. Shen, and Q. Li, "Deep learning based gait recognition using smartphones in the wild," *CoRR abs/1811.00338*, 2018.
- [94] I. Rida, N. Almaadeed, and S. Almaadeed, "Robust gait recognition: a comprehensive survey," *IET Biometrics*, vol. 8, no. 1, pp. 14–28, January 2019, doi: 10.1049/iet-bmt.2018.5063.
- [95] A. Vaswani *et al.*, "Attention is all you need," *arXiv preprint arXiv:1706.03762*, 2017.
- [96] Y. Sun, F. P. W. Lo, and B. Lo, "A deep learning approach on gender and age recognition using a single inertial sensor," in *Proc. IEEE 16th Int. Conf. Wearable and Implantable Body Sensor Networks (BSN'19). IEEE*, 2019, pp. 1–4, doi: 10.1109/BSN.2019.8771075.
- [97] M. C. Santana, L. A. Passos, T. P. Moreira, D. Colombo, V. H. C. Albuquerque, and J. P. Papa, "A novel siamese-based approach for scene change detection with applications to obstructed routes in hazardous environments," *IEEE Intell. Syst.*, vol. 35, no. 1, pp. 44–53, 2019, doi: 10.1109/MIS.2019.2949984.

BIOGRAPHIES OF AUTHORS



Md Shohel Sayeed    has been a member of Multimedia University since 2001 and now he serves as an Associate Professor and the Chairperson of Centre for Intelligent Cloud Computing (CICC) of the Faculty of Information Science and Technology. He received his Doctor of Philosophy from Multimedia University in Engineering, specializing in hand signature verification, and holds Master of Information Technology degree from the University Kebangsaan Malaysia, specializing in industrial computing. He received his Bachelor of Science degree in Agricultural science from Bangladesh Agricultural University. Dr. Shohel has more than 24 years meritorious working experience and he holds a challenging career which combines research, versatile administration, and excellent teaching. His core research interest is in the area of Biometrics, big data, cloud computing, artificial intelligence, information security, image and signal processing, pattern recognition and classification. He has published over 80 research papers in international peer-reviewed journals and international conference proceedings as a result of his research work. His research works have been published by high ranked peer-reviewed journals such as IEEE transactions on pattern analysis and machine intelligence (TPAMI), international journal of pattern recognition and artificial intelligence (IJPRI), expert systems with applications, discrete dynamics in nature and society (DDNS) as well as several peer-reviewed international journals. He can be contacted at email: shohel.sayeed@mmu.edu.my.



Ibrahim Bin Yusof    is a Lecturer at Multimedia University, Melaka, Malaysia. He also a trainer and consultant in ICT, Linux/open source software, computer security and computer forensics for the past 20 years. He has certified hacking forensic instigator (CHFI), linux professional institute certification (LPIC-1), SUSE 11 Tech Specialist, novell certified linux professional (NCLP), novell certified linux administrator (NCLA), data center technical specialist (DCTS), data center advanced technical specialist (DCATS), linux technical specialist (LTS). Besides his professional certifications, Ibrahim holds Master of Science in Information Technology, Advanced Diploma in Business Studies (Marketing), Diploma in Plantation Management and currently pursuing his Ph.D in ICT. He is also a professional member for The Association for Computing Machinery (ACM) and ordinary member for Internet Society (ISOC). He can be contacted at email: ibrahim.yusof@mmu.edu.my.



Mohd Fikri Azli bin Abdullah    is a Lecturer at Multimedia University, Melaka, Malaysia. He has completed his first degree in Bachelor of Information Technology in Information Science Engineering from Multimedia University (MMU). Then he further his studies for Master in Software System Engineering from Melbourne University, Australia. His interests are in several areas of IT like pattern recognition, image processing, vehicle system and technology, cloud computing, data analytics and internet of things. He can be contacted at email: mfikriazli.abdullah@mmu.edu.my.



Md Ahsanul Bari    is currently pursuing a Master of Science (Information Technology) degree in the Faculty of Information Science and Technology at Multimedia University, Malaysia. He received the Bachelor of Information Science and Technology (Hons) degree, majoring in data communication and networking. His research interests are in the internet of things (IoT), networking, cloud computing, and artificial intelligence. He can be contracted at email: ahsanulbari99@gmail.com.



Pa Pa Min    is currently pursuing the Doctor of Philosophy (Information Technology) in Faculty of Information Science and Technology at Multimedia University. She received the BIT(Hons) degree majoring in Information Technology Management and Master of Science (Information Technology). Her research interests are deep learning, pattern recognition and artificial intelligent. She can be contracted at email: mspapaminai@gmail.com.