A survey on gait recognition against occlusion: taxonomy, dataset and methodology

Tianhao Li[1,2], Weizhi Ma[1], Yujia Zheng[1,3], Xinchao Fan[1], Guangcan Yang[1], Lijun Wang[4] and Zhengping Li[1]

[1] School of Information Science and Technology, North China University of Technology, Beijing, China
[2] Department of Medical Physics, Duke University, Durham, North Carolina, United States
[3] State Key Laboratory of Intelligent Game, Institute of Software Chinese Academy of Sciences, Beijing, China
[4] Advance Vision Institute, Hangzhou Institute of Technology, Xidian University, Hangzhou, China

ABSTRACT

Traditional biometric techniques often require direct subject participation, limiting application in various situations. In contrast, gait recognition allows for human identification *via* computer analysis of walking patterns without subject cooperation. However, occlusion remains a key challenge limiting real-world application. Recent surveys have evaluated advances in gait recognition, but only few have focused specifically on addressing occlusion conditions. In this article, we introduces a taxonomy that systematically classifies real-world occlusion, datasets, and methodologies in the field of occluded gait recognition. By employing this proposed taxonomy as a guide, we conducted an extensive survey encompassing datasets featuring occlusion and explored various methods employed to conquer challenges in occluded gait recognition. Additionally, we provide a list of future research directions, which can serve as a stepping stone for researchers dedicated to advancing the application of gait recognition in real-world scenarios.

Subjects Computational Biology, Human-Computer Interaction, Artificial Intelligence, Data Mining and Machine Learning, Databases
Keywords Gait recognition, Occlusion, Biometric recognition, Forensic, Authentication security

INTRODUCTION

Authentication security is becoming increasingly important. Traditional knowledge-based authentication methods, such as passwords, are vulnerable to attack (*Sivaranjani et al., 2023*; *Wang et al., 2020*; *Liang et al., 2020*). Biometric authentication, which uses unique biological characteristics to verify identity, offers a more secure alternative (*Ryu et al., 2021*; *Maharjan et al., 2021*). However, traditional biometric authentication methods, such as fingerprint and facial recognition, are also susceptible to exploitation and forgery (*Coelho et al., 2023*; *Bodepudi & Reddy, 2020*; *Su, 2021*). Therefore, there is a need for new biometric traits that are difficult to forge. Such traits would offer a more robust and secure form of authentication for individuals and organizations alike.
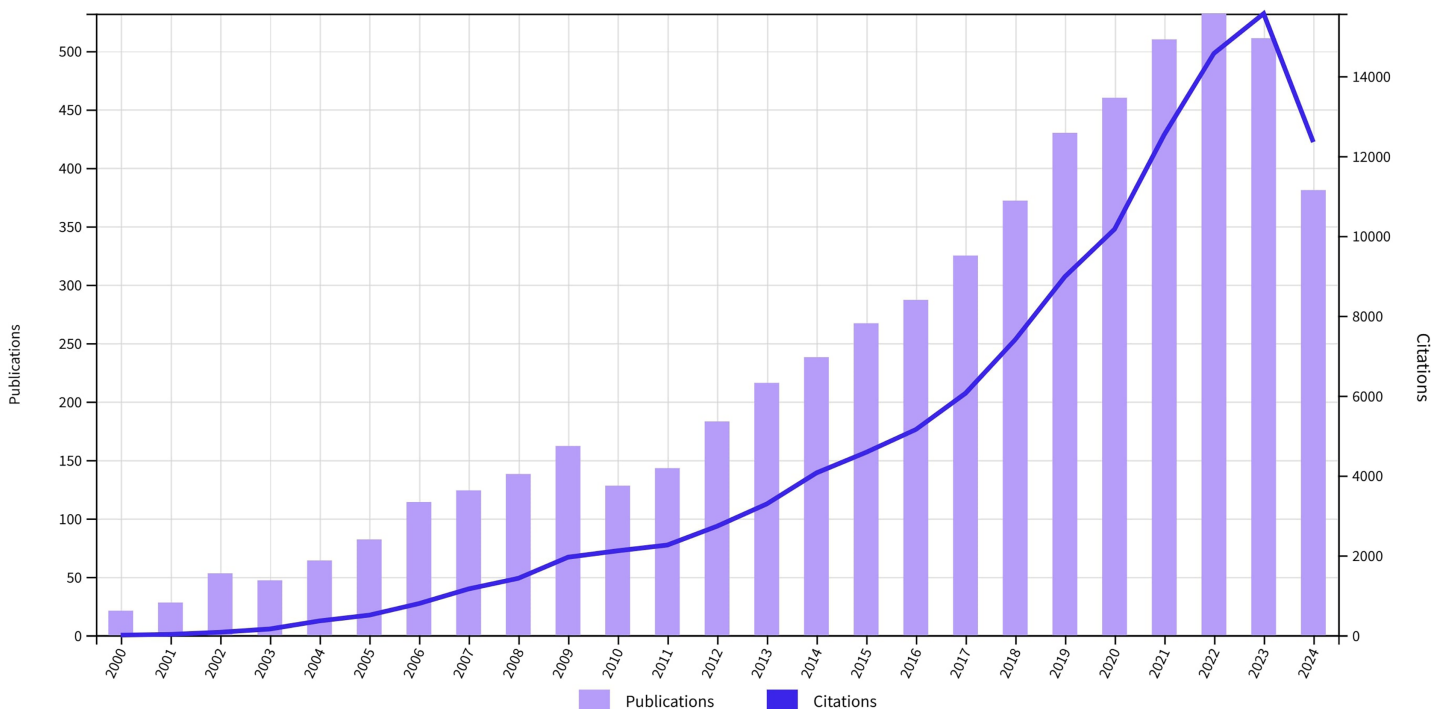
Figure 1 **Annual number of publications and citations for gait recognition research from 2000 to 2024, showing steady growth in publications and a significant increase in citations in recent years.** (Data source: Web of Science Citation Report, Keyword: gait recognition, Range: All databases except preprints from January 1, 2000, to October 26, 2024).   Full-size ⬙ DOI: 10.7717/peerj-cs.2602/fig-1

Gait, the manner in which an individual walks, is a unique identifier of individuals if all gait movements are considered (*Murray, Drought & Kory, 1964*; *Hug et al., 2019*; *Pataky et al., 2012*). In recent years, as Fig. 1 shown, gait recognition has increasingly gained prominence within the domain of biometric authentication. Survey articles (*Marsico & Mecca, 2019*; *Rida, Almaadeed & Almaadeed, 2019*; *Nambiar, Bernardino & Nascimento, 2019*; *Marsico & Mecca, 2019*; *Singh et al., 2018*; *Connor & Ross, 2018*; *Nordin & Saadoon, 2016*; *Singh et al., 2021*; *Sepas-Moghaddam & Etemad, 2022*; *Santos et al., 2022*; *Rani & Kumar, 2023*; *Khaliluzzaman et al., 2023*; *Güner Şahan, Şahin & Kaya Gülağız, 2024*; *Atrushi & Abdulazeez, 2024*; *Khan et al., 2024*; *Parashar et al., 2024*) have comprehensively examined and evaluated the latest advancements in the field of gait recognition. However, in practical applications, gait recognition remains an imperfect technique for industrial implementation. Carried objects (*Li et al., 2020a*; *Makihara et al., 2017*), walking speed (*Xu et al., 2017*; *Guan & Li, 2013*), observation view (*Wu et al., 2016*; *Makihara et al., 2006*), and occlusion (*Muramatsu, Makihara & Yagi, 2015*; *Uddin et al., 2019*; *Kumari & Bharti, 2023*) are challenging factor. In those factors, occlusion is the most common one, such as objects obstructing parts of the body during walking, can lead to reduced accuracy or even failure of gait recognition systems. In other words, certain gait movements remain concealed from the gait model due to occlusion, thereby greatly impacting the accuracy and robustness of gait recognition (*Kumari & Bharti, 2023*).

Figure 2 Example of real-world occlusion. The left image displays a case of static occlusion caused by bicycles parked between camera and subject. No motion is involved as the bicycles remain stationary. On the right is an example of dynamic occlusion from a pedestrian crowd. People walking both towards and away from the camera introduce temporally varying obstruction over time. These pictures demonstrate two prevalent types of occlusion encountered in everyday scenarios.

Full-size 🖼 DOI: 10.7717/peerj-cs.2602/fig-2

Motivation

Gait recognition has yet to be substantiated as a more dependable biometric modality compared to well-established techniques like fingerprinting, facial recognition, and iris recognition (*Bouchrika, 2018*), its practical application remains limited by performance degradation due to occlusion. Real-world conditions inevitably introduce occlusion, whether from clothing, carrying items, or partial view. Therefore, it is vital important to find the challenges of gait recognition in real-world settings. As discussed in introduction part, one of the foremost challenges encountered in real-world gait authentication pertains to occlusion (*Kumari & Bharti, 2023*). Imagine the scenarios below: (i) **Dynamic occlusion:** A significant number of individuals are present within the frame, their visibility may be compromised, thereby impeding the accurate detection of their walking patterns, as shown in Fig. 2. (ii) **Static occlusion:** When individuals walk behind objects like pillars, trees, or vehicles, their legs and lower body may be hidden from view, making it difficult to capture their distinctive gait features, as shown in Fig. 2. (iii) **Other occlusion:** When subject cut out or too close to lens, the occlusion increases. When the media quality cannot be guarantee cause by bad transmission or signal interference, salt and pepper noise (*Azzeh, Zahran & Alqadi, 2018*) or frame lost (*Lin et al., 2018*) may hinder the algorithm as well, become a potential occlusion.

The presence of occlusions are hindering successful implementation of gait recognition in real-world. Overcoming this obstacle is crucial to realizing the goal of deploying gait recognition technology in practical scenarios. Some researches have been conducted to solve this thorny problem. Unfortunately, to the best of our knowledge, there are only few review articles (*Kumari & Bharti, 2023*) focusing on gait recognition under occluded

conditions. This study aims to bridge the gap between academic research and real-world application by focusing on gait recognition under occluded conditions. It strives to provide insights, methodologies, and solutions to overcome the challenges posed by occlusions and advance gait recognition technology in practical scenarios.

Contribution

This article surveys recent advancements in gait recognition under occlusion conditions up to October 2024. The primary contributions are summarized as follows:

- We propose a structured taxonomy that systematically categorizes: (i) real-world occlusion types encountered in gait recognition, (ii) datasets that incorporate occlusions as sample variations, and (iii) methodologies designed to address the unique challenges posed by occluded gait sequences.
- We conduct a detailed review guided by this taxonomy, emphasizing datasets and state-of-the-art methods in occluded gait recognition. This comprehensive review offers critical insights for ongoing research and lays a foundation for developing novel algorithms in occlusion-resilient gait recognition.
- We outline potential directions for future research, grounded in current studies and established biometric frameworks, to guide researchers in addressing the complexities of occluded gait recognition in practical applications.

Target audience

This article will be of particular interest to researchers and developers working in the field of biometric identification, especially gait recognition. As occlusion presents a significant challenge to applying gait recognition in real-world scenarios, computer vision scientists seeking to advance occlusion-robust gait recognition algorithms will benefit from the proposed taxonomy and extensive literature review. Additionally, the suggested future research directions will aid investigators focused on improving gait recognition for practical applications. Beyond the research community, this survey offers value to government and private sector stakeholders interested in deploying biometric solutions for purposes such as surveillance and access control. By summarizing key datasets, methodologies, and gaps in current occluded gait recognition capabilities, this article provides guidance to support development of more robust systems capable of overcoming limitations imposed by occlusion. Ultimately, a wide audience encompassing technology researchers, developers, and end users stands to gain from progress toward accurate gait recognition under real-world conditions.

Organization

The structure of this article is as follows: In "Review Methodology", we detail the workflow and methodologies applied in this survey, including data sources, search strategies, and screening techniques used to identify and assess literature focused on gait recognition under occlusion conditions. "Deep Gait Recognition" provides a comprehensive overview of the gait recognition pipeline, covering essential processes such as data acquisition,
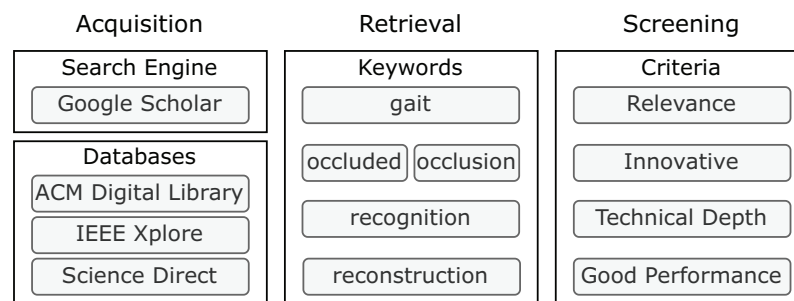
Figure 3 Process of review.   Full-size 🖼 DOI: 10.7717/peerj-cs.2602/fig-3

feature representation, dimension reduction, classification, and performance metrics, along with a discussion of potential application scenarios. This section establishes foundational knowledge for addressing challenges related to occlusions. In "Taxonomy", we present a structured taxonomy specifically developed to classify different approaches for tackling occluded gait recognition. "Review of Datasets" surveys datasets pertinent to occluded gait recognition, examining the resources available for algorithm development and testing. "Review of Literature" provides a comparative evaluation of state-of-the-art methods in occluded gait recognition, with experimental results on widely-used datasets. Finally, in "Future Directions", we synthesize our insights and outline prospective research directions for enhancing gait recognition systems under occlusion.

REVIEW METHODOLOGY

To capture high-quality, relevant research in this field and align our review methodology with established approaches in biometric recognition and processing (*Connor & Ross, 2018*; *Dargan & Kumar, 2020*; *Sepas-Moghaddam & Etemad, 2022*; *Ma et al., 2024b*, *2024a*; *Zheng et al., 2024d*), we conducted a rigorous search procedure to ensure a broad range and high caliber of research articles. The literature retrieval process consists of three main phases, as illustrated in Fig. 3: (i) To maximize the retrieval scope, we searched academic databases, including Google Scholar (https://scholar.google.com/); ACM Digital Library (https://dl.acm.org/); IEEE Xplore (https://ieeexplore.ieee.org/Xplore/home.jsp); and Science Direct (https://www.sciencedirect.com/). (ii) We used the keywords "gait occluded occlusion recognition reconstruction." (iii) To maintain relevance and research integrity, we manually screened each result by examining the title, abstract, methodology, experiments, and conclusions. Each article's relevance, innovation, technical depth, and performance were assessed to ensure that only impactful research was included in the survey. The selected articles then underwent a detailed review in line with the taxonomy we have proposed.

DEEP GAIT RECOGNITION

Benefiting from the advance of deep learning (*LeCun, Bengio & Hinton, 2015*; *Goodfellow, 2016*), deep learning-based models has dominated the field and achieved the state-of-the-art performance (*Alom et al., 2019*). Almost every recent approach uses the deep neural network as backbone in gait recognition task (*Sepas-Moghaddam & Etemad, 2022*). We

Figure 4 An example of silhouette. Full-size 🖼 DOI: 10.7717/peerj-cs.2602/fig-4

present general methods and popular models in this section. The pipeline of gait recognition typically consists of four main phases: data acquisition, feature representation, dimension reduction, and classification.

Data acquisition. This phase involves the collection of raw gait data, which can be obtained through various sensors such as cameras, accelerometers, floor sensors, or radar. For instance, camera-based systems capture video of a person walking, while accelerometer-based systems record the 3D accelerations of an individual's movements.

Feature representation. In this phase, the raw data collected is processed to extract meaningful features that represent the gait pattern. Model-based approaches may use kinematic features of human joints, such as trajectory-based features, while model-free approaches treat the silhouette, as shown in Fig. 4, as a whole and extract features for identification.

Dimension reduction. The extracted features are often high-dimensional, which can be challenging for subsequent processing. Dimension reduction techniques are applied to reduce the feature space without losing significant information. This step is crucial for improving the efficiency and performance of the gait recognition system.

Classification. The final phase involves the classification of the reduced-dimensional feature set to recognize or verify the identity of an individual. This is typically done using machine learning or deep learning models that have been trained on labeled datasets.

Beyond the four core processes above, we need metrics to evaluate the overall performance of gait recognition system. And we discuss some potential application scenarios in the last subsection.

Data acquisition: gait cycle capture

The gait cycle can be defined as the temporal interval between two consecutive instances of a repetitive characteristic associated with human locomotion by foot. Namely, the gait cycle demarcates the time from one particular phase or event of the walking motion to its subsequent repetition, such as from initial contact of the same foot to the next initial contact of that same foot (*Kharb et al., 2011*).

The human gait cycle can be delineated into seven distinct events that partition it into stance phase and swing phase, as shown in Fig. 5. Specifically, initial contact marking the
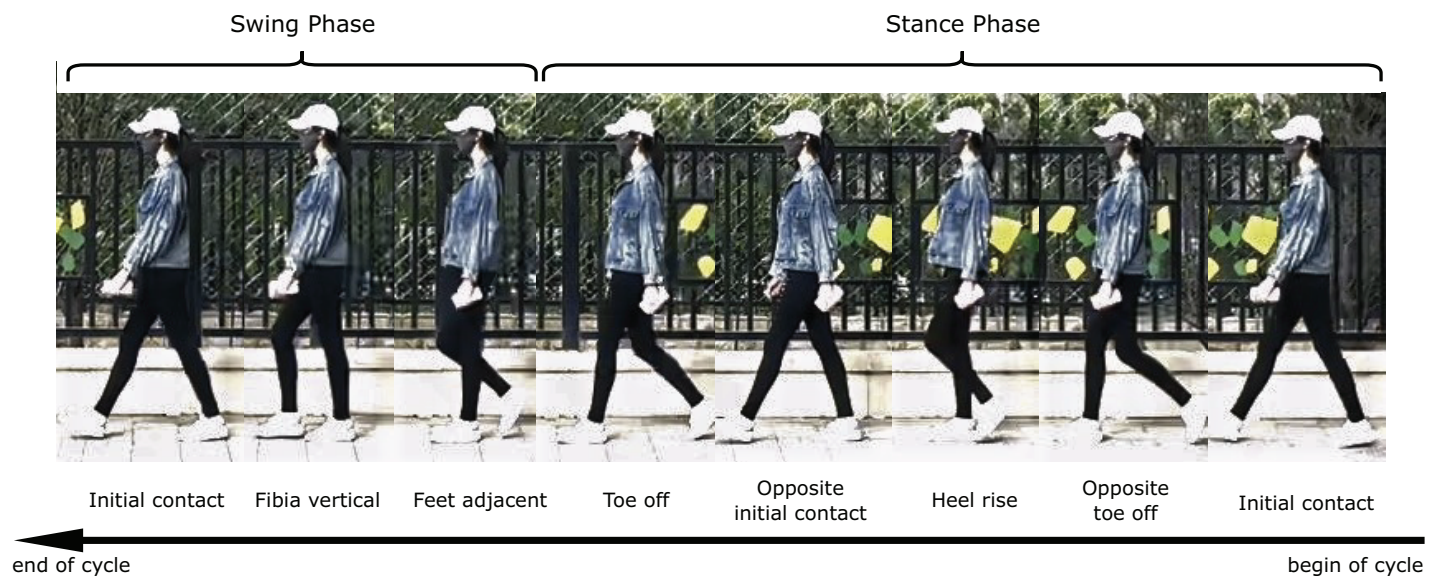
Swing Phase　　　　　　　　　　　　　　　　Stance Phase

Initial contact　Fibia vertical　Feet adjacent　Toe off　Opposite initial contact　Heel rise　Opposite toe off　Initial contact

end of cycle　　　　　　　　　　　　　　　　　　　　　begin of cycle

Figure 5 **The human gait cycle consists of stance and swing phases.** Stance begins with initial foot contact and ends with toe-off, involving opposite toe off, heel rise and opposite contact. Swing then occurs from toe-off to the next initial contact, comprising feet adjacent and tibia vertical positions. This sequence of seven events divides locomotion into balanced stance and swing durations with the foot either on or off the ground (*Kharb et al., 2011*). (The subject is one of the authors, and the photo was taken and used with her permission).

Full-size ☒ DOI: 10.7717/peerj-cs.2602/fig-5

beginning of stance phase followed by opposite toe off, heel rise, opposite initial contact and toe off which concludes stance and commences swing phase comprising of feet adjacent, tibia vertical. Stance, also termed the support or contact phase, encompasses four periods-loading response, mid-stance, terminal stance and pre-swing from initial contact to toe off. Meanwhile, swing spanning from toe off to the subsequent initial contact contains initial swing, mid-swing and terminal swing. Notably, a full gait cycle is quantified as the cycle time bifurcated into stance time spent on the ground and swing time airborne between footfalls. *Kharb et al. (2011)* Occlusion can impede accurate gait recognition when obstacles block view of the subject's motion. Occlusion obscure gait cycles and limit the information gained of later gait recognition process (*Kumari & Bharti, 2023*).

Gait cycle capture is a fundamental component of the gait recognition pipeline. This process involves the use of data acquisition devices to bridge the gap between the physical and digital domains, facilitating the translation of human locomotion into quantifiable data. The devices employed can be categorized into four primary types: camera, radar, accelerometer, and sensor (*Rani & Kumar, 2023*). Camera systems, for instance, are commonly employed in vision-based gait recognition and offer the advantage of non-intrusively capturing gait patterns from a distance. Radar technology provides an alternative means of capturing gait, offering advantages in terms of penetration through obstacles and the ability to measure velocity. Accelerometers, often used in wearable devices, are effective in capturing dynamic movements and have been integral in the development of model-based gait recognition techniques. Sensors, on the other hand, provide a complementary perspective by capturing pressure distribution and have been

Table 1 Overview of gait cycle capture devices: advantages and disadvantages across different sensing technologies (*Rani & Kumar, 2023*).

Name	Advantages	Disadvantages
Camera-based devices		
RGB camera	Non-intrusive	Poor depth perception, poor performance in low light
Microsoft Kinect	Non-intrusive, precise depth data	Sensitive to sunlight, infrared interference
Dynamic vision sensor	Responds to local changes	Expensive, limited availability
Event sense camera	Captures asynchronous events	Expensive, limited availability
Radar-based devices		
Radar micro-doppler	Efficient gait signature capture	Emits harmful microwaves
Radar	Efficient gait signature capture	Longer detection time
Millimetre Wave Radar Sensor	High spatial resolution, high velocity resolution	High penetration loss, poor diffraction
LiDAR (VLP16, HDL64)	Cost-effective, precise distance measurement	Emits harmful laser waves
Depth kinetic sensor	Accurate body dimension detection	Limited detection depth
Accelerometer-based devices		
Wearable accelerometer	Cost-effective, accurate	Requires contact with subject
Wearable gyroscope	Cost-effective, accurate	Requires contact with subject
Wearable body suit	High accuracy	Requires contact with subject
Inertial measurement unit (IMU)	Accurate human motion tracking	Expensive
Sensor-based devices		
Force platform pressure sensor	No risk of missing platform	Limited availability
Magnetometer	Low power consumption	Requires contact with subject
Gyroscope	High frequency response	Requires contact with subject
Accelerometer	High impedance, high sensitivity	Requires contact with subject
EEG headset	High signal quality, accurate reading	Requires contact with subject
Soft robotic sensor	High accuracy, increased flexibility	Requires contact with subject
Wearable pants	Accurate, reliable	Expensive, high power consumption
F-scan	Dynamic response	Expensive, high power consumption

useful in clinical gait analysis. In Table 1, we divided mainstream gait data acquisition devices into those four categories with their advantages and disadvantages.

Vision-based gait recognition has gained widespread attention as a practical biometric identification method due to its numerous advantages (*Singh et al., 2018*; *Khan, Farid & Grzegorzek, 2021*), as shown in Table 1. One key benefit is its non-intrusive nature, as it does not require physical contact with the subject (*Derbel, Vivet & Emile, 2015*). Unlike sensor-based approaches, such as accelerometers or gyroscopes, cameras can capture gait data from a distance, making them ideal for public and large-scale applications where attaching devices to individuals is impractical. This remote, real-time capability makes vision-based systems particularly well-suited for surveillance and security applications. The practicality of vision-based gait recognition is further enhanced by the widespread availability of camera technology. Cameras, particularly RGB cameras and depth sensors like Microsoft Kinect (*Zhang, 2012*), are already widely deployed in various environments such as public spaces and buildings. This existing infrastructure reduces the cost and

complexity of implementing gait recognition systems. In contrast, wearable sensor systems are limited in scalability, as they require equipping individuals with specific devices, whereas vision-based systems can monitor large populations simultaneously with minimal additional resources. Another significant advantage is the versatility of vision-based systems, which can perform multiple computer vision tasks beyond gait recognition, such as face identification and behavior analysis, using the same visual data. This multifunctionality adds value, especially in environments like law enforcement and public safety, where different forms of surveillance are needed. Moreover, the rich spatial and temporal data captured by vision-based systems, particularly those with depth-sensing capabilities, allows for detailed analysis of gait dynamics, improving the accuracy and reliability of these systems in real-world conditions. Recent advancements in deep learning and computer vision have further improved the performance of vision-based gait recognition, enabling automatic feature extraction and robust handling of variations in lighting, clothing, and walking surfaces. However, some limitations remain, such as sensitivity to environmental factors like lighting conditions and potential privacy concerns due to the detailed visual data collected. Despite these challenges, vision-based systems offer scalable, non-intrusive, and cost-effective solutions for biometric identification in various fields, from security to healthcare (*Parashar et al., 2024*).

Feature representation

To the best of our knowledge, only few recent researches achieved end-to-end gait recognition (*Song et al., 2019*; *Li et al., 2020b*; *Liang et al., 2022*; *Mehmood et al., 2020*; *Delgado-Escano et al., 2018*; *Li et al., 2021*; *Luo et al., 2019*). Most of methods follow step-by-step manner, have not achieved true end-to-end gait recognition from video alone. Direct extraction of gait features from a walking video remains a challenge task without intermediate representation. An effective way is needed to encode both bodily structure and motion sequence from visual walking data. An intermediate body model serves as a routing to encapsulate essential body and movement information, facilitating subsequent analysis and recognition.

Both silhouette (*Collins, Gross & Shi, 2002*; *Lam, Cheung & Liu, 2011*; *Wang et al., 2003*) and skeleton (*Rashmi & Guddeti, 2022*; *Teepe et al., 2022*; *Choi et al., 2019*) can be use as representation. Actually, most existing gait recognition based on silhouette or skeleton. Silhouette-based extraction offers computational efficiency but remains somewhat vulnerable to variabilities in environmental and subject-related factors such as occlusion. By contrast, skeleton-based modelling demonstrates greater robustness when faced variants. But this approach relies heavily on the accuracy of the extracted skeletal joint configuration. To represent temporal information, both silhouette and skeleton animations require constructing a sequence of frames rather than just a single frame. A sequence of silhouette or skeleton can effectively represent one's gait based on gait cycle theory illustrated in section Gait Cycle5.

Cai et al. (2021) proposed an integrated human body representation method combining both skeleton and silhouette information. By leveraging state-of-the-art pose estimation

techniques, robust human skeletons can be extracted from monocular RGB images. A fusion strategy is then applied to compactly encode both skeleton and silhouette cues into a single silhouette-skeleton gait descriptor. Preliminary results demonstrate the efficacy of this hybrid representation for accurate and efficient human body modeling from visual data alone.

Dimension reduction

Dimension reduction (*Ma & Zhu, 2013*; *Carreira-Perpinán, 1997*) plays a pivotal role in the gait recognition process, as it addresses the challenge of dealing with high-dimensional data while preserving the most relevant information for identification (*Santos et al., 2022*). The extracted features from the gait data, such as silhouettes, temporal sequences, or kinematic parameters, often result in a feature space that is too large to be processed efficiently. Moreover, not all features contribute equally to the recognition task; some may even introduce noise or redundancy. Dimension reduction techniques aim to mitigate these issues by transforming the original high-dimensional feature space into a lower-dimensional one, enhancing the performance and computational efficiency of the gait recognition system.

This subsection provides an overview of the significance and types of dimension reduction methods commonly employed in gait recognition research. We categorize these methods into two main types: feature reduction and outliers removal. Feature reduction techniques focus on selecting or transforming the most relevant features, while outliers removal aims to eliminate noisy or abnormal data points that may negatively impact the recognition process (*Marsico & Mecca, 2019*).

Feature reduction techniques encompass various algorithms, such as principal component analysis (PCA) (*Kherif & Latypova, 2020*), linear discriminant analysis (LDA) (*Balakrishnama & Ganapathiraju, 1998*), tensor-based methods, and wavelet transforms. These methods help in capturing the essential characteristics of the gait data, while reducing the dimensionality to a more manageable level. For instance, PCA (*Kherif & Latypova, 2020*) and LDA (*Balakrishnama & Ganapathiraju, 1998*) are widely used to identify the principal components or discriminative features that best represent the variance or class separation in the dataset. Tensor-based methods like general tensor discriminant analysis (GTDA) (*Khan, Farid & Grzegorzek, 2023*) and sparse bilinear discriminant analysis (SBDA) (*Lai et al., 2014*) are employed to exploit the multi-linear structure of gait data, preserving the most informative features. Additionally, wavelet transforms, such as discrete wavelet transform (DWT) (*Zhang, 2019*), are utilized for decomposing the gait signals into different frequency bands, filtering out noise, and reducing dimensionality.

By effectively reducing the dimensionality of the gait data, these methods contribute to improving the accuracy, speed, and robustness of gait recognition systems. Furthermore, they facilitate the visualization of data and enable more efficient storage and transmission of gait templates.

Deep learning-based classification approaches

In the gait recognition pipeline, classification constitutes a pivotal process. Traditional machine learning (*Mahesh, 2020*; *Alpaydin, 2021*; *Jordan & Mitchell, 2015*; *El Naqa & Murphy, 2015*; *Zheng et al., 2024b*; *Li et al., 2023a*) classifiers, including support vector machines (SVM) (*Xue et al., 2010*; *Kamruzzaman & Begg, 2006*), hidden Markov models (HMM) (*Panahandeh et al., 2013*; *Nickel et al., 2011*), K-nearest neighbors (K-NN) (*Choi et al., 2014*; *Bajwa, Garg & Saurabh, 2016*), Bayesian (*Chen, Liang & Zhu, 2011*; *Gupta et al., 2015*), and random forest (*Kececi et al., 2020*; *Shi et al., 2020*), have played a fundamental role in this domain, have been widely discussed in *Rani & Kumar (2023)*. Although these traditional machine learning classifiers have achieved some success in gait recognition, deep learning classifiers have garnered increasing attention in recent years, owing to their superior capabilities in handling large-scale datasets and complex feature extraction (*Parashar et al., 2024*). We mainly focus on deep learning-based classifier in this subsection.

Convolutional neural networks

Convolutional neural networks (CNNs) (*LeCun et al., 1998*) have emerged as a powerful tool for gait recognition, offering superior performance in capturing discriminative features from gait data. CNNs are capable of automatically learning hierarchical feature representations (*Pinaya et al., 2020*), eliminating the need for manual feature engineering (*Zheng & Casari, 2018*; *Adamczyk & Malawski, 2021*; *Krüger, Berger & Leich, 2019*). This is achieved through the use of convolutional layers to detect local features, followed by pooling layers to aggregate these features into more abstract representations.

Several studies have made significant contributions to the field of gait recognition using CNNs. *Shiraga et al. (2016)* introduced a CNN model designed for cross-view invariant gait recognition, demonstrating the effectiveness of a shallow network structure in this context. *Xu et al. (2020)* explored the use of spatial transformer networks to adaptively learn transformations that mitigate the effects of view variation, thereby enhancing the discrimination capability of the gait recognition system. *Huang et al. (2021)* presented a model that balances computational efficiency and recognition accuracy, making it ideal for wearable devices with limited computational resources.

Other studies have focused on improving the robustness and accuracy of gait recognition systems. *Hasan et al. (2022)* introduced an ensemble learning approach that leverages multiple CNN models and a secondary classifier to improve the precision of gait classification. *Bari & Gavrilova (2022)* proposed a CNN architecture that utilizes a unique 3D input representation to automate feature extraction for gait recognition, streamlining the process and enhancing performance.

Recent studies have built upon these contributions, introducing innovative approaches that further advance the field of gait recognition. *Fan et al. (2020)* proposed GaitPart, a temporal part-based model that addresses the varying shapes and moving patterns of different body parts during walking. This model employs a novel framework that includes

the frame-level part feature extractor (FPFE) and the micromotion capture module (MCM), effectively capturing local short-range spatio-temporal features crucial for discriminative gait representation.

In pursuit of practicality and ease of use, *Fan et al. (2023)* proposed GaitBase, a simple yet efficient model that serves as a new baseline for gait recognition studies. *Lin, Zhang & Yu (2021)* addressed the challenge of cross-view variation by presenting a framework that leverages both global and local feature representation, along with local temporal aggregation. This method enhances feature representation by incorporating 3D convolution and a statistical function to aggregate temporal information, adaptively capturing the spatial and temporal aspects of gait sequences.

Capsule networks

The advent of CNNs (*LeCun et al., 1998*) has revolutionized the field of computer vision, but they still possess certain limitations. One major drawback is their inability to capture spatial hierarchies and relationships between features, leading to a lack of robustness to affine transformations and viewpoint changes. To address these limitations, a new paradigm has emerged: capsule networks (CapNets) (*Sabour, Frosst & Hinton, 2017*). Building upon the strengths of CNNs, CapNets introduce a novel architecture that leverages the concept of "capsules" to represent objects as a nested set of vectors, rather than as a single scalar value.

The application of capsule networks in gait recognition has shown promising results, offering an effective solution to the challenges posed by varying viewing angles, clothing, and obstructions such as bags (*Sepas-Moghaddam et al., 2021*). Unlike traditional CNNs, which often struggle with the part-whole relationship of gait features, capsule networks have demonstrated the ability to capture deeper part-whole relationships and act as attention mechanisms to enhance robustness under different conditions (*Sepas-Moghaddam et al., 2021*).

The architecture of capsule networks typically comprises two main components: convolution-based and capsule-based parts (*Xu et al., 2019*). The convolution-based part serves as a local feature detector, converting pixel intensities into activities of local feature detectors. In contrast, the capsule-based part is akin to an inverse render process, striving to construct better representations for each entity (*Xu et al., 2019*). This dual structure allows the networks to predict the similarity between pairs of images, such as gait energy images (GEIs) or colored gait images (CGIs), by computing the differences between input features and synthesizing the properties of each entity (*Xu et al., 2019*).

One of the key advantages of capsule networks in gait recognition is their ability to learn complex projections and effectively recognize gait patterns under varying conditions using dynamic routing algorithms (*Xu et al., 2019*). This is in contrast to traditional CNNs, where the loss of information during pooling and the inability to accurately reflect hierarchical relationships can lead to reduced recognition accuracy (*Sepas-Moghaddam et al., 2021*). Research has also investigated the optimal depth of the capsule layer, with findings suggesting that increasing the depth does not always lead to better performance (*Xu et al., 2019*). For instance, in gait recognition, a four-capsule layer network has been

observed to outperform a meta-network with additional layers, indicating that there is an optimal depth for capturing the relevant properties of gait (*Xu et al., 2019*). The introduction of a feedback weight mechanism in capsule networks further enhances their capability in gait recognition (*Sepas-Moghaddam et al., 2021*). By iteratively refining the coupling coefficients between low-level and high-level capsules, the network can better measure the probability of the occurrence of an entity, thereby improving recognition accuracy (*Sepas-Moghaddam et al., 2021*).

The ability of CapNets (*Sabour, Frosst & Hinton, 2017*) to capture part-whole relationships, their robustness to viewpoint changes and clothing variations, and the introduction of dynamic routing and feedback weights all contribute to a more accurate and reliable gait recognition system.

Graph convolutional networks

The proliferation of graph-structured data in various real-world applications has underscored the limitations of traditional convolutional neural networks (CNNs) (*LeCun et al., 1998*) and capsule networks (CapsNets) (*Sabour, Frosst & Hinton, 2017*) in handling such data. To address this limitation, graph convolutional networks (GCNs) (*Zhang et al., 2022*) have emerged as a powerful tool for graph-based learning. By extending the convolutional operation to graph domains, GCNs enable the extraction of meaningful features from graph-structured data, such as social networks, molecular structures, and traffic patterns. Unlike CNNs and CapsNets, which operate on grid-like structures, GCNs can handle irregularly structured data, making them particularly well-suited for tasks such as node classification, graph classification, and link prediction.

Gait analysis, which involves the study of human walking patterns, is inherently graph-structured due to the complex relationships between joints and the temporal dynamics of movement. This has led to the exploration of GCNs for gait-based recognition tasks, including identity and emotion recognition from gait patterns (*Sheng & Li, 2021*), abnormal gait detection (*Tian et al., 2022*), gait phase classification for lower limb exoskeleton systems (*Wu et al., 2021*), and gait recognition based on 3D skeleton data (*Mao & Song, 2020*).

The attention-enhanced temporal graph convolutional network (AT-GCN) (*Sheng & Li, 2021*) introduces spatial-temporal attention mechanisms to capture discriminative features from both the spatial dependence of nodes and the temporal dynamics of frames. This allows for the simultaneous recognition of identity and emotion from gait, which is a novel approach that leverages multi-task learning to enhance performance, particularly in scenarios where data is limited. By sharing information across related tasks, the AT-GCN model demonstrates superior performance compared to previous state-of-the-art methods.

Furthermore, the application of spatio-temporal attention-enhanced gait-structural graph convolutional networks (AGC-GCN) (*Tian et al., 2022*) has shown promise in the detection of abnormal gait patterns. These networks utilize graph convolution to analyze the structural information of human gait, taking into account both spatial and temporal relationships between joints. This method addresses the challenges posed by

non-Euclidean data structures, where the adjacency of nodes is not fixed or orderly, and the input order of the dataset is not strict due to the graph-based nature of the data.

In the context of gait phase classification for exoskeleton systems (*Wu et al., 2021*), a GCN model is proposed that can learn from graph-structured data, avoiding many issues associated with gait disorders on Euclidean domains. This model employs a localized first-order approximation of spectral graph convolutions to smooth gait label information over the graph, requiring labeling for only a small subset of nodes.

For gait recognition based on 3D skeleton data (*Mao & Song, 2020*), researchers have adapted spatial-temporal graph convolutional networks (ST-GCN) to model the structural information between joints in both space and time dimensions. This approach treats joint points as vertices, joint coordinates as vertex attributes, and utilizes spatial and temporal edges to capture the dynamic relationships within the gait sequence.

Deep belief networks

Deep belief networks (DBNs) (*Hinton, Osindero & Teh, 2006*) are a type of neural network architecture composed of multiple layers of restricted Boltzmann machines (RBMs) (*Hinton, 2002*), which are trained in a greedy, layer-by-layer manner. This hierarchical structure enables DBNs to learn complex, abstract representations of data, rendering them particularly well-suited for tasks such as dimensionality reduction, feature learning, and generative modeling. In contrast to other deep learning models, DBNs are trained using an unsupervised learning approach, which allows them to capture underlying patterns and structures in the data without the need for labeled examples (*Hinton, Osindero & Teh, 2006*). The success of DBNs has paved the way for the development of other deep learning models, including deep autoencoders and generative adversarial networks (GANs).

The application of DBN architecture in gait recognition primarily focuses on two aspects: feature extraction and classification. Feature extraction is a crucial step in gait analysis, as it aims to capture the most relevant information from raw data, such as electromyography (EMG) signals, which can be noisy and high-dimensional. DBNs have been demonstrated to effectively learn hierarchical representations of these signals, facilitating the identification of key features that contribute to gait classification.

A recent study *He et al. (2022)* proposed a method that combines multi-feature fusion and an optimized DBN for EMG-based human gait classification. The layer-by-layer learning feature of DBN was leveraged to capture high-order correlations in the original input data, leading to improved accuracy in recognizing different stages of gait. Furthermore, the use of the sparrow search algorithm (SSA) (*Xue & Shen, 2020*) for optimizing DBN parameters, such as the number of neurons and fine-tuning iterations, eliminated human-made interference and enhanced the recognition rate and stability of the model. The DBN architecture is particularly advantageous in gait recognition, as it addresses the distribution differences in feature sets caused by gait variations (*He et al., 2022*). By learning layer-by-layer, DBN can discover features that are more robust to noise and individual differences, which are common challenges in gait analysis. The

incorporation of MoCA features into the model, for instance, improved the performance of DBN, demonstrating its flexibility in integrating multifaceted data (*Fernandes et al., 2018*).

A comparative study between DBN and other machine learning models, such as multi-layer perceptrons (MLPs) (*Haykin, 2009*), has also demonstrated the superior performance of DBN in terms of sensitivity and specificity for gait classification (*Fernandes et al., 2018*). This is likely attributed to DBN's ability to learn deeper representations and capture non-linear relationships present in gait data.

Recurrent neural networks

Recurrent neural networks (RNNs) (*Bengio, Simard & Frasconi, 1994*) are a paradigm of neural network architecture specifically designed to handle sequential data, such as time series data, speech, text, and video. Unlike traditional feedforward neural networks, RNNs possess feedback connections that enable them to maintain a hidden state that captures information from previous time steps. This allows RNNs to learn long-term dependencies and patterns in the data, rendering them particularly well-suited for tasks such as language modeling, machine translation, and speech recognition. However, traditional RNNs suffer from the vanishing gradient problem, which makes it challenging to train them on long sequences. To address this issue, variants of RNNs, such as long short-term memory (LSTM) (*Hochreiter & Schmidhuber, 1997*) and gated recurrent unit (GRU) (*Cho et al., 2014*) networks, have been developed, which utilize specialized memory cells to store information over long periods of time. These variants have achieved state-of-the-art performance in a wide range of natural language processing and sequence modeling tasks.

One of the pioneering works in applying RNNs to gait recognition is the use of Faster R-CNN combined with LSTM networks to recognize gait patterns both with and without carried objects (*Ghosh, 2022*). This approach addresses the challenge of gait recognition when individuals are carrying objects, which significantly alters their walking patterns. By employing Faster R-CNN (*Ren et al., 2016*) to detect and extract pedestrian features from video frames and LSTM to model the temporal dependencies within the gait sequences, this method demonstrates a significant advancement in gait recognition technology. Furthermore, the application of RNNs for continuous authentication through gait analysis has been explored (*Giorgi, Saracino & Martinelli, 2021*). This involves the use of an LSTM-based RNN that analyzes inertial signals from smartphones to classify user activities, such as walking, running, or climbing stairs. The ability of RNNs to process time-dependent inertial data makes them an effective tool for continuous authentication, enhancing security in mobile devices. The efficacy of RNNs in gait recognition has also been demonstrated in studies focusing on inertial gait user recognition in smartphones (*Fernandez-Lopez et al., 2019*). These studies utilize RNNs to extract feature vectors from segmented gait cycles, which are then compared to recognize users. The use of dynamic RNNs capable of handling gait cycles of varying lengths has proven to be particularly useful in this context, improving the scalability and performance of gait recognition systems.

Generative adversarial networks

GANs (*Goodfellow et al., 2020*) are a type of deep learning model that has revolutionized the field of generative modeling. GANs consist of two neural networks: a generator and a discriminator. The generator network takes a random noise vector as input and produces a synthetic data sample that attempts to mimic the real data distribution. The discriminator network, on the other hand, takes a data sample (either real or synthetic) as input and outputs a probability that the sample is real. The two networks are trained simultaneously, with the generator trying to produce realistic samples that can fool the discriminator, and the discriminator trying to correctly distinguish between real and synthetic samples. Through this adversarial process, GANs can learn to generate highly realistic data samples that are often indistinguishable from real data. In the context of gait recognition, GANs have been used to generate synthetic gait patterns that can be used to augment existing datasets, improve recognition performance, and enhance robustness to variations in gait patterns.

One of the pioneering works in this area is Silhouette Guided GAN (SG-GAN) (*Jia et al., 2019*), which proposes an attack method on gait recognition systems by synthesizing videos that combine a source walking sequence with a target scene image. This approach demonstrates the capability of GANs to generate realistic gait videos, which can be used to test and improve the robustness of recognition systems. A crucial aspect of gait recognition is identity preservation, which has been addressed by VN-GAN (*Zhang, Wu & Xu, 2019a*). This architecture employs a coarse-to-fine generation process that normalizes the view angle and preserves identity-related information. By integrating reconstruction and identity-preserving losses, VN-GAN ensures that the generated images are both view-identical and identity-discriminative, leading to promising performance in gait recognition tasks. Building upon these advancements, multi-view gait image generation (*Chen et al., 2021*) explores the use of GANs for synthesizing pedestrian images under different conditions, which can be used to enhance cross-view gait recognition. This work highlights the potential of GANs in generating diverse pedestrian appearances and poses, contributing to a more robust recognition system. Furthermore, VT-GAN (*Zhang, Wu & Xu, 2019b*) introduces a view transformation GAN that can translate gaits between any two views using a single model. This is a significant advancement over previous GAN-based models that normalize to a single unified view. VT-GAN's ability to preserve identity similarity while performing view transfer makes it a powerful tool for cross-view gait recognition and data augmentation. More recently, MaP-SGAN (*Liang et al., 2024b*) has been proposed as an innovative approach that incorporates multi-anchor point metrics for wireless signal spectrum data assessment. This method adapts the GAN framework to handle high-dimensional features, such as gait and domain information, by extending the discriminator's output to a distribution that captures data authenticity from multiple perspectives.

Large vision models

Large vision models (LVMs) are a class of deep neural networks that have achieved state-of-the-art performance in various computer vision tasks, including image

classification, object detection, and segmentation. LVMs are characterized by their massive scale, with millions of parameters and complex architectures that require significant computational resources to train. Despite their computational demands, LVMs have demonstrated remarkable capabilities in learning rich and robust representations of visual data, often surpassing human-level performance in certain tasks.

The application of LVMs in gait recognition has been a significant breakthrough, as evidenced by the pioneering work BigGait (*Ye et al., 2024*), which introduced a novel framework for extracting gait representations from task-agnostic LVMs. By leveraging the all-purpose knowledge encoded in these models, BigGait (*Ye et al., 2024*) demonstrated the potential to transcend the limitations of traditional, task-specific upstream models that depend heavily on supervised learning and annotated datasets. The LVMs-based approach has not only reduced the dependency on labor-intensive annotation for gait-specific features like silhouettes and skeletons but also avoided the error propagation that can occur from upstream task-specific models.

Performance evaluation metric

In evaluating gait recognition methods, a critical initial step is partitioning the dataset into a gallery set and a probe set. The gallery set, containing pre-enrolled gait sequences for each individual, serves as a reference for identifying individuals, while the probe set introduces new sequences to test recognition accuracy. This setup enables systematic performance assessment across varied occlusion scenarios, simulating real-world identification challenges. Evaluation protocols such as closed-set, open-set, and cross-dataset testing further enhance assessment rigor. In the closed-set protocol, all individuals in the probe set are present in the gallery, allowing for controlled testing of identification accuracy within known subjects. Conversely, the open-set protocol challenges the model with individuals in the probe set who are not in the gallery, simulating real-world applications where subjects may vary. Cross-dataset evaluation adds an additional layer of complexity, testing models on entirely different datasets to assess generalization across new environments, camera setups, or occlusion types. Complementing these partitioning strategies, several key metrics serve as indicators of performance under occlusion, notably Rank-N accuracy, false acceptance rate (FAR), and equal error rate (EER). Rank-N accuracy gauges the model's ability to identify the correct individual within the top-N results, an essential feature in occluded scenarios where visual data may be partially obstructed. FAR and EER provide insights into robustness under occlusion, measuring incorrect acceptances and the balance between acceptance and rejection errors, respectively. A low FAR, in particular, highlights the system's capability to avoid mistaking occluded patterns for unauthorized individuals. Together, these protocols and metrics offer a comprehensive framework for assessing a model's robustness, accuracy, and adaptability to occlusion in gait recognition, making them indispensable for understanding a method's suitability for real-world applications.

Dataset partitioning: gallery set and probe set

In gait recognition, performance evaluation typically involves splitting the dataset into two distinct sets: the gallery set, $\mathcal{G}$, and the probe set, $\mathcal{P}$. The gallery set contains gait sequences for enrolled individuals, where each individual is represented by a small number of gait sequences, denoted as $\{g_i\}_{i=1}^{N}$, with N as the number of individuals in the gallery. These sequences serve as templates or models for each individual's gait pattern. The probe set, $\mathcal{P}$, consists of gait sequences used to test the system's ability to recognize individuals. Each individual in this set is represented by gait sequences, $\{p_i\}_{i=1}^{M}$, where M is the number of individuals in the probe set. The probe set often includes variations in conditions such as lighting, camera angles, walking speed, and occlusion to assess the system's robustness. This partitioning allows researchers to evaluate how well the system generalizes to new, unseen data, which is essential for real-world applications. Formally, the dataset can be represented as:

$$\mathcal{D} = \mathcal{G} \cup \mathcal{P} \tag{1}$$

where $\mathcal{D}$ is the entire dataset, $\mathcal{G}$ is the gallery set, and $\mathcal{P}$ is the probe set (see Eq. (1)).

Evaluation protocols: closed set, open set, and cross set

Closed-set evaluation. In a closed-set protocol, all individuals in the probe set $\mathcal{P}$ are also present in the gallery set $\mathcal{G}$. This scenario is often used for controlled identification tasks, where the system must identify an individual from a known set. Formally:

$$\mathcal{P} \subseteq \mathcal{G} \tag{2}$$

Open-set evaluation. In open-set evaluation, some individuals in the probe set $\mathcal{P}$ are not present in the gallery set $\mathcal{G}$. This scenario is used in more challenging settings, such as verification tasks, where the system must handle unknown individuals. Formally:

$$\mathcal{P} \cap \mathcal{G} = \varnothing \tag{3}$$

Cross-dataset evaluation. For cross-dataset evaluation, the system is tested on a dataset different from the one used for training, evaluating the system's generalization to unseen conditions such as new camera viewpoints or lighting setups. Formally:

$$\mathcal{D}_{train} \neq \mathcal{D}_{test} \tag{4}$$

Single-shot and multi-shot evaluation. In single-shot evaluation, only one gait sequence per individual is used in the probe set, whereas in multi-shot evaluation, multiple sequences are used to represent each individual. Multi-shot evaluation provides more data for testing recognition performance, often yielding more accurate results.

$$\mathcal{P} = \{p_i\}_{i=1}^{M} \tag{5}$$

where M is the number of gait sequences per individual (see Eq. (5)).

Rank-N accuracy

Rank-N accuracy is a common metric in gait recognition that evaluates the system's ability to correctly identify an individual within the top-N matches from the gallery. A Rank-1 accuracy of 90% means that the system correctly identifies the individual in the top match 90% of the time, while Rank-5 accuracy of 95% means that the system identifies the correct individual within the top five matches 95% of the time. The cumulative match characteristic (CMC) (*Bolle et al., 2005*) curve is often used to visualize Rank-N accuracy across different values of N.

Rates: FAR, FRR, EER, DIR

To assess system reliability, several key metrics are used:

False acceptance rate (FAR). The FAR represents the probability that an imposter is incorrectly accepted as a genuine individual. It is defined as:

$$FAR = \frac{N_{FA}}{N_I} \tag{6}$$

where N_{FA} is the number of false acceptances, and N_I is the number of imposter attempts (see Eq. (6)).

False rejection rate (FRR). The FRR is the probability that a genuine individual is incorrectly rejected. It is given by:

$$FRR = \frac{N_{FR}}{N_G} \tag{7}$$

where N_{FR} is the number of false rejections, and N_G is the number of genuine attempts (see Eq. (7)).

Equal error rate (EER). The EER is the point at which the FAR and FRR are equal, serving as a balanced indicator of overall system performance.

Detection and identification rate (DIR). DIR measures the proportion of correctly identified individuals among all detected individuals, and is defined as:

$$DIR = \frac{N_{CI}}{N_D} \tag{8}$$

where N_{CI} is the number of correctly identified individuals, and N_D is the total number of detected individuals (see Eq. (8)).

These metrics provide a thorough evaluation of the system's accuracy and robustness in handling real-world gait recognition tasks.

Application scenarios

Gait recognition, with its non-intrusive identification approach (*Devarapalli, Dhanikonda & Gunturi, 2020*), offers promise in various applications. Unlike traditional biometrics requiring active cooperation, gait recognition identifies individuals through their walking patterns, often without their knowledge. As conventional biometrics face increased forgery risks, gait recognition's resistance to imitation makes it a more secure option. In this section, we outline potential gait recognition applications.

Forensic gait recognition technology, as demonstrated in recent research, finds crucial applications in crime investigation (*Marra, Di Silvestro & Somma, 2023*) and forensics (*Xu et al., 2023a*). When video footage records a suspect's walking pattern, algorithms analyze it to extract unique biometric data, facilitating identity verification through database matching. Gait's inherent resistance to disguise or forgery enhances police tracking of suspect movements over time.

Access security control gait recognition technology enhances security access control systems (*Ali, Sabir & Ullah, 2019*) by verifying individual identities based on unique walking patterns. Unlike traditional methods susceptible to forgery, such as ID cards, passwords, or recent biometrics like face and fingerprint recognition, gait authentication is exceptionally hard to fake or mimic. This robust approach bolsters security and deters unauthorized access to sensitive areas and buildings.

Border security control. The gait recognition authentication system, when used at border checkpoints (*Ngo & Hung, 2019*), enhances security by accurately identifying travelers. It collects and analyzes gait data, extracting distinctive features and cross-referencing them with an identity database for swift and precise identification. This non-intrusive approach eliminates the need for additional documents or passwords, reducing the risk of errors.

High traffic scenarios. In highly populated areas such as transportation hubs, shopping districts, and streets, surveillance systems play a critical role in urban infrastructure. Within surveillance videos, "pedestrians" are a common focal point. Manual visual identification is labor-intensive, time-consuming, and prone to both visual fatigue (*Souchet et al., 2022*) and misidentification (*Jung et al., 2020*) over extended periods. Therefore, the adoption of gait recognition authentication technology is essential for efficient and accurate surveillance in such scenarios.

In such situational contexts, the phenomenon of occlusion is quite a regular occurrence. Various real-world scenarios involve complex spatial arrangements where one aspect of the visual scene is partially or fully blocked from sight by intervening obstructions. Figure 2 demonstrates two most common types of real-world occlusion.

TAXONOMY

Although occluded gait recognition is not a widely explored area, it remains crucial to propose a taxonomy to provide a comprehensive understanding of various occlusions encountered in real-life scenarios, datasets, and methodologies, as depicted in Fig. 6. In practical situations, occlusions can be broadly categorized as static or dynamic, but other types also exist, including salt and pepper noise (*Azzeh, Zahran & Alqadi, 2018*), frame loss, and limited field of view. Regarding datasets, beyond those with native occlusions, researchers frequently introduce synthetic occlusions into public datasets, which can be categorized into three types: code framework, model-based, and customized approaches. From a methodological perspective, there are diverse ways to capture body and temporal information. Techniques for addressing occluded gait recognition can be divided into two primary groups: reconstruction-based and reconstruction-free methods. Building on this taxonomy, we can proceed with our review.
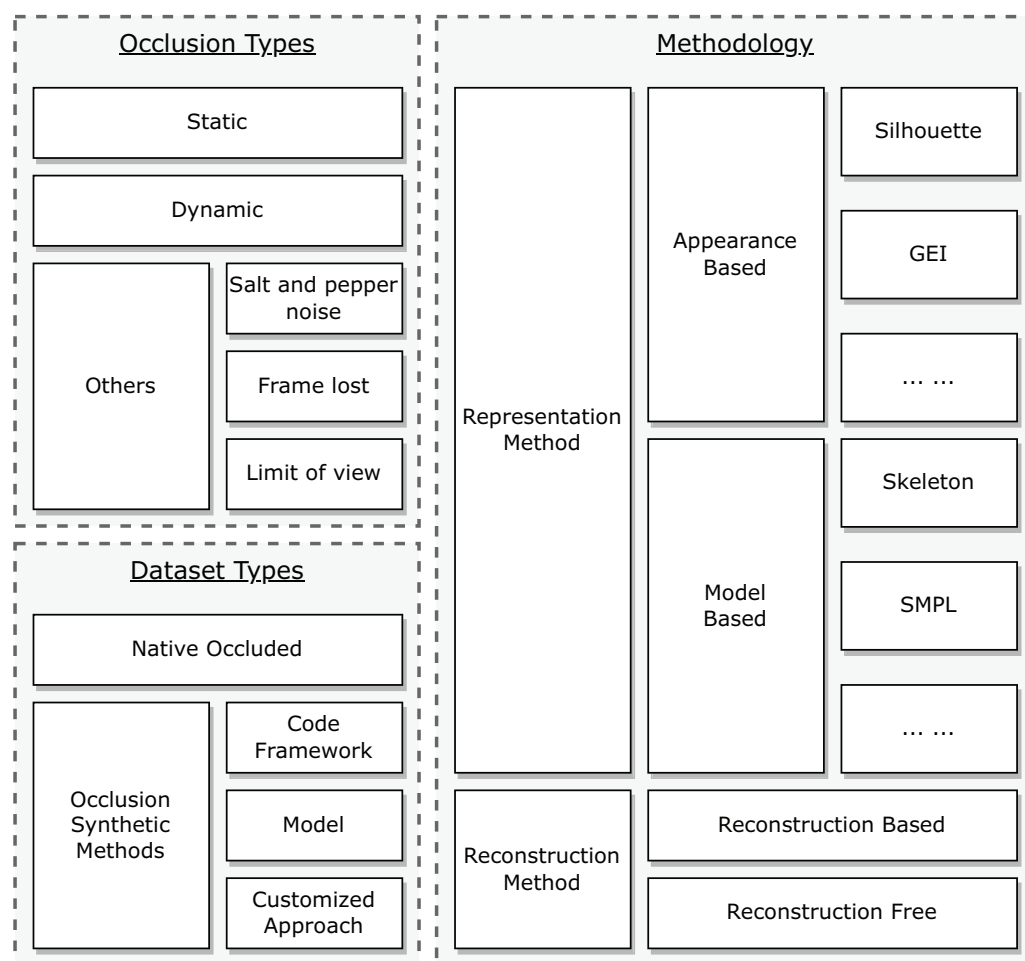
Figure 6 Taxonomy: occlusion, dataset, and methodology.

Full-size ☐ DOI: 10.7717/peerj-cs.2602/fig-6

REVIEW OF DATASETS

Deep gait recognition has emerged as the dominant approach in the field, primarily owing to its exceptional performance (*Prakash, Kumar & Mittal, 2018*; *Sepas-Moghaddam & Etemad, 2022*; *Narayan et al., 2023*). It is worth noting that data plays a fundamental role in both the training and testing phases of deep models. In this section, datasets are divided into two categories according to the taxonomy we proposed in section Taxonomy, namely public datasets with native occlusion and occlusion synthetic methods on public datasets. We conducted a thorough review of both occluded gait datasets and occlusion synthetic methods on available literature.

Public datasets with native occlusion

The study of gait recognition under occlusion conditions has been greatly facilitated by the release of various datasets. Table 2 provides a detailed overview of 10 popular datasets that incorporate occlusion scenarios, each offering a distinct set of challenges and environments. For instance, OccGait (*Huang et al., 2025*) and SUSTech1K

Table 2 Public gait datasets with occlusion.

Dataset	Year	Data type	Subjects	Sequences	Views	Challenges	Environment
OccGait (Huang et al., 2025)	2025	RGB Silhouette	101	>80,000	8	Occlusion	Indoor
SUSTech1K (Shen et al., 2023)	2023	3D Point Cloud RGB Silhouette	1,050	25,239	12	Views occlusion clothing illumination	Outdoor
CCPG (Li et al., 2023b)	2023	RGB	200	>16,000	10	Views occlusion illumination	Indoor Outdoor
OG RGB+D (Li & Zhao, 2023)	2023	RGB Depth Depth joints Silhouette The optimized joints	96	58,752	8	Views occlusion clothing illumination	Indoor
VersatileGait (Zhang et al., 2023)	2023	Silhouette	10,000	>1,000,000	44	View distractor carrying dressing occlusion surface illumination speed shoes trajectories	Wild
Gait3D (Zheng et al., 2022)	2022	Silhouette 2D Pose 3D Pose 3D Mesh SMPL	4,000	25,309	39	View distractor carrying dressing occlusion surface illumination speed shoes trajectories	Wild
GREW (Zhu et al., 2021)	2021	Silhouette Flow 2D Skeleton 3D Skeleton	26,345	128,671	882	View distractor carrying dressing occlusion surface illumination speed shoes trajectories	Wild

Table 2 (continued)							
Dataset	Year	Data type	Subjects	Sequences	Views	Challenges	Environment
Multi Gait (*Singh et al., 2019*)	2019	RGB	20	240	3	Occlusion	Outdoor
Depth Gait (*Chattopadhyay, Sural & Mukherjee, 2015*)	2015	Depth 2D Skeleton	29	464	2	View occlusion speed	Indoor
TUM-IITKGP (*Hofmann, Sural & Rigoll, 2011*)	2011	Silhouette	35	840	1	Time carrying shoes occlusion	Indoor

(*Shen et al., 2023*) capture data under indoor and outdoor conditions, respectively, with different types of occlusions such as static and crowd occlusions for OccGait and more diverse challenges including clothing and illumination for SUSTech1K (*Shen et al., 2023*). Datasets such as OG RGB+D (*Li & Zhao, 2023*) further leverage multi-modal data to enhance recognition under occluded conditions. The figure setups in Fig. 7 highlight the experimental environments for these datasets. For example, SUSTech1K (*Shen et al., 2023*) uses a LiDAR sensor and an RGB camera setup, while the OG RGB+D (*Li & Zhao, 2023*) dataset relies on a multi-Kinect SDAS system. Similarly, Depth Gait captures depth and skeleton data using Kinect sensors. VersatileGait (*Zhang et al., 2023*), another significant dataset, simulates occlusions in a game-engine environment, creating realistic walking animations. These datasets collectively provide a comprehensive platform for testing gait recognition algorithms under varying occlusion conditions.

OccGait. The OccGait benchmark is a pioneering occluded gait recognition dataset designed to address the challenges posed by real-world occlusions in gait recognition. It includes 101 subjects, with over 80,000 sequences captured from eight camera views. OccGait introduces four types of occlusion scenarios—None (normal walking), Carrying (with items like umbrellas or luggage), Crowd (interactions with other individuals), and Static (obstructions like plants or chairs). The dataset provides detailed annotations of occlusion types, offering a robust platform for the analysis and evaluation of gait recognition algorithms under varying occluded conditions. Data collection took place in an indoor lab, and silhouettes were extracted using MaskFormer for preprocessing. By simulating realistic occlusions, OccGait facilitates both qualitative and quantitative assessments, serving as a foundation for developing more resilient gait recognition systems. Extensive experiments on this benchmark demonstrate significant improvements in handling occlusion compared to previous methods (*Huang et al., 2025*).

SUSTech1K. The SUSTech1K dataset comprises a total of 1,050 subjects, encompassing 25,239 gait sequences. Each subject is captured from 12 different views in outdoor environment, collected by a LiDAR sensor and an RGB camera. Occlusion is considered in
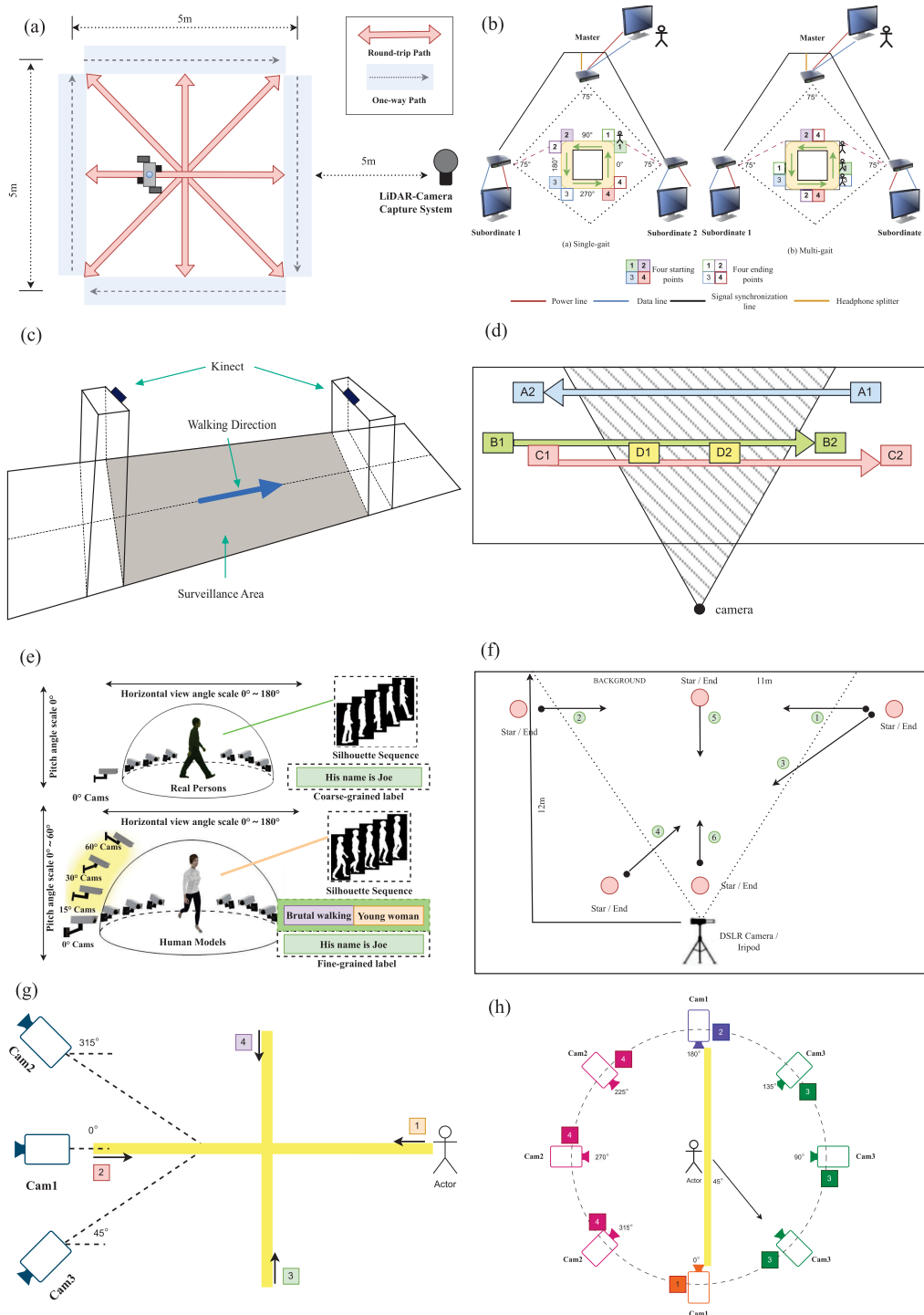
Figure 7 (A) SUSTech1K: Participants walked four round-trip and four one-way paths, first conventionally and then randomly varied. (B) OG RGB+D: Single-person gait data was captured using a multi-Kinect system, while multi-person data was collected as subjects moved in a straight line. (C) Depth Gait: Two cameras capture individuals from the front and back, using Kinect's depth and skeleton streams for recognition. (D) TUM-IITKGP: A camera positioned at 1.85 m in a corridor captures a subject walking from A1 to A2, with dynamic and static occlusions from other individuals. (E) VersatileGait: The dataset was made by creating 3D models, retargeting animations, simulating scenes, capturing silhouettes, and annotating attributes. (F) Multi Gait: The dataset includes recordings of pairs walking together and the same participants walking alone to introduce dynamic occlusion. (G) & (H) OccGait: The layout and camera angles were optimized for visibility and data capture. Full-size ☑ DOI: 10.7717/peerj-cs.2602/fig-7

process of data collection. SUSTech1K encompasses the diverse range of variations observed within established datasets, incorporating factors such as Normal, Bag, Clothes Changing, Views, and Object Carrying. Furthermore, it acknowledges the presence of other prevalent and demanding variations encountered in outdoor scenarios, including Occlusion, Illumination, Uniform, and Umbrella. By systematically categorizing these walking sequences into distinct subsets based on their specific variations, we can gain a comprehensive understanding of how different variations impact the performance of gait recognition for both modalities (*Shen et al., 2023*).

CCPG. The CCPG dataset consists of 200 subjects and over 16,000 sequences captured both indoors and outdoors. Each subject appears in seven distinct outfits, making it one of the few datasets specifically designed to handle variations in clothing, including changes to tops, pants, entire outfits, and the presence of bags. Unlike previous datasets that focus on controlled conditions, CCPG introduces real-world challenges such as varying illumination, occlusion, and viewing angles. Both RGB and silhouette data are available, enabling a comprehensive evaluation of person re-identification and gait recognition methods under changing clothing conditions. This dataset provides a valuable resource for researchers aiming to improve the robustness of gait recognition systems in practical applications (*Li et al., 2023b*).

OG RGB+D. The OG RGB+D dataset consists of a total 96 subjects, encompassing 58,752 gait sequences. Each subject is captured from 12 different views in indoor environment. By utilizing a multiple synchronous Azure Kinect DK sensors data acquisition system (multi-Kinect SDAS), this database overcomes the limitations of existing gait databases by including multimodal gait data with various types of occlusions, such as self-occlusion, active occlusion, and passive occlusion. The Azure Kinect DK sensors allow for the simultaneous collection of multimodal data, supporting different gait recognition algorithms and enabling the effective capture of camera-centric multi-person 3D poses. The use of multiple views provided by the multi-Kinect setup is advantageous in handling occlusion scenarios compared to single-view systems. Moreover, the OG RGB+D database offers accurate silhouettes and optimized human 3D joints data (OJ) by combining data collected from the multi-Kinects, which ensures improved accuracy in representing human poses under occlusion (*Li & Zhao, 2023*).

VersatileGait. The VersatileGait dataset consist of a total 10,000 subjects, encompassing more than one million gait sequences. Each subject is captured from 44 different views in wild simulation environment by toolkit Unity3D. It offers a synthetic yet realistic representation of gait variations encountered in real-world scenarios. By incorporating diverse pedestrian attributes, walking styles, and a complex camera layout, the dat aset aims to bridge the domain gap between synthetic and real data (*Zhang et al., 2023*).

Gait3D. The Gait3D dataset is an extensive collection of gait sequences, comprising a total of 4,000 subjects and 25,309 gait sequences. This dataset is particularly noteworthy due to its large sample size, which allows for a more comprehensive analysis of gait characteristics and variations among individuals. Furthermore, the Gait3D dataset encompasses a diverse range of viewing angles, with a total view number of 39. This multifaceted perspective facilitates the examination of gait patterns from different angles,

enabling a more thorough understanding of gait dynamics and kinematics. The Gait3D dataset is an invaluable resource for researchers in the field of gait analysis, providing a rich source of data for developing and validating gait recognition algorithms, as well as for studying the underlying mechanisms of gait disorders (*Zheng et al., 2022*).

GREW. The GREW dataset consists of a total of 26,345 subjects, encompassing 128,671 gait sequences. Each subject is captured from 882 different perspectives. Unlike other outdoor datasets, GREW is uncontrolled, which exhibits greater irregularity and complexity compared to controlled outdoor settings, representing what is commonly referred to as a wild environment. This research endeavors to shift the testing environment to real-life conditions. Existing benchmarks for gait recognition inadequately address the challenges encountered in fully unconstrained real-world scenarios. This underscores the necessity for more comprehensive and practical evaluation standards. GREW represents the pioneering effort in focusing on the wild environment, with the potential to serve as a significant milestone in the advancement of gait recognition (*Zhu et al., 2021*).

Multi Gait. The Multi Gait dataset is specifically designed to address the challenges associated with gait recognition in diverse conditions. With a total of 20 subjects, each captured from three different views, this dataset provides a comprehensive collection of 240 gait sequences. These sequences include variations in walking directions, such as back and forth, as well as the presence of dynamic occlusions, which are crucial for evaluating the robustness of gait recognition algorithms. Furthermore, the dataset also includes sequences without any occlusion, serving as a baseline for comparison. By incorporating such a wide range of conditions, the Multi Gait dataset offers valuable resources for researchers to develop and test gait recognition algorithms that can perform effectively in real-world scenarios. This dataset is an essential contribution to the field of computer vision and biometrics, and it will undoubtedly facilitate advancements in gait analysis and recognition technologies (*Singh et al., 2019*).

Depth Gait. The Depth Gait dataset comprises a total of 29 subjects, encompassing 464 gait sequences. Each subject is captured from both the back and the front views using two Kinect cameras mounted at the entry and exit points of a surveillance zone. Kinect, introduced by Microsoft in 2010, encompasses a series of motion sensing input devices. These devices are equipped with RGB cameras, infrared projectors, and detectors that facilitate depth mapping using structured light or time of flight calculations. The utilization of the dataset requires the involvement of Microsoft SDK due to the use of Kinect for data acquisition. The Depth Gait dataset specifically focuses on occlusion scenarios that occur in congested places such as airports, railway stations, and shopping malls, where multiple people enter the surveillance zone one after another, leading to the occlusion of the target (*Chattopadhyay, Sural & Mukherjee, 2015*).

TUM-IITKGP. The TUM-IITKGP dataset comprises a total of 35 subjects, encompassing 840 gait sequences. Each subject is captured from only one single view in indoor environment. This datset was created to simulate real-world surveillance scenarios. Specifically, this dataset addresses the problem of dynamic and static inter object occlusion. The physical setup involves a camera positioned at a height of 1.85 m in a narrow hallway, perpendicular to the direction of people walking from right to left and left

to right in the image, reflecting a potential real-world surveillance application (*Hofmann, Sural & Rigoll, 2011*).

While current datasets have made significant strides in advancing gait recognition under occluded conditions, there are several areas for improvement that could further elevate their utility. Future datasets should prioritize more realistic and diverse occlusion scenarios, integrate additional data modalities such as thermal or radar imaging, and expand environmental coverage to include a wider range of real-world conditions. Increasing subject diversity and including longitudinal data will also enhance the generalizability of models. Furthermore, better annotation and metadata, along with synthetic data augmentation, will allow for the creation of richer datasets that can comprehensively address the challenges posed by occlusion in gait recognition. These improvements will ultimately drive the development of more robust and reliable gait recognition systems.

Occlusion synthetic methods based on public gait dataset

Introducing occlusion into normal gait datasets is a relatively straightforward process. It is feasible to generate a occluded version dataset derived from a publicly available gait dataset that lacks occlusion. In the literature, many different kind of methods to generate occlusion pattern exist. In this section, we present a collection of synthetic methods employed by researchers to contaminate normal gait video or it's representation, thereby facilitating their investigations into occluded gait recognition. These methods serve as valuable tools for simulating occlusion scenarios and evaluating recognition algorithms under controlled conditions, with a summary provided in Table 3.

Frameworks and models

Peng, Cao & He (2023) proposed an occlusion gait dataset composition framework called OccSilGait. The proposed framework allows control of occlusion probability and severity through two key parameters: occlusion probability (*occ_prob*) and occlusion ratio (*occ_ratio*). Higher *occ_prob* results in more occluded sequences in the generated dataset, thus increasing the difficulty of gait recognition. Higher *occ_ratio* signifies greater occlusion severity. By carefully considering obstacle size and location conditional placement avoids unrealistic results. The framework considers various occlusion scenarios-no occlusion, crowd occlusion, static occlusion, and detection occlusion. Based on proposed framework and CASIA-B gait dataset (*Yu, Tan & Tan, 2006*), OccCASIA-B was developed. It also introduces the SpaAlignTemOccRecover network, demonstrating superior performance when applied to this new dataset containing controlled occlusions. By systematically varying occlusion parameters and network design, this work advances gait recognition under real-world occlusion conditions.

Delgado-Escano et al. (2020) proposed an occlusion gait dataset composition framework called MuPeG that leverages two normal gait datasets as input. The framework merges the data from input datasets to generate a synthetic gait dataset with incorporated occlusions. Specifically, the procedure involves two input frames from different videos, each featuring distinct subjects. Through segmentation, a binary mask is obtained for the lower frame,

Table 3 Occlusion synthetic methods based on public gait dataset. FM stands for 'Framework and model', CA stands for 'Customized approaches'.

Reference	Year	Category	Static	Dynamic	Other	Dataset usage
Peng, Cao & He (2023)	2023	FM	✓	✓	✓	CASIA-B (*Yu, Tan & Tan, 2006*)
Xu et al. (2023b)	2023	CA	✓	✓	✗	OU-MVLP (*Takemura et al., 2018*)
Paul et al. (2022)	2022	CA	✗	✗	✓	CASIA-B (*Yu, Tan & Tan, 2006*), OU-ISLR LP (*Iwama et al., 2012*)
Delgado-Escano et al. (2020)	2020	FM	✓	✗	✗	CASIA-B (*Yu, Tan & Tan, 2006*), TUM-GAID (*Hofmann et al., 2014*)
Singh et al. (2020)	2020	CA	✗	✓	✗	Non-public Dataset
Uddin et al. (2019)	2019	CA	✓	✓	✗	OU-MVLP (*Takemura et al., 2018*)
Ortells et al. (2017)	2017	CA	✓	✓	✓	OU-ISIR Treadmill Dataset B (*Makihara et al., 2012*), Human ID Gait Database (*Sarkar et al., 2005*)
Nangtin, Kumhom & Chamnongthai (2016)	2016	CA	✓	✗	✗	CASIA-B (*Yu, Tan & Tan, 2006*)
Tang et al. (2016)	2016	CA	✓	✗	✗	CMU MoBo (*Collins, Gross & Shi, 2002*)
Roy et al. (2015)	2015	FM	✓	✓	✗	PETS 2010 (*Ellis & Ferryman, 2010*)
Roy et al. (2011)	2011	CA	✓	✗	✗	CMU MoBo (*Collins, Gross & Shi, 2002*)
Chen et al. (2009)	2009	CA	✓	✗	✗	CMU MoBo (*Collins, Gross & Shi, 2002*)

while a negative mask is computed for the upper frame. By performing a cropping operation using the masks and their corresponding frames, specific regions of interest are isolated. These modified frames are then aggregated by an element-wise addition

operation, resulting in a final output that incorporates both subjects, allowing for their simultaneous representation. The proposed framework presents a novel approach that enables the creation of previously non-existent types of datasets, thereby introducing new challenges for researchers in the field.

Roy et al. (2015) proposed an occlusion synthesis model. This study formulated occlusion synthesis as an optimization problem, seeking occluder parameters iteratively to meet occlusion goals. Particle swarm optimization (PSO) (*Eberhart & Kennedy, 1995*; *Wang, Tan & Liu, 2018*) is used explored the parameter space. The proposed occlusion model and methodology can generate static occlusion or dynamic occlusion from an occluder moving with or against the target seamlessly. Depending on the specified occlusion type, only the search domain for occluder parameters requires adjustment. By framing synthesis as optimization and leveraging particle swarm search, the approach allows for flexible generation of diverse occlusion effects in a computationally efficient manner.

Customized approaches

Xu et al. (2023b) added occlusion patterns by taking the OU-MVLP gait dataset (*Takemura et al., 2018*) containing over 10,000 subjects captured from 14 views as the base dataset. They selected four views for training and testing, and focused on vertical body occlusion. Two occlusion scenarios were simulated-fixed occlusion ratio where 20%, 40% and 60% of the height was occluded, and changing ratio where the occlusion varied from 60–20%, 40–0% and 20–0% over the gait sequence. A rectangular region was used to simulate the occlusion, and the remaining visible part was cropped as input. Samples with different views, occlusion patterns of fixed/changing ratios, and degrees of occlusion were then generated to train a unified gait recognition model, effectively introducing occlusion variations to the dataset in a systematic manner.

Paul et al. (2022) experiment on the TUM-IITKGP (*Hofmann, Sural & Rigoll, 2011*), the CASIA-B (*Yu, Tan & Tan, 2006*) and the OU-ISIR LP (*Iwama et al., 2012*) gait datasets. The latter two need synthetic occlusion to evaluate the performance. For each of them, two complete gait cycles were used to train recognition models, while four occluded cycles served as test sets. Synthetic occlusion was introduced by randomly selecting and blanking frames from test sequences, removing all silhouette data. This evaluation protocol allowed researcher to systematically study how varying degrees of synthetically inserted missing data impacted gait identification.

Singh et al. (2020) added occlusion patterns to the home-made dataset by creating random gaps, representing occlusions, at different time frames within the sequence of frames for various regions of interest (ROIs) such as the left ankle, left knee, left wrist, left elbow, and left shoulder. Specifically, gaps were inserted in the frontal view to simulate occlusion events, and the lengths of the gaps were varied between 0.2 to 0.3 s for each ROI and across different samples, with the gaps occurring in the first half, middle, and last half of each 30-frame sample sequence on average. This created occlusion patterns by simulating blocked visibility at different time points for the various ROIs.

Uddin et al. (2019) synthesized occlusion patterns on OU-MVLP (*Takemura et al., 2018*) using both static and dynamic occlusion. For dynamic occlusion, which involved simulating occlusion caused by objects moving from right to left. To create this pattern, a background rectangle mask was added to cover a certain area of the silhouette in the left-most position of the first frame. The position of the mask was gradually changed towards the right-most position as the frames progressed. This type of occlusion was referred to as relative dynamic occlusion from left to right (RDLR). Similarly, the author created dynamic occlusion from bottom to top (RDBT) by occluding the person from the bottom to the top. For static occlusion, which involved adding a fixed background mask in the bottom (RSB), top (RST), left (RSL), or right (RSR) positions for all frames in a sequence. Finally, random occlusion was simulated by adding a background mask in random positions horizontally (RandH) and vertically (RandV) across the silhouette sequence. Different degrees of occlusion (30%, 40%, and 50%) were applied to each type of occlusion, resulting in a total of 24 occlusion patterns.

Ortells et al. (2017) experiment on OU-ISIR Treadmill Dataset B (*Makihara et al., 2012*) and USF Human ID Gait Database (*Sarkar et al., 2005*). Three kinds of occlusion are used, namely static occlusion, dynamic occlusion, and salt & pepper noise. Salt and pepper noise was applied by randomly setting 75% of pixels in silhouette images to black or white, mimicking variable lighting conditions. For static occlusion, a weed silhouette was sequentially overlaid on human silhouettes to represent background objects blocking lower body parts. Dynamic occlusion used a car silhouette moving across human silhouettes, differently from static objects it segmented as foreground distorting silhouette shape.

Nangtin, Kumhom & Chamnongthai (2016) synthesizes occlusion on the CASIA-B gait dataset (*Yu, Tan & Tan, 2006*) by superimposing a black strip over normal side view sequences to create occluded samples. Specifically, 123 subjects with six normal sequences each (three for training, three for testing) from the CASIA-B dataset (*Yu, Tan & Tan, 2006*) are selected. Occlusion is introduced by overlaying a black strip over frames in the normal sequences. This is done for nine different combinations: the strip is positioned on the left, middle, or right side of the frame, and in the upper, middle, or lower position within that frame. This yields occluded sequences with the strip occluding different spatial regions across different samples. In total, this process synthesizes 2,952 occluded sequences from the CASIA-B dataset (*Yu, Tan & Tan, 2006*) for experiments, allowing the author to evaluate gait recognition performance under controlled, synthesized occlusions.

Tang et al. (2016) added occlusion patterns to the CMU MoBo dataset (*Collins, Gross & Shi, 2002*) by randomly overlaying simulated horizontal or vertical bars on the silhouette images with a probability ranging from 10% to 100%. The width of the horizontal bars varied from 40 to 100 pixels in increments of 20 pixels, while the width of the vertical bars varied from 20 to 50 pixels in increments of 10 pixels. The position of the added bars was uniformly distributed within the height or width of each silhouette image. This process simulated occlusion in the binary silhouette images to demonstrate how the proposed method performs on incomplete silhouettes with occlusion, adding occlusion variability to the dataset in a controlled manner for experimental purposes.

Roy et al. (2011) experiment both on TUM-IITKGP (*Hofmann, Sural & Rigoll, 2011*) and CMU Mobo gait dataset (*Collins, Gross & Shi, 2002*). They introduced occlusion pattern for later dataset. The authors considered sequences of two gait cycles long starting from midstance after left leg forward position. They modeled low, medium and severe occlusion by degrading frames according to normal distributions $\mathcal{N}(8,3)$, $\mathcal{N}(12,3)$ and $\mathcal{N}(17,3)$ respectively, generating 100 samples indicating the number of degraded silhouette frames for each distribution. The frames were degraded in three different positions of each sequence-start, middle and end-to observe the effect of occlusion in different parts of the gait cycle. Specifically, the start was from midstance to right leg forward, middle was from left leg forward to right leg forward (end of one cycle to beginning of next), and end was from left leg forward to midstance. The degraded frames were treated as if the full silhouette was occluded by a large static object, allowing the author to test models with different levels and positions of occlusion within the gait cycle sequences.

Chen et al. (2009) experiment on CMU Mobo gait dataset (*Collins, Gross & Shi, 2002*) with synthesized occlusion pattern. To simulate occlusion, the researchers added either horizontal or vertical bars to the gallery sequences/silhouettes. These bars were the same color as the background in order to occlude parts of the silhouettes. A bar was added to each gallery silhouette with a probability varying from 10% to 100%. This probability represented the percentage of frames in a given sequence that contained an occluding bar. The position of the added bar within the height or width of each silhouette was randomly distributed in a uniform manner. In this way, the researchers could systematically introduce occlusion at different levels by varying the bar addition probability from 10% up to 100% across the frames of each gallery sequence. The bars themselves were matched to the background color to synthetically create realistic occlusion effects.

REVIEW OF LITERATURE

Existing research on gait recognition shares a common objective: to identify individuals based on their unique walking style or pattern. This serves as an alternative solution for surveillance when conventional methods fail. For instance, face recognition is ineffective when subjects wear masks. To accomplish this objective, various methods have been proposed in the literature. Our primary focus is on vision-based gait recognition; methods utilizing other data sources such as wearable sensors are not addressed in this article.

In this section, we examine various contemporary methodologies that involve the utilization of corrupt body or temporal representation to extract gait features. Specifically, we will focus on both reconstruction-free and reconstruction-based approaches, directly employ the corrupt body or temporal representation for gait feature extraction and reconstruct before extraction respectively. In Table 4, we clearly demonstrated the key features of reviewed methods.

Reconstruction free

Reconstruction-free approaches for gait recognition are designed to bypass the conventional step of reconstructing occluded or corrupted input data, opting instead to

Table 4 Methodologies of gait recognition using corrupt representative media caused by occlusion. For occlusion type, 'ST' stands for static, 'DY' stands for dynamic, 'OT' stands for others. For dataset type, 'NO' stands for native-occluded, 'OS' stands for occlusion synthetic. For reconstruction method, 'RB' stands for reconstruction-based, 'RF' stands for reconstruction-free. For body and temporal representation type, 'SIL' stands for silhouette, 'SMPL' stands for skinned multi-person linear.

Reference	Category	Advantages	Disadvantages	Dataset usage
GaitMoE (*Huang et al., 2025*)	OT NO RF SIL	Enhances occluded gait recognition by leveraging dynamic contextual information and a mixture of experts, improving robustness across various occlusion scenarios.	The complexity of the mixture of experts model may introduce higher computational demands, potentially limiting its real-time application.	OccGait (Proposed), OccCASIA-B (*Peng, Cao & He, 2023*), Gait3D (*Zheng et al., 2022*), GREW (*Takemura et al., 2018*)
GMSN (*Wei et al., 2024b*)	ST OS RF SIL	Improves gait recognition accuracy under occlusion by utilizing multi-scale feature extraction and part-based horizontal mapping to focus on key body parts.	The method's complexity in handling multi-scale features and occlusions may result in increased computational costs	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), GREW (*Takemura et al., 2018*)
GaitASMS (*Sun et al., 2024*)	ST OS RF SIL	Improves gait recognition performance under occlusion by using adaptive spatial representation and multi-scale temporal aggregation	The use of complex temporal aggregation methods increases computational cost.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*)
GaitAE (*Li et al., 2024a*)	ST OS RB SIL	Enhances gait recognition accuracy by using an autoencoder-based framework to effectively reconstruct occluded gait sequences	The reliance on synthetic occlusion data for training may limit the generalizability of the model.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), SUSTech1K (*Shen et al., 2023*)
BGaitR-Net (*Kumar et al., 2024*)	ST OS RB SIL	Effectively reconstructs occluded gait sequences using a novel deep learning approach, improving recognition accuracy even under high occlusion levels.	The reliance on computationally intensive reconstruction methods may limit its real-time application.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), OU-ISIR LP (*Iwama et al., 2012*), TUM-IITKGP (*Hofmann, Sural & Rigoll, 2011*)
GaitLRDF (*Pan et al., 2024*)	OT NO RF SIL	Enhances gait recognition by using local relevant feature representation and discriminative feature learning, leading to improved accuracy in complex occlusion scenarios.	The method's reliance on detailed feature extraction may increase computational complexity.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*)
GLGait (*Peng et al., 2024*)	OT NO RF SIL	Enhances gait recognition in real-world scenarios by introducing a global-local temporal receptive field	The model's complexity, especially with its temporal attention mechanisms, increases computational costs, which may limit scalability.	Gait3D (*Zheng et al., 2022*), GREW (*Takemura et al., 2018*)
SSGait (*Xi et al., 2024*)	OT NO RB SIL	Improves gait recognition performance by utilizing semi-supervised and self-supervised learning with innovative strategies like tube-masked reconstruction and clip order prediction.	May introduce additional computational overhead, potentially limiting efficiency in large-scale applications.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*)

Table 4 (continued)

Reference	Category	Advantages	Disadvantages	Dataset usage
Gaitdlf (*Wei et al., 2024a*)	OT NO RF SIL	Enhances gait recognition by incorporating dynamic local and global feature fusion, improving robustness in complex real-world environments.	The method's complexity and higher computational cost may limit its practicality for real-time applications.	CASIA-B (*Yu, Tan & Tan, 2006*), Gait3D (*Zheng et al., 2022*), GREW (*Takemura et al., 2018*)
MCDGait (*Xiong et al., 2024*)	OT NO RF SIL	Enhances gait recognition by integrating multimodal data and spatial-temporal graph reasoning, significantly improving performance in real-world environments.	Increases computational overhead, potentially hindering real-time application.	Gait3D (*Zheng et al., 2022*), GREW (*Takemura et al., 2018*)
QAGait (*Wang et al., 2024*)	OT NO RB SIL	Improves gait recognition in complex, real-world environments by incorporating cost-effective quality assessment and feature learning strategies.	It requires additional computational steps for quality assessments, which may limit efficiency in real-time applications.	Gait3D (*Zheng et al., 2022*), GREW (*Takemura et al., 2018*), CASIA-B (*Yu, Tan & Tan, 2006*)
GaitSTR (*Zheng et al., 2024a*)	OT NO RB SIL	Offers flexibility across static, dynamic, and native-occluded datasets.	Relies on computationally intensive methods, may limit real-time application performance.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*), GREW (*Takemura et al., 2018*)
Gupta & Chellappa (2024)	ST OS RF SIL	Can collaborate with any backbone networks.	Synthetic occlusion may not accurately simulate real-world occlusion.	GREW (*Takemura et al., 2018*), BRIAR (*Cornett et al., 2023*)
HSTL (*Wang et al., 2023a*)	OT NO RF SIL	Enhances gait recognition by employing hierarchical spatio-temporal representation learning, improving robustness to occlusion and varied environmental conditions.	The complexity of the hierarchical model increases computational demands.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), GREW (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*)
DyGait (*Wang et al., 2023b*)	DY NO RF SIL	Improves gait recognition by exploiting dynamic representations, enhancing robustness and performance in scenarios with varying occlusions.	The reliance on dynamic feature extraction increases model complexity	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), GREW (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*)
GaitGCI (*Dou et al., 2023*)	OT NO RB SIL	Improves gait recognition by using generative counterfactual interventions to mitigate confounders	The reliance on singular value decomposition (SVD) for counterfactual learning introduces computational costs, which may limit efficiency in high-dimensional datasets and multi-dataset scenarios.	GREW (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*)
DANet (*Ma et al., 2023*)	DY NO RF SIL	Enhances gait recognition by dynamically aggregating local and global motion patterns.	The reliance on dynamic attention mechanisms and complex feature aggregation increases computational costs.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*)

(Continued)

Table 4 (continued)

Reference	Category	Advantages	Disadvantages	Dataset usage
GaitBase (*Fan et al., 2023*)	OT NO RF SIL	Improves the practicality and robustness of gait recognition in real-world scenarios by offering a flexible, extensible codebase and a strong baseline model.	The reliance on outdoor datasets for testing reveals that the model may struggle with performance consistency due to complex real-world conditions.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), GREW (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*)
Xu et al. (2023c)	ST OS RB SIL	More in line with practical application scenarios by directly with a bounding box that contains only the visible area of the contour.	Only for upper body shading and did not consider shading of other body parts.	OU-MVLP (*Takemura et al., 2018*)
Xu et al. (2023b)	ST OS RB SMPL	Extracting SMPL models directly from occluded gait videos without a full body bounding box	Difficult to well estimate a full human body model just from relatively small non-occluded body parts.	OU-MVLP (*Takemura et al., 2018*)
OccSilGait (*Peng, Cao & He, 2023*)	ST & DY OS RB SIL	A new framework, OccSilGait, is proposed to generate realistic occluded pedestrian sequences, considering a variety of occlusion scenarios	The proposed model contains multiple modules, cause high computational costs for training and inference.	OccCASIA-B (*Peng, Cao & He, 2023*)
MTSGait (*Zheng et al., 2022*)	OT NO RF SIL	Enhances gait recognition by introducing a multi-hop temporal switch mechanism.	The reliance on 2D convolution and non-cyclic sampling may still struggle with certain complex temporal dynamics, limiting the model's ability to fully capture long-term temporal dependencies.	GREW (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*)
SMPLGait (*Zheng et al., 2022*)	OT NO RF SMPL	Improves gait recognition by leveraging dense 3D SMPL models, providing more detailed representations of body shape and viewpoint	The method requires significant computational resources to handle 3D meshes, and its reliance on the Gait3D dataset may limit its applicability to environments with similar conditions.	Gait3D (*Zheng et al., 2022*)
Das, Agarwal & Chattopadhyay (2022)	ST & DY NO & OS RB GEI	It is first work on appearance-based gait recognition that employs DNNs to detect occlusion and predict missing frames.	It is not easy to generalize to real world due to complex occlusion scenario.	TUM-IITKGP (*Hofmann, Sural & Rigoll, 2011*), CASIA-B (*Yu, Tan & Tan, 2006*)
GaitGL (*Lin et al., 2022*)	ST & DY NO & OS RF SIL	By combining global and local features, made it possible to generate more discriminative gait feature representations.	High computational complexity, requires large of training set.	CASIA-B (*Yu, Tan & Tan, 2006*), OU-MVLP (*Takemura et al., 2018*), GREW (*Takemura et al., 2018*), Gait3D (*Zheng et al., 2022*)
BGaitR-Net (*Kumara, Chattopadhyaya & Wang, 2021*)	ST & DY NO & OS RB GEI	Outstanding performance when dealing with moderate occlusion.	Performance degrades when dealing with high levels occlusion; Requires a large training set.	OU-ISIR Treadmill Dataset B (*Makihara et al., 2012*), CASIA-B (*Yu, Tan & Tan, 2006*)

Table 4 (continued)

Reference	Category	Advantages	Disadvantages	Dataset usage
Rpnet (*Qin et al., 2021*)	ST & DY NO & OS RB GEI	A new module PFRE capable of discovering relationships between different body parts of pedestrians is proposed and the RPNet network is constructed base on PFRE.	The extent to which different body parts contribute to recognition is unclear, and feature aggregation methods are simpler.	CASIA-B (*Yu, Tan & Tan, 2006*) OU-ISLR LP (*Iwama et al., 2012*) OU-MVLP (*Takemura et al., 2018*)
GaitPart (*Fan et al., 2020*)	OT NO RF SIL	Enhances gait recognition by introducing a part-based temporal model that captures fine-grained spatial and temporal features.	The method's focus on short-range temporal features may limit its ability to model long-range dependencies, which could affect recognition performance in scenarios involving complex gait cycles.	CASIA-B (*Yu, Tan & Tan, 2006*) OU-MVLP (*Takemura et al., 2018*)
GaitSet (*Chao et al., 2019*)	OT NO RF SIL	This work enhances gait recognition by introducing a set-based approach that efficiently captures temporal and spatial information.	The approach may struggle with further scalability and efficiency due to the computational cost of processing large sets of silhouettes.	CASIA-B (*Yu, Tan & Tan, 2006*) OU-MVLP (*Takemura et al., 2018*)
Uddin et al. (2019)	ST & DY OS RB SIL	Efficient reconstruction of gait sequences without occlusion.	Requires a large amount of training data and the model needs to be retrained for different occlusion patterns.	OU-MVLP (*Takemura et al., 2018*)
Rgait-net (*Das et al., 2019*)	ST & DY & OT OS RB GEI	For the first time, a neural-based approach is proposed to detect occlusions and predict missing frames.	Accuracy reduce for sequences with high levels of occlusion.	CASIA-B (*Yu, Tan & Tan, 2006*), TUM-IITKGP (*Hofmann, Sural & Rigoll, 2011*)
Chen et al. (2017)	DY NO & OS RF GEI	Automatic detection of stable and invariant characteristics of gait in multiple walkers' gait.	Recognition rate still needs to be improved in the case of multiple walkers, especially in the case of multiple people with heavy occlusion.	CASIA-B (*Yu, Tan & Tan, 2006*), EEPIT Multi-gait (Proposed)
Ortells et al. (2017)	ST & DY & OT OS RF SIL	Can provide reliable gait characterization in severely occluded gait sequences.	Single weight function with universal parameters may not optimally model feature-specific criteria to separate genuine and outlier feature deviations.	OU-ISIR Treadmill Dataset B (*Makihara et al., 2012*), USF Human ID Gait Database (*Sarkar et al., 2005*)
Nangtin, Kumhom & Chamnongthai (2016)	ST OS RF GEI	Improved recognition robustness by dividing the human body into multiple modules and performing occlusion detection and exclusion.	Not set up to handle part occlusion and assume all body parts appear in all frames.	CASIA-B (*Yu, Tan & Tan, 2006*), EEPIT (Proposed)

directly extract features from the available, unoccluded portions of the subject's gait. By skipping the reconstruction process, these methods reduce computational overhead, making them advantageous for real-time applications where processing speed and low latency are critical. This capability is particularly valuable in scenarios with partial occlusions, where the visible gait patterns offer sufficient information to achieve accurate identification. These approaches are built on specialized algorithms capable of emphasizing unaffected portions of the data, making them highly suitable for dynamic environments, such as crowded surveillance areas or live-streamed footage.

One standout example of a reconstruction-free strategy is GaitMoE (*Huang et al., 2025*), which uses a mixture of experts (MoE) model to adaptively handle different occlusion scenarios. In this framework, distinct expert models specialize in recognizing specific occlusion patterns, and the system dynamically selects the most suitable expert for each input. This adaptive approach enables GaitMoE to perform robustly across a range of environments with varying occlusion levels. However, this method involves a trade-off: while it increases flexibility and robustness, the MoE architecture adds complexity and computational demands due to the dynamic expert selection mechanism. This challenge becomes especially apparent in applications requiring continuous, real-time processing, where the multi-expert setup may affect overall efficiency. GMSN (*Wei et al., 2024b*) adopts a different approach by focusing on static occlusions through multi-scale feature extraction and part-based horizontal mapping. In this method, the silhouette of the subject is divided into segments, with emphasis placed on unoccluded body regions to optimize feature extraction. By concentrating on visible parts of the body, GMSN effectively mitigates the impact of occlusions, enhancing its resilience in scenarios with consistent occlusion patterns. Nevertheless, the reliance on multi-scale features introduces additional computational costs, particularly in environments where diverse occlusion patterns necessitate repeated adjustments. The segmentation and multi-scale strategy allow GMSN to perform well in moderately occluded environments, though handling fine-tuned adjustments across large datasets may require substantial computational resources. Further expanding on reconstruction-free methods, GaitASMS (*Sun et al., 2024*) incorporates adaptive spatial representation with multi-scale temporal aggregation to improve occluded gait recognition. This approach enables the model to emphasize visible portions of the gait while maintaining adaptability to changing occlusion patterns. GaitASMS's temporal aggregation methods are particularly effective in dealing with partial occlusions across extended gait sequences; however, the complexity of aggregating temporal features at multiple scales can lead to increased processing time, especially when managing long gait sequences or high-resolution data. The GLGait method (*Peng et al., 2024*) offers an innovative take on reconstruction-free approaches by combining global and local temporal receptive fields. This fusion allows GLGait to capture both macro-level gait trends and micro-level body movements, thus providing robustness across varying degrees of occlusion. By integrating global and local cues, GLGait excels in recognizing gait patterns in real-world scenarios where occlusion levels fluctuate. Nonetheless, the complex temporal attention mechanisms utilized in this approach can result in higher

computational requirements, making it less ideal for applications with constrained resources or where large-scale data must be processed rapidly.

While these reconstruction-free approaches demonstrate effectiveness in moderate occlusion scenarios, their performance tends to decline as occlusion severity increases. In cases where substantial portions of the body are obscured, the ability of these methods to extract discriminative features diminishes due to the lack of visible gait information. This limitation highlights a core difference from reconstruction-based methods, which are designed to infer missing data and perform better in extreme occlusion settings. The primary advantage of reconstruction-free approaches lies in their efficiency, as they directly process the available data without the need for intermediate reconstruction. This streamlined operation is particularly suitable for environments where processing speed is essential, such as in surveillance systems deployed in crowded public spaces, transportation hubs, or live video feeds, where rapid response is prioritized over absolute precision. These methods, by focusing on extracting features from visible, unoccluded regions, allow for a significant reduction in computational load compared to reconstruction-based techniques. The use of models like GaitMoE (*Huang et al., 2025*), GMSN (*Wei et al., 2024b*), and GLGait (*Peng et al., 2024*) demonstrates how specialized, feature-focused extraction can provide robust performance with partial occlusions. However, the reliance on visible data limits the use of reconstruction-free methods in high-occlusion scenarios, where critical gait features may be obscured. The multi-expert setups in models like GaitMoE and the complex temporal mechanisms in GLGait, while enhancing flexibility, also contribute to computational overhead that may constrain their real-time capabilities in highly dynamic environments. As such, reconstruction-free methods are best suited for applications with low-to-moderate occlusions and environments prioritizing speed over exhaustive detail. For more severe occlusions, reconstruction-based approaches may still offer superior recognition accuracy by filling in missing gait information, thereby preserving discriminative features necessary for reliable identification.

Reconstruction based

Reconstruction-based gait recognition methods offer a robust solution for handling occlusion by first reconstructing missing or occluded portions of gait sequences before proceeding to feature extraction. These methods commonly leverage deep learning architectures, such as autoencoders (*Tschannen, Bachem & Lucic, 2018*; *Yang et al., 2022*; *Li, Pei & Li, 2023*) and generative adversarial networks (GANs) (*Goodfellow et al., 2020*; *Saxena & Cao, 2021*), to restore incomplete representations of gait data. By providing a more holistic view of the occluded data, reconstruction-based approaches enable the extraction of rich and informative features, significantly enhancing identification accuracy, even in challenging scenarios with substantial data loss. These methods are particularly valuable in high-security environments and critical applications where maximizing recognition accuracy is essential, albeit at the cost of higher computational demands.

One notable example is GaitAE (*Li et al., 2024a*), which utilizes an autoencoder model to map corrupted input gait sequences into a latent space and subsequently reconstructs

them into a more complete form. The reconstruction process in GaitAE allows for the recovery of essential gait features, making it highly effective in scenarios with extensive occlusions. However, GaitAE's dependence on synthetic occlusion data during training introduces potential challenges in generalizability, as real-world occlusions may differ from synthetic patterns, potentially limiting the model's performance in authentic, unpredictable occlusion conditions. Adaptations, such as transfer learning, could enhance GaitAE's applicability in real-world environments by enabling it to generalize better to diverse occlusion patterns encountered outside the training data. BGaitR-Net (*Kumar et al., 2024*), another influential reconstruction-based model, targets static occlusions by employing a deep learning framework to restore occluded gait segments before feature extraction. This model is particularly effective under high levels of occlusion, as it reconstructs missing portions of the gait data, allowing for robust recognition despite significant visual data loss. Nevertheless, BGaitR-Net's computationally intensive reconstruction process can be a limiting factor in real-time applications, especially in high-stakes settings where rapid decision-making is required. SSGait (*Xi et al., 2024*) introduces a more sophisticated approach by combining semi-supervised and self-supervised learning techniques to enhance recognition performance in dynamic occlusion scenarios. Utilizing methods such as tube-masked reconstruction and clip order prediction, SSGait effectively handles occlusions without the need for extensive labeled data, improving data efficiency. However, the complexity of these self-supervised techniques increases computational demands, making SSGait less feasible for real-time deployment in environments that require quick processing and response times. OccSilGait (*Peng, Cao & He, 2023*) provides an advanced reconstruction-based framework specifically designed to address both static and dynamic occlusions. By using multiple modules to generate realistic occluded gait sequences, OccSilGait achieves high recognition accuracy even in severe occlusion scenarios. The multi-module architecture, however, introduces substantial computational costs during both training and inference, limiting its practicality in resource-constrained or real-time applications. Additional models, such as GaitSTR (*Zheng et al., 2024a*) and QAGait (*Wang et al., 2024*), also adopt reconstruction-based approaches to tackle occlusion. GaitSTR is notable for its adaptability across various occlusion types, using complex reconstruction techniques to address a wide range of missing data scenarios, while QAGait incorporates a quality assessment stage to ensure that only high-quality reconstructed features are used in the identification process. This added quality assessment step enhances QAGait's performance in real-world environments where data fidelity is often compromised.

Although reconstruction-based methods offer an effective approach to managing severe occlusions, they are computationally intensive, which limits their applicability in real-time scenarios. The reconstruction process itself requires significant computational resources, and the added feature extraction step further contributes to the overall processing load. As such, these methods are generally more suited for high-security or high-stakes environments, such as forensic analysis or post-event security assessments, where recognition accuracy is prioritized over speed. With advancements in hardware acceleration and optimized neural network architectures, however, there is potential to

bring reconstruction-based methods closer to real-time applicability. Improved computational efficiency could expand the utility of these methods to broader domains, enabling more responsive gait recognition systems capable of managing high levels of occlusion in dynamic, real-world environments. For applications requiring immediate responses, though, the trade-off between computational complexity and accuracy makes reconstruction-free methods a more practical option, albeit at the potential cost of reduced effectiveness under heavy occlusions.

FUTURE DIRECTIONS

Bridging the lab-real world divide. Despite recent progress, most gait recognition datasets lack a comprehensive representation of occlusion patterns, especially those found in real-world settings. Datasets focusing on occlusions are often gathered in controlled environments, creating a significant gap between laboratory conditions and real-world applications. Addressing this disparity requires developing gait datasets captured in diverse, uncontrolled environments, as exemplified by the Gait Recognition in the Wild (GREW) dataset (*Zhu et al., 2021*). Future datasets should expand to include various occlusion types, such as static and dynamic obstacles, and span diverse lighting, backgrounds, and subject variations to capture the complexities encountered in practical applications. Additionally, benchmarks and evaluation metrics must evolve to account for these factors, enabling a more accurate assessment of model robustness and real-world viability. In high-impact areas like live surveillance, border security, and mobile device authentication, the practicality of gait recognition systems is often hindered by the computational intensity of state-of-the-art occlusion-resilient models. Although these models achieve robustness using complex architectures or extensive datasets, they typically demand high computational power and memory, which limits their utility on resource-constrained devices. A key future direction lies in developing lightweight models that can maintain high accuracy despite occlusions, without sacrificing speed or efficiency. Research efforts could focus on techniques such as model pruning, quantization, and efficient architecture designs to yield compact, high-performing models. Furthermore, advances in edge AI and hardware optimization would allow these models to run on embedded systems, facilitating real-time gait recognition even in environments with limited computational resources. Such developments could enable wide-scale deployment of occlusion-resilient gait recognition systems in diverse, real-time scenarios.

Balancing speed and accuracy through adaptive strategies. In evaluating reconstruction-free and reconstruction-based methods, several strategic considerations arise regarding their operational deployment and adaptability. Reconstruction-free methods, while invaluable for applications requiring fast, efficient processing, exhibit limitations under extensive occlusion. Conversely, reconstruction-based methods provide superior accuracy in challenging occlusion scenarios, albeit at the expense of higher computational demands. A promising direction for future research lies in the development of adaptive frameworks capable of dynamically selecting between these two approaches based on occlusion severity. For example, a hybrid gait recognition system could utilize reconstruction-free techniques in cases of minimal occlusion, switching to reconstruction-

Li et al. (2024), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.2602

39/55

based methods when occlusion levels exceed a predetermined threshold, thereby optimizing both speed and accuracy. Furthermore, enhancing model efficiency through innovations such as model pruning, quantization, or edge computing may make reconstruction-based methods feasible for real-time applications. Exploring such hybrid and adaptive approaches not only broadens the practical applicability of gait recognition technology across various settings but also aligns with the growing demand for efficient, high-performance models in real-world surveillance, security, and forensic applications.

Enhancing robustness through multi-modal integration and advanced data augmentation. Several methodologies in occluded gait recognition have shown promise but face limitations in specific scenarios, necessitating innovative approaches to enhance robustness. CNNs (*LeCun et al., 1998*), for example, perform well with silhouette-based occlusions but struggle with complex, dynamic occlusions, such as those found in crowded environments, due to challenges in capturing part-whole relationships. Similarly, while GCNs (*Zhang et al., 2022*) excel in modeling joint structures across multi-view datasets, they may underperform under irregular occlusion patterns or low visibility, such as in dim lighting. CapNets (*Sabour, Frosst & Hinton, 2017*) offer increased spatial awareness and robustness to viewpoint changes, though their high computational demands limit their suitability for real-time or edge-based applications. To address these constraints, future research can explore integrating multi-modal inputs like thermal or radar imaging to improve visibility under challenging conditions, as well as leveraging GANs (*Goodfellow et al., 2020*) to synthetically augment training datasets and enhance model resilience. These hybrid approaches—combining data augmentation with multi-modal strategies—could significantly improve model generalization across diverse occlusion scenarios and support the deployment of robust gait recognition systems in high-stakes applications, such as public surveillance, crowded spaces, and security settings.

Biomechanical foundations of gait analysis. Gait refers to human walking style, which results from the coordinated function of muscle groups (*Zajac, Neptune & Kautz, 2003, 2002*). Prior research has attempted to quantify gait through various parameters. However, current biomechanical parameters (*Stergiou, 2020*) alone may not fully capture the complexity of walking motions (*Roberts, Mongeon & Prince, 2017*; *Shull et al., 2014*; *Simon, 2004*). Further progress could involve exploring relationships between parameters and the specific muscle activations that produce gait. A more comprehensive biomechanical understanding linking gait analysis to muscle group dynamics may help develop more sophisticated gait recognition technologies. Moving forward, researchers could aim to better explain gait through associating measurable parameters with the underlying neuromuscular control mechanisms responsible for diverse walking patterns across individuals. Many challenges encountered in the field of occluded gait recognition can be overcome through thorough research into the fundamental causes of gait movement. Such insights may lead to enhanced models capable of interpreting subtle gait variations and handling occlusions by predicting missing biomechanical data points, thus enabling more robust recognition in complex scenarios.

Safety and privacy matters. In the development of methodologies aimed at enhancing gait recognition accuracy, ensuring the safety of these systems is a critical concern. For

instance, the use of large language models in gait recognition (*Ye et al., 2024*) might raise security challenges. As demonstrated by *Liu et al. (2024)*, *Zhao et al. (2024)*, *Liang et al. (2024a)*, *Li et al. (2024b)*, while these models significantly improve performance in multiple tasks, they are also vulnerable to various forms of attacks, including adversarial manipulation. One important research question is whether a malicious actor could exploit these vulnerabilities to manipulate prediction outcomes. By exploring both model-level and system-level attack techniques for gait recognition (*Hou et al., 2024*; *Ji et al., 2024*; *He et al., 2023*; *Kwon, 2023*; *Bukhari et al., 2022*; *Zhu et al., 2020*; *Jia et al., 2019*), defense technique should also be extensively research on *Wu, Hassan & Hu (2022)*. Moreover, as gait recognition technologies become more pervasive, researchers and developers must implement robust privacy-preserving techniques (*Li et al., 2024c*; *Zheng et al., 2024c*; *Xia, Munakata & Sugiura, 2023*; *Su, Li & Cao, 2023*; *Delgado-Santos et al., 2022*; *Parashar & Shekhawat, 2022*; *Xu et al., 2021*), such as data anonymization and encryption, to ensure that these systems do not infringe upon individuals' rights. Regulatory frameworks must also be updated to address the unique challenges posed by gait recognition, establishing clear guidelines for its ethical use in both public and private sectors. Addressing the above risks and considerations is essential to maintain the integrity and trustworthiness of gait recognition systems, especially in high-stakes environments where security and privacy are paramount. Additionally, privacy research must explore occlusion scenarios' effects on privacy protection, balancing transparency with individual privacy needs across various public and surveillance contexts.

CONCLUSION

In this survey, we address a key challenge hindering the industrial application of gait recognition: occlusion. Through an extensive literature review, we present a systematic taxonomy that classifies different types of occlusions encountered in real-world scenarios. This taxonomy serves as a foundation for evaluating existing datasets and methods designed to tackle occlusion in gait recognition, providing a structured lens through which to assess current research efforts. We identify essential datasets that should underpin future studies and review current methods that specifically address occlusion challenges, applying rigorous criteria to enhance the practical relevance of our analysis. Furthermore, we propose potential research directions to encourage progress in addressing occlusion within gait recognition systems. By highlighting these directions, we hope this survey will serve as a valuable resource for researchers and practitioners, advancing the development of gait recognition toward practical, real-world applications.

ADDITIONAL INFORMATION AND DECLARATIONS

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Tianhao Li conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the article, and approved the final draft.
- Weizhi Ma analyzed the data, prepared figures and/or tables, and approved the final draft.
- Yujia Zheng analyzed the data, prepared figures and/or tables, and approved the final draft.
- Xinchao Fan analyzed the data, prepared figures and/or tables, and approved the final draft.
- Guangcan Yang conceived and designed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Lijun Wang performed the experiments, authored or reviewed drafts of the article, and approved the final draft.
- Zhengping Li performed the computation work, authored or reviewed drafts of the article, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

This is a literature review.

REFERENCES

Adamczyk J, Malawski F. 2021. Comparison of manual and automated feature engineering for daily activity classification in mental disorder diagnosis. *Computing and Informatics* **40(4)**:850–879 DOI 10.31577/cai_2021_4_850.

Ali I, Sabir S, Ullah Z. 2019. Internet of things security, device authentication and access control: a review. ArXiv preprint DOI 10.48550/arXiv.1901.07309.

Alom MZ, Taha TM, Yakopcic C, Westberg S, Sidike P, Nasrin MS, Hasan M, Van Essen BC, Awwal AA, Asari VK. 2019. A state-of-the-art survey on deep learning theory and architectures. *Electronics* **8(3)**:292 DOI 10.3390/electronics8030292.

Alpaydin E. 2021. *Machine learning.* Massachusetts, United States: MIT press.

Atrushi D, Abdulazeez AM. 2024. Human gait recognition based on deep learning: a review. *Indonesian Journal of Computer Science* **13(1)**:345–358 DOI 10.33022/ijcs.v13i1.3719.

Azzeh J, Zahran B, Alqadi Z. 2018. Salt and pepper noise: effects and removal. *JOIV: International Journal on Informatics Visualization* **2(4)**:252–256.

Bajwa TK, Garg S, Saurabh K. 2016. Gait analysis for identification by using SVM with K-NN and NN techniques. In: *2016 Fourth International Conference on Parallel, Distributed and Grid Computing (PDGC).* Piscataway: IEEE, 259–263.

Li et al. (2024), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.2602

42/55

Balakrishnama S, Ganapathiraju A. 1998. Linear discriminant analysis-a brief tutorial. *Institute for Signal and Information Processing* **18(1998)**:1–8.

Bari AH, Gavrilova ML. 2022. Kinectgaitnet: kinect-based gait recognition using deep convolutional neural network. *Sensors* **22(7)**:2631 DOI 10.3390/s22072631.

Bengio Y, Simard P, Frasconi P. 1994. Learning long-term dependencies with gradient descent is difficult. *IEEE Transactions on Neural Networks* **5(2)**:157–166 DOI 10.1109/72.279181.

Bodepudi A, Reddy M. 2020. Spoofing attacks and mitigation strategies in biometrics-as-a-service systems. *Eigenpub Review of Science and Technology* **4(1)**:1–14.

Bolle RM, Connell JH, Pankanti S, Ratha NK, Senior AW. 2005. The relation between the ROC curve and the CMC. In: *Fourth IEEE Workshop on Automatic Identification Advanced Technologies (AutoID'05)*. Piscataway: IEEE, 15–20.

Bouchrika I. 2018. A survey of using biometrics for smart visual surveillance: gait recognition. In: *Surveillance in Action: Technologies for Civilian, Military and Cyber Surveillance*. Cham: Springer, 3–23.

Bukhari M, Durrani MY, Gillani S, Yasmin S, Rho S, Yeo S-S. 2022. Exploiting vulnerability of convolutional neural network-based gait recognition system. *The Journal of Supercomputing* **78(17)**:18578–18597 DOI 10.1007/s11227-022-04611-3.

Cai N, Feng S, Gui Q, Zhao L, Pan H, Yin J, Lin B. 2021. Hybrid silhouette-skeleton body representation for gait recognition. In: *2021 13th International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC)*. Piscataway: IEEE, 216–220.

Carreira-Perpinán MA. 1997. A review of dimension reduction techniques. Department of Computer Science. University of Sheffield Technical Report. Vol. 9, 1–69.

Chao H, He Y, Zhang J, Feng J. 2019. GaitSet: regarding gait as a set for cross-view gait recognition. *Proceedings of the AAAI Conference on Artificial Intelligence* **33**:8126–8133 DOI 10.1609/aaai.v33i01.33018126.

Chattopadhyay P, Sural S, Mukherjee J. 2015. Frontal gait recognition from occluded scenes. *Pattern Recognition Letters* **63(4)**:9–15 DOI 10.1016/j.patrec.2015.06.004.

Chen C, Liang J, Zhao H, Hu H, Tian J. 2009. Frame difference energy image for gait recognition with incomplete silhouettes. *Pattern Recognition Letters* **30(11)**:977–984 DOI 10.1016/j.patrec.2009.04.012.

Chen C, Liang J, Zhu X. 2011. Gait recognition based on improved dynamic Bayesian networks. *Pattern Recognition* **44(4)**:988–995 DOI 10.1016/j.patcog.2010.10.021.

Chen X, Luo X, Weng J, Luo W, Li H, Tian Q. 2021. Multi-view gait image generation for cross-view gait recognition. *IEEE Transactions on Image Processing* **30**:3041–3055 DOI 10.1109/TIP.2021.3055936.

Chen X, Weng J, Lu W, Xu J. 2017. Multi-gait recognition based on attribute discovery. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **40(7)**:1697–1710 DOI 10.1109/TPAMI.2017.2726061.

Cho K, Van Merriënboer B, Bahdanau D, Bengio Y. 2014. On the properties of neural machine translation: encoder-decoder approaches. ArXiv preprint DOI 10.48550/arXiv.1409.1259.

Choi S, Kim J, Kim W, Kim C. 2019. Skeleton-based gait recognition via robust frame-level matching. *IEEE Transactions on Information Forensics and Security* **14(10)**:2577–2592 DOI 10.1109/TIFS.2019.2901823.

Choi S, Youn IH, LeMay R, Burns S, Youn JH. 2014. Biometric gait recognition based on wireless acceleration sensor using k-nearest neighbor classification. In: *2014 International Conference on Computing, Networking and Communications (ICNC)*. Piscataway: IEEE, 1091–1095.

Coelho KK, Tristão ET, Nogueira M, Vieira AB, Nacif JA. 2023. Multimodal biometric authentication method by federated learning. *Biomedical Signal Processing and Control* **85(6)**:105022 DOI 10.1016/j.bspc.2023.105022.

Collins RT, Gross R, Shi J. 2002. Silhouette-based human identification from body shape and gait. In: *Proceedings of Fifth IEEE International Conference on Automatic Face Gesture Recognition*. Piscataway: IEEE, 366–371.

Connor P, Ross A. 2018. Biometric recognition by gait: a survey of modalities and features. *Computer Vision and Image Understanding* **167(4)**:1–27 DOI 10.1016/j.cviu.2018.01.007.

Cornett D, Brogan J, Barber N, Aykac D, Baird S, Burchfield N, Dukes C, Duncan A, Ferrell R, Goddard J, Jager G, Larson M, Murphy B, Johnson C, Shelley I, Srinivas N, Stockwell B, Thompson L, Yohe M, Zhang R, Dolvin S, Santos-Villalobos HJ, Bolme DS. 2023. Expanding accurate person recognition to new altitudes and ranges: the briar dataset. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 593–602.

Dargan S, Kumar M. 2020. A comprehensive survey on the biometric recognition systems based on physiological and behavioral modalities. *Expert Systems with Applications* **143(c)**:113114 DOI 10.1016/j.eswa.2019.113114.

Das D, Agarwal A, Chattopadhyay P. 2022. Gait recognition from occluded sequences in surveillance sites. In: *European Conference on Computer Vision*. Cham: Springer, 703–719.

Das D, Agarwal A, Chattopadhyay P, Wang L. 2019. Rgait-Net: an effective network for recovering missing information from occluded gait cycles. ArXiv preprint DOI 10.48550/arXiv.1912.06765.

Delgado-Escano R, Castro FM, Cózar R, Marin-Jimenez MJ, Guil N. 2020. Mupeg—the multiple person gait framework. *Sensors* **20(5)**:1358 DOI 10.3390/s20051358.

Delgado-Escano R, Castro FM, Cózar JR, Marín-Jiménez MJ, Guil N. 2018. An end-to-end multi-task and fusion CNN for inertial-based gait recognition. *IEEE Access* **7**:1897–1908 DOI 10.1109/ACCESS.2018.2886899.

Delgado-Santos P, Tolosana R, Guest R, Vera-Rodriguez R, Deravi F, Morales A. 2022. Gaitprivacyon: privacy-preserving mobile gait biometrics using unsupervised learning. *Pattern Recognition Letters* **161(4)**:30–37 DOI 10.1016/j.patrec.2022.07.015.

Derbel A, Vivet D, Emile B. 2015. Access control based on gait analysis and face recognition. *Electronics Letters* **51(10)**:751–752 DOI 10.1049/el.2015.0767.

Devarapalli HP, Dhanikonda VSS, Gunturi SB. 2020. Non-intrusive identification of load patterns in smart homes using percentage total harmonic distortion. *Energies* **13(18)**:4628 DOI 10.3390/en13184628.

Dou H, Zhang P, Su W, Yu Y, Lin Y, Li X. 2023. Gaitgci: generative counterfactual intervention for gait recognition. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 5578–5588.

Eberhart R, Kennedy J. 1995. Particle swarm optimization. In: *Proceedings of the IEEE International Conference on Neural Networks*. Vol. 4. Citeseer, 1942–1948.

El Naqa I, Murphy MJ. 2015. *What is machine learning?* Cham: Springer.

Ellis A, Ferryman J. 2010. Pets2010 and pets2009 evaluation of results using individual ground truthed single views. In: *2010 7th IEEE International Conference on Advanced Video and Signal Based Surveillance*. Piscataway: IEEE, 135–142.

Fan C, Liang J, Shen C, Hou S, Huang Y, Yu S. 2023. Opengait: revisiting gait recognition towards better practicality. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Piscataway: IEEE, 9707–9716.

Fan C, Peng Y, Cao C, Liu X, Hou S, Chi J, Huang Y, Li Q, He Z. 2020. Gaitpart: temporal part-based model for gait recognition. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition.* Piscataway: IEEE, 14225–14233.

Fernandes C, Fonseca L, Ferreira F, Gago M, Costa L, Sousa N, Ferreira C, Gama J, Erlhagen W, Bicho E. 2018. Artificial neural networks classification of patients with parkinsonism based on gait. In: *2018 IEEE International Conference on Bioinformatics and Biomedicine (BIBM).* Piscataway: IEEE, 2024–2030.

Fernandez-Lopez P, Liu-Jimenez J, Kiyokawa K, Wu Y, Sanchez-Reillo R. 2019. Recurrent neural network for inertial gait user recognition in smartphones. *Sensors* **19(18)**:4054 DOI 10.3390/s19184054.

Ghosh R. 2022. A faster R-CNN and recurrent neural network based approach of gait recognition with and without carried objects. *Expert Systems with Applications* **205(13)**:117730 DOI 10.1016/j.eswa.2022.117730.

Giorgi G, Saracino A, Martinelli F. 2021. Using recurrent neural networks for continuous authentication through gait analysis. *Pattern Recognition Letters* **147(12)**:157–163 DOI 10.1016/j.patrec.2021.03.010.

Goodfellow I. 2016. Book review: deep learning. *Healthcare Informatics Research* **22(4)**:351–354 DOI 10.4258/hir.2016.22.4.351.

Goodfellow I, Pouget-Abadie J, Mirza M, Xu B, Warde-Farley D, Ozair S, Courville A, Bengio Y. 2020. Generative adversarial networks. *Communications of the ACM* **63(11)**:139–144 DOI 10.1145/3422622.

Guan Y, Li CT. 2013. A robust speed-invariant gait recognition system for walker and runner identification. In: *2013 International Conference on Biometrics (ICB).* Piscataway: IEEE, 1–8.

Gupta A, Chellappa R. 2024. You can run but not hide: improving gait recognition with intrinsic occlusion type awareness. In: *Proceedings of the IEEE/CVF Winter Conference on Applications of Computer Vision*, 5893–5902.

Gupta A, Prasad PC, Alsadoon A, Bajaj K. 2015. Hybrid method for gait recognition using SVM and Baysian network. In: *2015 IEEE 8th International Workshop on Computational Intelligence and Applications (IWCIA).* Piscataway: IEEE, 89–94.

Güner Şahan P, Şahin S, Kaya Gülağız F. 2024. A survey of appearance-based approaches for human gait recognition: techniques, challenges, and future directions. *The Journal of Supercomputing* **80(13)**:1–38 DOI 10.1007/s11227-024-06172-z.

Hasan MAM, Al Abir F, Al Siam M, Shin J. 2022. Gait recognition with wearable sensors using modified residual block-based lightweight CNN. *IEEE Access* **10**:42577–42588 DOI 10.1109/ACCESS.2022.3168019.

Haykin S. 2009. *Neural networks and learning machines, 3/E.* Bengaluru, Karnataka: Pearson Education India.

He J, Gao F, Wang J, Wu Q, Zhang Q, Lin W. 2022. A method combining multi-feature fusion and optimized deep belief network for EMG-based human gait classification. *Mathematics* **10(22)**:4387 DOI 10.3390/math10224387.

He Z, Wang W, Dong J, Tan T. 2023. Temporal sparse adversarial attack on sequence-based gait recognition. *Pattern Recognition* **133**:109028 DOI 10.1016/j.patcog.2022.109028.

Hinton GE. 2002. Training products of experts by minimizing contrastive divergence. *Neural Computation* **14(8)**:1771–1800 DOI 10.1162/089976602760128018.

Hinton GE, Osindero S, Teh YW. 2006. A fast learning algorithm for deep belief nets. *Neural Computation* **18(7)**:1527–1554 DOI 10.1162/neco.2006.18.7.1527.

Hochreiter S, Schmidhuber J. 1997. Long short-term memory. *Neural Computation* **9(8)**:1735–1780 DOI 10.1162/neco.1997.9.8.1735.

Hofmann M, Geiger J, Bachmann S, Schuller B, Rigoll G. 2014. The tum gait from audio, image and depth (GAID) database: multimodal recognition of subjects and traits. *Journal of Visual Communication and Image Representation* **25(1)**:195–206 DOI 10.1016/j.jvcir.2013.02.006.

Hofmann M, Sural S, Rigoll G. 2011. Gait recognition in the presence of occlusion: a new dataset and baseline algorithms. *Available at https://mediatum.ub.tum.de/doc/1107259/659988.pdf.*

Hou S, Wang Z, Zhang M, Cao C, Liu X, Huang Y. 2024. Edge-oriented adversarial attack for deep gait recognition. *International Journal of Computer Vision* **2(4)**:1–15 DOI 10.1007/s11263-024-02225-1.

Huang P, Peng Y, Hou S, Cao C, Liu X, He Z, Huang Y. 2025. Occluded gait recognition with mixture of experts: an action detection perspective. In: *European Conference on Computer Vision.* Cham: Springer, 380–397.

Huang H, Zhou P, Li Y, Sun F. 2021. A lightweight attention-based CNN model for efficient gait recognition with wearable imu sensors. *Sensors* **21(8)**:2866 DOI 10.3390/s21082866.

Hug F, Vogel C, Tucker K, Dorel S, Deschamps T, Le Carpentier É, Lacourpaille L. 2019. Individuals have unique muscle activation signatures as revealed during gait and pedaling. *Journal of Applied Physiology* **127(4)**:1165–1174 DOI 10.1152/japplphysiol.01101.2018.

Iwama H, Okumura M, Makihara Y, Yagi Y. 2012. The OU-ISIR gait database comprising the large population dataset and performance evaluation of gait recognition. *IEEE Transactions on Information Forensics and Security* **7(5)**:1511–1521 DOI 10.1109/TIFS.2012.2204253.

Ji B, Chen X, Yang W, Zhu F. 2024. Boosting robustness of silhouette-based gait recognition against adversarial attacks. In: *International Conference on Intelligent Computing.* Cham: Springer, 72–84.

Jia M, Yang H, Huang D, Wang Y. 2019. Attacking gait recognition systems via silhouette guided GANS. In: *Proceedings of the 27th ACM International Conference on Multimedia,* 638–646.

Jordan MI, Mitchell TM. 2015. Machine learning: trends, perspectives, and prospects. *Science* **349(6245)**:255–260 DOI 10.1126/science.aaa8415.

Jung J, Garnett E, Vispo B, Chen X, Cao J, Elghetany MT, Devaraj S. 2020. Misidentification of unstable, low oxygen affinity hemoglobin variant. *Clinica Chimica Acta* **509(9)**:177–179 DOI 10.1016/j.cca.2020.06.023.

Kamruzzaman J, Begg RK. 2006. Support vector machines and other pattern recognition approaches to the diagnosis of cerebral palsy gait. *IEEE Transactions on Biomedical Engineering* **53(12)**:2479–2490 DOI 10.1109/TBME.2006.883697.

Kececi A, Yildirak A, Ozyazici K, Ayluctarhan G, Agbulut O, Zincir I. 2020. Implementation of machine learning algorithms for gait recognition. *Engineering Science and Technology, an International Journal* **23(4)**:931–937 DOI 10.1016/j.jestch.2020.01.005.

Khaliluzzaman M, Uddin A, Deb K, Hasan MJ. 2023. Person recognition based on deep gait: a survey. *Sensors* **23(10)**:4875 DOI 10.3390/s23104875.

Khan MH, Farid MS, Grzegorzek M. 2021. Vision-based approaches towards person identification using gait. *Computer Science Review* **42(2)**:100432 DOI 10.1016/j.cosrev.2021.100432.

Khan MH, Farid MS, Grzegorzek M. 2023. A comprehensive study on codebook-based feature fusion for gait recognition. *Information Fusion* **92(8)**:216–230 DOI 10.1016/j.inffus.2022.12.001.

Khan A, Galarraga O, Garcia-Salicetti S, Vigneron V. 2024. Deep learning for quantified gait analysis: a systematic literature review. *IEEE Access* **12**:138932–138957 DOI 10.1109/ACCESS.2024.3434513.

Kharb A, Saini V, Jain Y, Dhiman S. 2011. A review of gait cycle and its parameters. *IJCEM International Journal of Computational Engineering & Management* **13**:78–83.

Kherif F, Latypova A. 2020. Principal component analysis. In: *Machine Learning*. Amsterdam, Netherlands: Elsevier, 209–225.

Krüger J, Berger T, Leich T. 2019. Features and how to find them: a survey of manual feature location. In: *Software Engineering for Variability Intensive Systems*. United States: Auerbach Publications, 153–172.

Kumar SS, Singh B, Chattopadhyay P, Halder A, Wang L. 2024. BGaitR-Net: an effective neural model for occlusion reconstruction in gait sequences by exploiting the key pose information. *Expert Systems with Applications* **246(1)**:123181 DOI 10.1016/j.eswa.2024.123181.

Kumara SS, Chattopadhyaya P, Wang L. 2021. BGaitR-Net: occluded gait sequence reconstructionwith temporally constrained model for gait recognition. ArXiv preprint DOI 10.48550/arXiv.2110.09564.

Kumari B, Bharti J. 2023. A comprehensive review of the occlusion identification based gait recognition mechanism. In: *2023 IEEE International Students' Conference on Electrical, Electronics and Computer Science (SCEECS)*. Piscataway: IEEE, 1–6.

Kwon H. 2023. Toward selective adversarial attack for gait recognition systems based on deep neural network. *IEICE Transactions on Information and Systems* **106(2)**:262–266 DOI 10.1587/transinf.2021EDL8080.

Lai Z, Xu Y, Jin Z, Zhang D. 2014. Human gait recognition via sparse discriminant projection learning. *IEEE Transactions on Circuits and Systems for Video Technology* **24(10)**:1651–1662 DOI 10.1109/TCSVT.2014.2305495.

Lam TH, Cheung KH, Liu JN. 2011. Gait flow image: a silhouette-based gait representation for human identification. *Pattern Recognition* **44(4)**:973–987 DOI 10.1016/j.patcog.2010.10.011.

LeCun Y, Bengio Y, Hinton G. 2015. Deep learning. *Nature* **521(7553)**:436–444 DOI 10.1038/nature14539.

LeCun Y, Bottou L, Bengio Y, Haffner P. 1998. Gradient-based learning applied to document recognition. *Proceedings of the IEEE* **86(11)**:2278–2324 DOI 10.1109/5.726791.

Li W, Hou S, Zhang C, Cao C, Liu X, Huang Y, Zhao Y. 2023b. An in-depth exploration of person re-identification and gait recognition in cloth-changing conditions. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13824–13833.

Li R, Li H, Qiu Y, Ren J, Ng WW, Zhao H. 2024a. Gaitae: a cognitive model-based autoencoding technique for gait recognition. *Mathematics* **12(17)**:2780 DOI 10.3390/math12172780.

Li T, Lu J, Chu C, Zeng T, Zheng Y, Li M, Huang H, Wu B, Liu Z, Ma K, Yuan X, Wang X, Ding K, Chen H, Zhang Q. 2024b. Scisafeeval: a comprehensive benchmark for safety alignment of large language models in scientific tasks. ArXiv preprint DOI 10.48550/arXiv.2410.03769.

Li X, Makihara Y, Xu C, Yagi Y. 2021. End-to-end model-based gait recognition using synchronized multi-view pose constraint. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 4106–4115.

Li X, Makihara Y, Xu C, Yagi Y, Ren M. 2020a. Gait recognition via semi-supervised disentangled representation learning to identity and covariate features. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 13309–13319.

Li X, Makihara Y, Xu C, Yagi Y, Yu S, Ren M. 2020b. End-to-end model-based gait recognition. In: *Proceedings of the Asian Conference on Computer Vision*.

Li P, Pei Y, Li J. 2023. A comprehensive survey on design and application of autoencoder in deep learning. *Applied Soft Computing* **138**:110176 DOI 10.1016/j.asoc.2023.110176.

Li N, Zhao X. 2023. A multi-modal dataset for gait recognition under occlusion. *Applied Intelligence* **53(2)**:1517–1534 DOI 10.1007/s10489-022-03474-8.

Li T, Zheng Y, Ma W, Wang G, Li Z, Wang L. 2024c. Trustworthy metaverse: a comprehensive investigation into security risks and privacy issues in artificial intelligence-extended reality systems. *SID Symposium Digest of Technical Papers* **55**:872–877 DOI 10.1002/sdtp.17226.

Li T, Zheng Y, Zhang H, Ma W, Li Y. 2023a. Research on real-time early warning mechanism of aviation safety based on finite state machine underlying in qar stream data. In: *Proceedings of the 2023 7th International Conference on Big Data and Internet of Things*, 216–222.

Liang J, Fan C, Hou S, Shen C, Huang Y, Yu S. 2022. Gaitedge: beyond plain end-to-end gait recognition for better practicality. In: *European Conference on Computer Vision*. Cham: Springer, 375–390.

Liang Y, Li H, Wu W, Xu P. 2024b. MaP-SGAN: multi-anchor point siamese GAN for Wi-Fi CSI-based cross-domain gait recognition. *Expert Systems with Applications* **251(3)**:124083 DOI 10.1016/j.eswa.2024.124083.

Liang S, Liang J, Pang T, Du C, Liu A, Chang EC, Cao X. 2024a. Revisiting backdoor attacks against large vision-language models. ArXiv preprint DOI 10.48550/arXiv.2406.18844.

Liang Y, Samtani S, Guo B, Yu Z. 2020. Behavioral biometrics for continuous authentication in the internet-of-things era: an artificial intelligence perspective. *IEEE Internet of Things Journal* **7(9)**:9128–9143 DOI 10.1109/JIOT.2020.3004077.

Lin TL, Tseng HW, Wen Y, Lai FW, Lin CH, Wang CJ. 2018. Reconstruction algorithm for lost frame of multiview videos in wireless multimedia sensor network based on deep learning multilayer perceptron regression. *IEEE Sensors Journal* **18(23)**:9792–9801 DOI 10.1109/JSEN.2018.2865916.

Lin B, Zhang S, Wang M, Li L, Yu X. 2022. Gaitgl: learning discriminative global-local feature representations for gait recognition. ArXiv preprint DOI 10.48550/arXiv.2208.01380.

Lin B, Zhang S, Yu X. 2021. Gait recognition via effective global-local feature representation and local temporal aggregation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 14648–14656.

Liu D, Yang M, Qu X, Zhou P, Hu W, Cheng Y. 2024. A survey of attacks on large vision-language models: resources, advances, and future trends. ArXiv preprint DOI 10.48550/arXiv.2407.07403.

Luo R, Sun S, Zhang X, Tang Z, Wang W. 2019. A low-cost end-to-end sEMG-based gait sub-phase recognition system. *IEEE Transactions on Neural Systems and Rehabilitation Engineering* **28(1)**:267–276 DOI 10.1109/TNSRE.2019.2950096.

Ma K, Fu Y, Zheng D, Cao C, Hu X, Huang Y. 2023. Dynamic aggregated network for gait recognition. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 22076–22085.

Ma W, Li Z, Li T, Zheng Y, Wang L. 2024a. Application of virtual reality technology in the diagnosis and treatment of psychological disorders: an electroencephalography (EEG)-based approach. *SID Symposium Digest of Technical Papers* **55**:882–886 DOI 10.1002/sdtp.17228.

Ma W, Zheng Y, Li T, Li Z, Li Y, Wang L. 2024b. A comprehensive review of deep learning in eeg-based emotion recognition: classifications, trends, and practical implications. *PeerJ Computer Science* **10(1)**:e2065 DOI 10.7717/peerj-cs.2065.

Ma Y, Zhu L. 2013. A review on dimension reduction. *International Statistical Review* **81(1)**:134–150 DOI 10.1111/j.1751-5823.2012.00182.x.

Maharjan P, Shrestha K, Bhatta T, Cho H, Park C, Salauddin M, Rahman MT, Rana SS, Lee S, Park JY. 2021. Keystroke dynamics based hybrid nanogenerators for biometric authentication and identification using artificial intelligence. *Advanced Science* **8(15)**:2100711 DOI 10.1002/advs.202100711.

Mahesh B. 2020. Machine learning algorithms-a review. *International Journal of Science and Research (IJSR)* **9(1)**:381–386 DOI 10.21275/ART20203995.

Makihara Y, Mannami H, Tsuji A, Hossain MA, Sugiura K, Mori A, Yagi Y. 2012. The OU-ISIR gait database comprising the treadmill dataset. *IPSJ Transactions on Computer Vision and Applications* **4**:53–62 DOI 10.2197/ipsjtcva.4.53.

Makihara Y, Sagawa R, Mukaigawa Y, Echigo T, Yagi Y. 2006. Gait recognition using a view transformation model in the frequency domain. In: *Computer Vision–ECCV 2006: 9th European Conference on Computer Vision, May 7–13, 2006, Proceedings, Part III 9*. Graz, Austria: Springer, 151–163.

Makihara Y, Suzuki A, Muramatsu D, Li X, Yagi Y. 2017. Joint intensity and spatial metric learning for robust gait recognition. In: *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, 5705–5715.

Mao M, Song Y. 2020. Gait recognition based on 3D skeleton data and graph convolutional network. In: *2020 IEEE International Joint Conference on Biometrics (IJCB)*. Piscataway: IEEE, 1–8.

Marra AC, Di Silvestro G, Somma R. 2023. Palaeontology applied to criminal investigation. *Atti della Accademia Peloritana dei Pericolanti-Classe di Scienze Fisiche, Matematiche e Naturali* **101(S1)**:4 DOI 10.1478/AAPP.101S1A4.

Marsico MD, Mecca A. 2019. A survey on gait recognition via wearable sensors. *ACM Computing Surveys (CSUR)* **52(4)**:1–39 DOI 10.1145/3340293.

Mehmood A, Khan MA, Sharif M, Khan SA, Shaheen M, Saba T, Riaz N, Ashraf I. 2020. Prosperous human gait recognition: an end-to-end system based on pre-trained CNN features selection. *Multimedia Tools and Applications* **83(5)**:1–21 DOI 10.1007/s11042-020-08928-0.

Muramatsu D, Makihara Y, Yagi Y. 2015. Gait regeneration for recognition. In: *2015 International Conference on Biometrics (ICB)*. Piscataway: IEEE, 169–176.

Murray MP, Drought AB, Kory RC. 1964. Walking patterns of normal men. *JBJS* **46(2)**:335–360 DOI 10.2106/00004623-196446020-00009.

Nambiar A, Bernardino A, Nascimento JC. 2019. Gait-based person re-identification: a survey. *ACM Computing Surveys (CSUR)* **52(2)**:1–34.

Nangtin P, Kumhom P, Chamnongthai K. 2016. Gait identification with partial occlusion using six modules and consideration of occluded module exclusion. *Journal of Visual Communication and Image Representation* **36(2)**:107–121 DOI 10.1016/j.jvcir.2016.01.008.

Narayan V, Awasthi S, Fatima N, Faiz M, Srivastava S. 2023. Deep learning approaches for human gait recognition: a review. In: *2023 International Conference on Artificial Intelligence and Smart Communication (AISC)*. Piscataway: IEEE, 763–768.

Ngo T-W, Hung EP. 2019. The political economy of border checkpoints in shadow exchanges. DOI 10.1080/00472336.2018.1555273.

Nickel C, Busch C, Rangarajan S, Möbius M. 2011. Using hidden markov models for accelerometer-based biometric gait recognition. In: *2011 IEEE 7th International Colloquium on Signal Processing and its Applications*. Piscataway: IEEE, 58–63.

Nordin MJ, Saadoon A. 2016. A survey of gait recognition based on skeleton model for human identification. *Research Journal of Applied Sciences, Engineering and Technology* **12(7)**:756–763 DOI 10.19026/rjaset.12.2751.

Ortells J, Mollineda RA, Mederos B, Martín-Félez R. 2017. Gait recognition from corrupted silhouettes: a robust statistical approach. *Machine Vision and Applications* **28(1–2)**:15–33 DOI 10.1007/s00138-016-0798-y.

Pan X, Xie H, Zhang N, Li S. 2024. GaitLRDF: gait recognition via local relevant feature representation and discriminative feature learning. *Applied Intelligence* **54(23)**:1–16 DOI 10.1007/s10489-024-05837-9.

Panahandeh G, Mohammadiha N, Leijon A, Händel P. 2013. Continuous hidden markov model for pedestrian activity classification and gait analysis. *IEEE Transactions on Instrumentation and Measurement* **62(5)**:1073–1083 DOI 10.1109/TIM.2012.2236792.

Parashar A, Parashar A, Shabaz M, Gupta D, Sahu AK, Khan MA. 2024. Advancements in artificial intelligence for biometrics: a deep dive into model-based gait recognition techniques. *Engineering Applications of Artificial Intelligence* **130(21)**:107712 DOI 10.1016/j.engappai.2023.107712.

Parashar A, Shekhawat RS. 2022. Protection of gait data set for preserving its privacy in deep learning pipeline. *IET Biometrics* **11(6)**:557–569 DOI 10.1049/bme2.12093.

Pataky TC, Mu T, Bosch K, Rosenbaum D, Goulermas JY. 2012. Gait recognition: highly unique dynamic plantar pressure patterns among 104 individuals. *Journal of the Royal Society Interface* **9(69)**:790–800 DOI 10.1098/rsif.2011.0430.

Paul A, Jain MM, Jain J, Chattopadhyay P. 2022. Gait cycle reconstruction and human identification from occluded sequences. ArXiv preprint DOI 10.48550/arXiv.2206.13395.

Peng Y, Cao C, He Z. 2023. Occluded gait recognition. In: *2023 International Joint Conference on Neural Networks (IJCNN)*. Piscataway: IEEE, 1–8.

Peng G, Wang Y, Zhao Y, Zhang S, Li A. 2024. Glgait: a global-local temporal receptive field network for gait recognition in the wild. ArXiv preprint DOI 10.48550/arXiv.2408.06834.

Pinaya WHL, Vieira S, Garcia-Dias R, Mechelli A. 2020. Convolutional neural networks. In: *Machine learning*. Amsterdam, Netherlands: Elsevier, 173–191.

Prakash C, Kumar R, Mittal N. 2018. Recent developments in human gait research: parameters, approaches, applications, machine learning techniques, datasets and challenges. *Artificial Intelligence Review* **49(1)**:1–40 DOI 10.1007/s10462-016-9514-6.

Qin H, Chen Z, Guo Q, Wu QJ, Lu M. 2021. RPNet: gait recognition with relationships between each body-parts. *IEEE Transactions on Circuits and Systems for Video Technology* **32(5)**:2990–3000 DOI 10.1109/TCSVT.2021.3095290.

Rani V, Kumar M. 2023. Human gait recognition: a systematic review. *Multimedia Tools and Applications* **82(24)**:1–35 DOI 10.1007/s11042-023-15079-5.

Rashmi M, Guddeti RMR. 2022. Human identification system using 3D skeleton-based gait features and lstm model. *Journal of Visual Communication and Image Representation* **82(10)**:103416 DOI 10.1016/j.jvcir.2021.103416.

Ren S, He K, Girshick R, Sun J. 2016. Faster R-CNN: towards real-time object detection with region proposal networks. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39(6)**:1137–1149 DOI 10.1109/TPAMI.2016.2577031.

Rida I, Almaadeed N, Almaadeed S. 2019. Robust gait recognition: a comprehensive survey. *IET Biometrics* **8(1)**:14–28 DOI 10.1049/iet-bmt.2018.5063.

Roberts M, Mongeon D, Prince F. 2017. Biomechanical parameters for gait analysis: a systematic review of healthy human gait. *Physical Therapy and Rehabilitation* **4(6)**:6.

Roy A, Chattopadhyay P, Sural S, Mukherjee J, Rigoll G. 2015. Modelling, synthesis and characterisation of occlusion in videos. *IET Computer Vision* **9(6)**:821–830 DOI 10.1049/iet-cvi.2014.0170.

Roy A, Sural S, Mukherjee J, Rigoll G. 2011. Occlusion detection and gait silhouette reconstruction from degraded scenes. *Signal, Image and Video Processing* **5(4)**:415–430 DOI 10.1007/s11760-011-0245-5.

Ryu R, Yeom S, Kim SH, Herbert D. 2021. Continuous multimodal biometric authentication schemes: a systematic review. *IEEE Access* **9**:34541–34557 DOI 10.1109/ACCESS.2021.3061589.

Sabour S, Frosst N, Hinton GE. 2017. Dynamic routing between capsules. *Advances in Neural Information Processing Systems* **30**:3859–3869 DOI 10.5555/3294996.3295142.

Santos CFGd, Oliveira DDS, Passos LA, Pires RG, Santos DFS, Valem LP, Moreira TP, Santana MCS, Roder M, Papa JP, Colombo D. 2022. Gait recognition based on deep learning: a survey. ArXiv preprint DOI 10.48550/arXiv.2201.03323.

Sarkar S, Phillips PJ, Liu Z, Vega IR, Grother P, Bowyer KW. 2005. The humanid gait challenge problem: data sets, performance, and analysis. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **27(2)**:162–177 DOI 10.1109/TPAMI.2005.39.

Saxena D, Cao J. 2021. Generative adversarial networks (GANS) challenges, solutions, and future directions. *ACM Computing Surveys (CSUR)* **54(3)**:1–42 DOI 10.1145/3446374.

Sepas-Moghaddam A, Etemad A. 2022. Deep gait recognition: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **45(1)**:264–284 DOI 10.1109/TPAMI.2022.3151865.

Sepas-Moghaddam A, Ghorbani S, Troje NF, Etemad A. 2021. Gait recognition using multi-scale partial representation transformation with capsules. In: *2020 25th International Conference on Pattern Recognition (ICPR)*. Piscataway: IEEE, 8045–8052.

Shen C, Fan C, Wu W, Wang R, Huang GQ, Yu S. 2023. Lidargait: benchmarking 3D gait recognition with point clouds. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1054–1063.

Sheng W, Li X. 2021. Multi-task learning for gait-based identity recognition and emotion recognition using attention enhanced temporal graph convolutional network. *Pattern Recognition* **114(1)**:107868 DOI 10.1016/j.patcog.2021.107868.

Shi LF, Qiu CX, Xin DJ, Liu GX. 2020. Gait recognition via random forests based on wearable inertial measurement unit. *Journal of Ambient Intelligence and Humanized Computing* **11(11)**:5329–5340 DOI 10.1007/s12652-020-01870-x.

Shiraga K, Makihara Y, Muramatsu D, Echigo T, Yagi Y. 2016. Geinet: view-invariant gait recognition using a convolutional neural network. In: *2016 International Conference on Biometrics (ICB)*. Piscataway: IEEE, 1–8.

Shull PB, Jirattigalachote W, Hunt MA, Cutkosky MR, Delp SL. 2014. Quantified self and human movement: a review on the clinical impact of wearable sensing and feedback for gait analysis and intervention. *Gait & Posture* **40(1)**:11–19 DOI 10.1016/j.gaitpost.2014.03.189.

Simon SR. 2004. Quantification of human motion: gait analysis—benefits and limitations to its application to clinical problems. *Journal of Biomechanics* **37(12)**:1869–1880 DOI 10.1016/j.jbiomech.2004.02.047.

Singh JP, Arora S, Jain S, SoM UPS. 2019. A multi-gait dataset for human recognition under occlusion scenario. In: *2019 International Conference on Issues and Challenges in Intelligent Computing Techniques (ICICT)*. Vol. 1, Piscataway: IEEE, 1–6.

Singh JP, Jain S, Arora S, Singh UP. 2018. Vision-based gait recognition: a survey. *IEEE Access* **6**:70497–70527 DOI 10.1109/ACCESS.2018.2879896.

Singh JP, Jain S, Arora S, Singh UP. 2020. Reconstruction of occluded ROI in multi-person gait based on numerical methods. *Multimedia Systems* **26(3)**:249–266 DOI 10.1007/s00530-019-00641-9.

Singh JP, Jain S, Arora S, Singh UP. 2021. A survey of behavioral biometric gait recognition: current success and future perspectives. *Archives of Computational Methods in Engineering* **28(1)**:107–148 DOI 10.1007/s11831-019-09375-3.

Sivaranjani P, Sasikala S, Ragul R, Riyazur Razak N, Saran V. 2023. Design and development of smart security key for knowledge based authentication. In: *2023 8th International Conference on Communication and Electronics Systems (ICCES)*. Piscataway: IEEE, 558–563.

Song C, Huang Y, Huang Y, Jia N, Wang L. 2019. Gaitnet: an end-to-end network for gait based human identification. *Pattern Recognition* **96(12)**:106988 DOI 10.1016/j.patcog.2019.106988.

Souchet AD, Philippe S, Lourdeaux D, Leroy L. 2022. Measuring visual fatigue and cognitive load via eye tracking while learning with virtual reality head-mounted displays: a review. *International Journal of Human-Computer Interaction* **38(9)**:801–824 DOI 10.1080/10447318.2021.1976509.

Stergiou N. 2020. *Biomechanics and gait analysis.* Cambridge: Academic Press.

Su P. 2021. Immersive online biometric authentication algorithm for online guiding based on face recognition and cloud-based mobile edge computing. *Distributed and Parallel Databases* **11(8)**:1–22 DOI 10.1007/s10619-021-07351-0.

Su Y, Li Y, Cao Z. 2023. Gait-based privacy protection for smart wearable devices. *IEEE Internet of Things Journal* **11(2)**:3497–3509 DOI 10.1109/JIOT.2023.3296650.

Sun Y, Long H, Feng X, Nixon M. 2024. Gaitasms: gait recognition by adaptive structured spatial representation and multi-scale temporal aggregation. *Neural Computing and Applications* **36(13)**:7057–7069 DOI 10.1007/s00521-024-09445-z.

Takemura N, Makihara Y, Muramatsu D, Echigo T, Yagi Y. 2018. Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. *IPSJ Transactions on Computer Vision and Applications* **10(1)**:1–14 DOI 10.1186/s41074-018-0039-6.

Tang J, Luo J, Tjahjadi T, Guo F. 2016. Robust arbitrary-view gait recognition based on 3D partial similarity matching. *IEEE Transactions on Image Processing* **26(1)**:7–22 DOI 10.1109/TIP.2016.2612823.

Teepe T, Gilg J, Herzog F, Hörmann S, Rigoll G. 2022. Towards a deeper understanding of skeleton-based gait recognition. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1569–1577.

Tian H, Ma X, Wu H, Li Y. 2022. Skeleton-based abnormal gait recognition with spatio-temporal attention enhanced gait-structural graph convolutional networks. *Neurocomputing* **473(6)**:116–126 DOI 10.1016/j.neucom.2021.12.004.

Tschannen M, Bachem O, Lucic M. 2018. Recent advances in autoencoder-based representation learning. ArXiv preprint DOI 10.48550/arXiv.1812.05069.

Uddin MZ, Muramatsu D, Takemura N, Ahad MAR, Yagi Y. 2019. Spatio-temporal silhouette sequence reconstruction for gait recognition against occlusion. *IPSJ Transactions on Computer Vision and Applications* **11(1)**:1–18 DOI 10.1186/s41074-019-0061-3.

Wang M, Guo X, Lin B, Yang T, Zhu Z, Li L, Zhang S, Yu X. 2023b. DyGait: exploiting dynamic representations for high-performance gait recognition. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*. Piscataway: IEEE, 13424–13433.

Wang Z, Hou S, Zhang M, Liu X, Cao C, Huang Y, Li P, Xu S. 2024. Qagait: revisit gait recognition from a quality perspective. *Proceedings of the AAAI Conference on Artificial Intelligence* **38**:5785–5793 DOI 10.1609/aaai.v38i6.28391.

Wang L, Liu B, Liang F, Wang B. 2023a. Hierarchical spatio-temporal representation learning for gait recognition. In: *2023 IEEE/CVF International Conference on Computer Vision (ICCV)*. Piscataway: IEEE, 19582–19592.

Wang D, Tan D, Liu L. 2018. Particle swarm optimization algorithm: an overview. *Soft Computing* **22(2)**:387–408 DOI 10.1007/s00500-016-2474-6.

Wang L, Tan T, Ning H, Hu W. 2003. Silhouette analysis-based gait recognition for human identification. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **25(12)**:1505–1518 DOI 10.1109/TPAMI.2003.1251144.

Wang C, Wang Y, Chen Y, Liu H, Liu J. 2020. User authentication on mobile devices: approaches, threats and trends. *Computer Networks* **170(2)**:107118 DOI 10.1016/j.comnet.2020.107118.

Wei S, Liu W, Wei F, Wang C, Xiong NN. 2024a. Gaitdlf: global and local fusion for skeleton-based gait recognition in the wild. *The Journal of Supercomputing* **80(12)**:1–27 DOI 10.1007/s11227-024-06089-7.

Wei T, Liu M, Zhao H, Li H. 2024b. Gmsn: an efficient multi-scale feature extraction network for gait recognition. *Expert Systems with Applications* **252(7)**:124250 DOI 10.1016/j.eswa.2024.124250.

Wu Y, Hassan M, Hu W. 2022. Safegait: safeguarding gait-based key generation against vision-based side channel attack using generative adversarial network. *Proceedings of the ACM on Interactive, Mobile, Wearable and Ubiquitous Technologies* **6(2)**:1–27 DOI 10.1145/3534607.

Wu Z, Huang Y, Wang L, Wang X, Tan T. 2016. A comprehensive study on cross-view gait based human identification with deep CNNs. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39(2)**:209–226 DOI 10.1109/TPAMI.2016.2545669.

Wu X, Yuan Y, Zhang X, Wang C, Xu T, Tao D. 2021. Gait phase classification for a lower limb exoskeleton system based on a graph convolutional network model. *IEEE Transactions on Industrial Electronics* **69(5)**:4999–5008 DOI 10.1109/TIE.2021.3082067.

Xi H, Ren K, Lu P, Li Y, Hu C. 2024. Ssgait: enhancing gait recognition via semi-supervised self-supervised learning. *Applied Intelligence* **54(7)**:5639–5657 DOI 10.1007/s10489-024-05385-2.

Xia C, Munakata A, Sugiura Y. 2023. Privacy-aware gait identification with ultralow-dimensional data using a distance sensor. *IEEE Sensors Journal* **23(9)**:10109–10117 DOI 10.1109/JSEN.2023.3260846.

Xiong J, Zou S, Tang J, Tjahjadi T. 2024. Mcdgait: multimodal co-learning distillation network with spatial-temporal graph reasoning for gait recognition in the wild. *The Visual Computer* **40(10)**:1–14 DOI 10.1007/s00371-024-03426-y.

Xu Z, Lu W, Zhang Q, Yeung Y, Chen X. 2019. Gait recognition based on capsule network. *Journal of Visual Communication and Image Representation* **59(2)**:159–167 DOI 10.1016/j.jvcir.2019.01.023.

Xu C, Makihara Y, Li X, Yagi Y. 2023a. Gait recognition from fisheye images. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 1030–1040.

Xu C, Makihara Y, Li X, Yagi Y. 2023b. Occlusion-aware human mesh model-based gait recognition. *IEEE Transactions on Information Forensics and Security* **18(6)**:1309–1321 DOI 10.1109/TIFS.2023.3236181.

Xu C, Makihara Y, Li X, Yagi Y, Lu J. 2017. Speed invariance vs. stability: cross-speed gait recognition using single-support gait energy image. In: *Computer Vision–ACCV 2016: 13th*

Asian Conference on Computer Vision, November 20–24, 2016, Revised Selected Papers, Part II 13. Taipei, Taiwan: Springer, 52–67.

Xu C, Makihara Y, Li X, Yagi Y, Lu J. 2020. Cross-view gait recognition using pairwise spatial transformer networks. *IEEE Transactions on Circuits and Systems for Video Technology* **31(1)**:260–274 DOI 10.1109/TCSVT.2020.2975671.

Xu C, Tsuji S, Makihara Y, Li X, Yagi Y. 2023c. Occluded gait recognition via silhouette registration guided by automated occlusion degree estimation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 3199–3209.

Xu W, Xue W, Lin Q, Lan G, Feng X, Wei B, Luo C, Li W, Zomaya AY. 2021. Privgait: an energy-harvesting-based privacy-preserving user-identification system by gait analysis. *IEEE Internet of Things Journal* **9(22)**:22048–22060 DOI 10.1109/JIOT.2021.3089618.

Xue Z, Ming D, Song W, Wan B, Jin S. 2010. Infrared gait recognition based on wavelet transform and support vector machine. *Pattern Recognition* **43(8)**:2904–2910 DOI 10.1016/j.patcog.2010.03.011.

Xue J, Shen B. 2020. A novel swarm intelligence optimization approach: sparrow search algorithm. *Systems Science & Control Engineering* **8(1)**:22–34 DOI 10.1080/21642583.2019.1708830.

Yang Z, Xu B, Luo W, Chen F. 2022. Autoencoder-based representation learning and its application in intelligent fault diagnosis: a review. *Measurement* **189(5–6)**:110460 DOI 10.1016/j.measurement.2021.110460.

Ye D, Fan C, Ma J, Liu X, Yu S. 2024. BigGait: learning gait representation you want by large vision models. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 200–210.

Yu S, Tan D, Tan T. 2006. A framework for evaluating the effect of view angle, clothing and carrying condition on gait recognition. In: *18th International Conference on Pattern Recognition (ICPR'06)*. Vol. 4, Piscataway: IEEE, 441–444.

Zajac FE, Neptune RR, Kautz SA. 2002. Biomechanics and muscle coordination of human walking: part I: Introduction to concepts, power transfer, dynamics and simulations. *Gait & Posture* **16(3)**:215–232 DOI 10.1016/S0966-6362(02)00068-1.

Zajac FE, Neptune RR, Kautz SA. 2003. Biomechanics and muscle coordination of human walking: part II: lessons from dynamical simulations and clinical implications. *Gait & Posture* **17(1)**:1–17 DOI 10.1016/S0966-6362(02)00069-3.

Zhang Z. 2012. Microsoft kinect sensor and its effect. *IEEE Multimedia* **19(2)**:4–10 DOI 10.1109/MMUL.2012.24.

Zhang D. 2019. Wavelet Transform. In: *Fundamentals of Image Data Mining. Texts in Computer Science*. Cham: Springer, 35–44 DOI 10.1007/978-3-030-17989-2_3.

Zhang P, Dou H, Zhang W, Zhao Y, Qin Z, Hu D, Fang Y, Li X. 2023. A large-scale synthetic gait dataset towards in-the-wild simulation and comparison study. *ACM Transactions on Multimedia Computing, Communications and Applications* **19(1)**:1–23 DOI 10.1145/3517199.

Zhang H, Lu G, Zhan M, Zhang B. 2022. Semi-supervised classification of graph convolutional networks with laplacian rank constraints. *Neural Processing Letters* **54**:2645–2656 DOI 10.1007/s11063-020-10404-7.

Zhang P, Wu Q, Xu J. 2019a. VN-Gan: identity-preserved variation normalizing GAN for gait recognition. In: *2019 International Joint Conference on Neural Networks (IJCNN)*. Piscataway: IEEE, 1–8.

Zhang P, Wu Q, Xu J. 2019b. VT-GAN: view transformation gan for gait recognition across views. In: *2019 International Joint Conference on Neural Networks (IJCNN)*. Piscataway: IEEE, 1–8.

Zhao T, Zhang L, Ma Y, Cheng L. 2024. A survey on safe multi-modal learning systems. In: *Proceedings of the 30th ACM SIGKDD Conference on Knowledge Discovery and Data Mining*, 6655–6665.

Zheng A, Casari A. 2018. *Feature engineering for machine learning: principles and techniques for data scientists*. United States: O'Reilly Media, Inc.

Zheng Y, Li T, Ma W, Li Z, Wang L, Li Y. 2024b. Automated pricing and replenishment decision for vegetable products based on hybrid machine learning models. In: *Electronics, Communications and Networks*. Amsterdam: IOS Press, 81–87.

Zheng Y, Li Z, Ma W, Li T, Zheng J, Wang L. 2024d. Virtual reality content generation for anxiety and stress management: current status, challenges, and future developments. *SID Symposium Digest of Technical Papers* **55**:878–881 DOI 10.1002/sdtp.17227.

Zheng Y, Li T, Ma W, Zheng J, Li Z, Wang L. 2024c. Unveiling privacy challenges: big data-driven digital twins in smart city applications. *SID Symposium Digest of Technical Papers* **55**:49–52 DOI 10.1002/sdtp.16992.

Zheng J, Liu X, Liu W, He L, Yan C, Mei T. 2022. Gait recognition in the wild with dense 3D representations and a benchmark. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*, 20228–20237.

Zheng W, Zhu H, Zheng Z, Nevatia R. 2024a. GaitSTR: gait recognition with sequential two-stream refinement. *IEEE Transactions on Biometrics, Behavior, and Identity Science* **6**(4):528–538 DOI 10.1109/TBIOM.2024.3390626.

Zhu T, Fu L, Liu Q, Lin Z, Chen Y, Chen T. 2020. One cycle attack: fool sensor-based personal gait authentication with clustering. *IEEE Transactions on Information Forensics and Security* **16**:553–568 DOI 10.1109/TIFS.2020.3016819.

Zhu Z, Guo X, Yang T, Huang J, Deng J, Huang G, Du D, Lu J, Zhou J. 2021. Gait recognition in the wild: a benchmark. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision*, 14789–14799.