



Master Thesis Proposal

Human Gait Recognition Using Deep Learning

Waleed Mohammed Rasheedy

Master's Student, Master of Science in Artificial Intelligence Updated,
College of Informatics
Midocean University

Under Supervision of

Prof. Dr. Yasser Abdelhamid

Associate Professor, College of Informatics

Research Area

Biometric Recognition with a focus on Human Gait Recognition using
Deep learning

Abstract

Several techniques are being used widely for human authentication and recognition including fingerprints, iris, retina, etc. but human gait recognition plays a vital role due to its data acquisition approach. In gait recognition, the subject is not required to be in contact for data retrieval as gait patterns can be captured from a distance with the help of cameras. However, this method also has some factors that need to be addressed including variation of data due to clothing carrying, speed, shadow, and lightning. Previously many techniques are implemented by researchers for gait recognition. This work aims to develop a deep learning-based gait recognition method that would use publicly available datasets for evaluation. We aim to improve the accuracy with the help of improvement of the dataset and conducting extensive experiments with the help of different hyperparameters including learning rate, number of epochs, and optimizer. This work will help practical implementation of the gait method in a real-world environment.

Introduction

Several biometric authentication methods are being used for personal identification. These methods can be divided into physical or behavioral categories [1]. The face, fingerprints, palm, and ears lie under the physical biometric technique while in the behavioral category, voice, signature, and stride pattern are considered [2]. In the case of gait, the walking pattern is used without the consent of the subject under consideration [3]. The gait is considered appealing because it is comprised of 24 distinct composites such as angles and displacement of joints of the human body. The gait method is being widely used to ensure security [4].

Gait recognition has captured much attention in computer vision due to its potential in data capturing using non-invasive methods [5]. Researchers made a significant effort to the practical application of gait recognition in the past few decades. However, the majority of techniques are based on lateral-view data of human silhouettes [6] features extraction which makes it difficult due to a data collection requirement. This problem occurs in the real-time indoor surveillance environment where the large sequence of data collection is not feasible [7]. Furthermore, the cameras installed for security and surveillance are mostly installed in corridors, exits, or narrow places in which only frontal view is captured [8]. Therefore, the frontal view gait analysis is preferred to handle this problem.

These methods addressed the problem of frontal-view recognition at some level but still, there are problems of low accuracy due to environmental factors and complicated background [9]. Moreover, an extra layer of computational cost also makes it impractical due to the usage of camera depth and radar sensors.

Despite of importance, robustness, and efficiency of gait recognition, there are still problems interfering during the authentication and recognition of an individual. These factors include variations of view angles, clothing, load carriage, walking speed, shadow, etc. [10]. These problems reduce the recognition rate of a gait recognition system. To address these problems, several gait recognition methods are implemented that can be categorized as model-free and model-based.

The model-based approach aims to obtain the characterize of human stride which uses the human body [11]. Mostly this technique faces the problem of a lack of reliable approximation of the human body. However, this method needs high-quality images which makes it a challenging task. The model-based method can further be classified as pose-based and sensor-based. The model-free approach on the other hand focuses on the extraction of human silhouettes directly from the human body [12].

The gait recognition tasks can be implemented with the help of two broadly used methods including hand-crafted feature extraction [13] and deep learning-based features [14]. The hand-crafted features are referred to as the classical methods and are inaccurate and time-consuming. The deep learning-based methods are accurate and do not require human effort for feature extraction and automatically obtain features from the input data. The deep learning methods for gait recognition introduce new approaches to process the data and address the covariant problem such as clothing, carrying, and walking variation. However, these deep learning models reflect the attributes of a black box nature and may be problematic if not properly fine-tuned and need information about the pre-training conditions and parameters. To overcome these problems there is a demand for robust accurate, and efficient gait recognition methods that process the data correctly.

Problem Definition

In human gait recognition, the features are extracted from the walking style of an individual for authentication or recognition. This method is used in different domains such as security, health, sports, etc. In the past, many techniques have been implemented for gait recognition. These methods face the problem of different variations in the data due to changes in the view angles, clothes, carrying, walking speed, lightning, etc. Moreover, the data biases also create problems during recognition due to complexity. There are many techniques based on classical methods and deep learning approaches introduced in the past. There is a need to study these approaches to design the roust, accurate, and efficient gait recognition method that can address these challenges.

Human Gait Datasets

There are several publicly available datasets being utilized by researchers for human gait recognition.

Dataset Name	Link	Biomarker(s)	Size & Samples	Limitations
CASIA-A	http://www.cbsr.ia.ac.cn/english/Gait%20Data%20bases.asp	Silhouettes	20 subjects, 3 views (0°, 45°, 90°)	Small dataset, treadmill-based

CASIA-B	http://www.cbsr.ia.ac.cn/english/Gait%20Data%20bases.asp	Silhouettes, multi-view gait	124 subjects, 11 angles (0°–180°)	Limited real-world diversity, controlled environment
CASIA-C	http://www.cbsr.ia.ac.cn/english/Gait%20Data%20bases.asp	Infrared silhouettes	153 subjects, outdoor setting	Occlusions make recognition difficult
OU-ISIR Large Population	http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitMVLP.html#:~:text=Introduction,for%20cross%2Dview%20gait%20recognition.	Silhouettes, gait features	4,007 subjects	Limited pose variations
OU-ISIR Multi-view	http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitMVLP.html	Silhouettes, multi-view gait	10 views , indoor setting	Only side views are well-represented
OU-ISIR Age & Gender	http://www.am.sanken.osaka-u.ac.jp/BiometricDB/GaitLPAGE.html	Silhouettes, age & gender variations	1,615 subjects (ages 2-87)	No depth or RGB data
TUM-GAID	https://www.ce.cit.tum.de/mmk/verschiedenes/tum-gaid-database/	RGB, Depth, Silhouettes, Audio	305 subjects , 3 conditions (normal, backpack, coated)	Small dataset
SOTON Gait Database	https://web-archive.southampton.ac.uk/www.gait.ecs.soton.ac.uk/database/	Silhouettes, multi-view	115 subjects	Outdated dataset, low diversity

Florence 3D Gait (FVG)	https://cvlab.cse.msu.edu/frontal-view-gaitfvg-database.html	RGB, Depth, 3D Skeletons	180 subjects, indoor & outdoor	Small sample size
---------------------------------------	---	-----------------------------	--	----------------------

Related Work

To assess the contribution of important gait features in pattern recognition, a new method is presented for gait recognition that utilizes metaheuristic optimization for feature selection [15]. Initially, the features are extracted from the input data, and optimization is performed afterward with the help of the metaheuristic method. The optimized features are passed to various classification models for recognition. The method was tested with the help of 800 individual data and obtained 99.81% accuracy. The comparison of the method is also performed with the recent features selection method and this method offered improved results.

A gait recognition method to acquire cooperative motion of the human exoskeleton is designed based on the feature extraction of the deep learning model [16]. Initially, data is acquired with the help of a lightweight device using three sensors. Different conditions are considered to capture the data such as walking on stairs, slop, and being still. A Bi-LSTM model with an attention mechanism is used to obtain features from input data. the method was tested on 27 classes and achieved improved results.

A frontal view gait representation method called LFHEI based on the spatiotemporal features is implemented [17]. The method uses dynamic parts of the body and offers robust features. The LFHEI feature descriptor is passed to the deep learning model for feature extraction and gait recognition is performed using nearest neighbours. The CASIA dataset was used to test the method and obtained 95.0% accuracy on CASIA-A and 89.9% on CASIA-B. the method is also compared with the recent method and shows 4% improved results.

A new deep-learning model based on the attention mechanism is introduced to minimize the variability in the gait data [18]. Initially, the data is obtained in the form of gait energy images based on multi-recurrent features. Secondly, the deep learning model is used to obtain the relevant features from input data. they used attention-based blocks to highlight the important gait features for improved performance. The temporal relation of frames is obtained using the Bi-GRU layer. The method testing is carried out with the help of CASIA-B and OULP gait achieving accuracy of 93.4% and 97.4%.

To obtain spatiotemporal information from the contour of the human body, a gait recognition method used on the GaitSet model is presented [19]. This method aims to improve the recognition rate of video data based on multi multi-modality of silhouette and gait energy images. The method is tested with the help of CASIA-B and attained 95.8% on NM, 89.3% on BG, and 72.5% on CL which is improved in contrast to recent gait recognition methods.

A new gait recognition algorithm is designed for indoor data called GaitTS which is based on spatiotemporal features of different scales [20]. The method aims to address the problem of lacking

attention to the features of data. The GaitTS aggregates temporal features of multiple frames. The spatial features are divided into different maps having several sizes to give attention to different parts of the human body. The method obtained an accuracy of 97.9% on NM, 95.4% on BG, and 85.7% on CL respectively. An average accuracy of 76.50% is obtained on occCASIA-B and 93.0% on CCPG.

By employing a deformable convolution, a gait recognition method is called DCNGait which uses the temporal dimension [21]. The deformable convolution aims to enhance the temporal dynamics representation. The proposed TD3D convolution is capable of obtaining diverse features using temporal patterns. Careful feature extraction is performed by designing spatial and temporal blocks. The method DCNGait obtained 93.70% accuracy on CASIA-B.

A new gait recognition technique called Stacked-Gait [14] is designed that is based on several steps including image resizing, image conversion to grayscale, training, and stacking of autoencoders, and classification using machine learning algorithms. The Stacked-Gait was tested using CASIA-B and attained an average accuracy of 98.18% using CASIA-B.

The effectiveness of deep learning models including CNN, MLP, SOMs, and transfer learning with EfficientNet are used for gait recognition [22]. These models are utilized for robust feature extraction from the image data. They used 92,596 images of 64×64 size obtained from 10 individual sequences. Furthermore, a modified deep learning model based on a dense classification layer is also implemented which obtained 97.12% on the CASIA-B dataset.

Summary of Related Work

References	Datatype	Training Model	Accuracy	Dataset
Xu et al. [15]	Kinematics and kinetics data waveform	Machine Learning Models	99.81%	800 group gait data
Zhan et al. [16]	sEMG, GRF, and joint motion information	Bi-LSTM model	More than 90%	25 classes
Deng et al. [17]	Video data	Optical Flow and Deep Features	95.0% on CASIA-A and 89.9% CASIA-B	CASIA-A and CASIA-B
Junaid et al. [18]	Video Data	Attention-based Deep Learning	93.4% and 97.4%	CASIA-B and OULP

Shi et al. [19]	Multi-Modal	GaitSet Model	95.8%, 89.3%, and 72.5%	CASIA-B
Zhang et al. [20]	Image Sequence	Attention based CNN model	97.9% on NM, 95.4% on BG, and 85.7% on CL. 76.50% on occCASIA- B, 93.0% on CCPG.	CASIA-B, occCASIA-B, CCPG-
Kim et al. [21]	Image Sequence	DCNGait	93.70%	CASIA-B
Mehmood et al. [14]	Image Sequence	Stacked Autoencoders	98.10%	CASIA-B
Aman et al. [22]	Image Sequence	Deep Learning	97.12%	CASIA-B

Research Objectives

The main objectives of this work are given as:

- To design a deep learning model to address factors that affect the gait recognition method.
- To select a publicly available dataset and conduct experiments on that dataset to improve the recognition rate.
- To train the model after selecting hyperparameters after extensive experiments. These parameters include epochs, learning rate, optimizer, etc.
- To conduct a comparison of the proposed method with the recent gait recognition method to confirm the variability.

Research Plan

The following steps will be followed to write the survey paper on gait recognition.

Data Acquisition: To obtain a publicly available dataset that is famous by researchers to conduct experiments.

Data Preprocessing: This step includes improving data quality by noise moving, augmentation, or background removal.

Designing the Model: In this step a deep learning model will be designed based on various layers such as input, convolution, dropout, fully connected, and output.

Model Training and Feature Extraction: In this step, the designed model will be trained with the help of hyperparameters such as learning rate, epochs, and optimizer.

Validation and Testing: This step will include the validation and testing of the model with the help of test set of acquired datasets.

References

- [1] M. I. Sharif et al., "Deep learning and kurtosis-controlled, entropy-based framework for human gait recognition using video sequences," vol. 11, no. 3, p. 334, 2022.<https://doi.org/10.3390/electronics11030334>.
- [2] T. Teepe, J. Gilg, F. Herzog, S. Hörmann, and G. Rigoll, "Towards a deeper understanding of skeleton-based gait recognition," in Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, 2022, pp. 1569-1577.
- [3] H. Arshad et al., "A multilevel paradigm for deep convolutional neural network features selection with an application to human gait recognition," p. e12541, 2020.<https://doi.org/10.1111/exsy.12541>.
- [4] A. Mehmood, J. Amin, M. Sharif, and S. J. P. R. L. Kadry, "Human Gait Recognition by using Two Stream Neural Network along with Spatial and Temporal Features," vol. 180, pp. 16-25, 2024.<https://doi.org/10.1016/j.patrec.2024.02.010>.
- [5] K. Soni, D. P. Dogra, A. A. Sekh, S. Kar, H. Choi, and I.-J. J. E. S. w. A. Kim, "Person re-identification in indoor videos by information fusion using Graph Convolutional Networks," vol. 210, p. 118363, 2022.<https://doi.org/10.1016/j.eswa.2022.118363>.
- [6] M. Zhai, X. Xiang, N. Lv, and X. J. P. R. Kong, "Optical flow and scene flow estimation: A survey," vol. 114, p. 107861, 2021.<https://doi.org/10.1016/j.patcog.2021.107861>.
- [7] M. Munsif, S. U. Khan, N. Khan, A. Hussain, M. J. Kim, and S. W. J. K.-B. S. Baik, "Contextual visual and motion salient fusion framework for action recognition in dark environments," vol. 304, p. 112480, 2024.<https://doi.org/10.1016/j.knosys.2024.112480>.
- [8] M. Deng, Z. Fan, P. Lin, and X. J. I. T. o. M. Feng, "Human gait recognition based on frontal-view sequences using gait dynamics and deep learning," vol. 26, pp. 117-126, 2023. 10.1109/TMM.2023.3262131.
- [9] M. Deng, Z. Zhong, Y. Zou, Y. Wang, K. Wang, and J. J. N. P. L. Liao, "Human Gait Recognition Based on Frontal-View Walking Sequences Using Multi-modal Feature Representations and Learning," vol. 56, no. 2, p. 133, 2024.<https://doi.org/10.1007/s11063-024-11554-8>.
- [10] E. J. Harris, I.-H. Khoo, E. J. F. i. R. Demircan, and AI, "A survey of human gait-based artificial intelligence applications," vol. 8, p. 749274, 2022.<https://doi.org/10.3389/frobt.2021.749274>.
- [11] W. Xu, C. Zhu, and Z. J. P. R. L. Wang, "Multiview max-margin subspace learning for cross-view gait recognition," vol. 107, pp. 75-82, 2018.<https://doi.org/10.1016/j.patrec.2017.10.033>.

- [12] Z. Zhang et al., "Gait Recognition via Disentangled Representation Learning," in 2019 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), 2019, pp. 4705-4714, Long Beach, CA, USA: IEEE.
- [13] G. Liu, S. Zhong, and T. Li, "Gait recognition method of temporal-spatial HOG features in critical separation of Fourier correction points," *Future Generation Computer Systems*, vol. 94, pp. 11-15, May 1, 2019. [10.1016/j.future.2018.09.012](https://doi.org/10.1016/j.future.2018.09.012).
- [14] A. Mehmood, J. Amin, M. Sharif, S. Kadry, and J. J. P. o. Kim, "Stacked-gait: A human gait recognition scheme based on stacked autoencoders," vol. 19, no. 10, p. e0310887, 2024. <https://doi.org/10.1371/journal.pone.0310887>.
- [15] D. Xu et al., "A new method proposed for realizing human gait pattern recognition: Inspirations for the application of sports and clinical gait analysis," vol. 107, pp. 293-305, 2024. <https://doi.org/10.1016/j.gaitpost.2023.10.019>.
- [16] H. Zhan, J. Kou, Y. Cao, Q. Guo, J. Zhang, and Y. J. M. Shi, "Human Gait phases recognition based on multi-source data fusion and BiLSTM attention neural network," vol. 238, p. 115396, 2024. <https://doi.org/10.1016/j.measurement.2024.115396>.
- [17] M. Deng, Y. Zou, Z. Zeng, Y. Wang, X. Feng, and Y. J. E. A. o. A. I. Liu, "Human gait recognition based on frontal-view sequence using discriminative optical flow feature representations and learning," vol. 145, p. 110213, 2025. <https://doi.org/10.1016/j.engappai.2025.110213>.
- [18] M. I. Junaid, S. Madarapu, S. J. S. Ari, Image, and V. Processing, "Human gait recognition using attention based convolutional network with sequential learning," vol. 19, no. 1, pp. 1-8, 2025. <https://doi.org/10.1007/s11760-024-03765-2>.
- [19] X. Shi, W. Zhao, H. Pei, H. Zhai, and Y. J. I. A. Gao, "Research on gait recognition based on GaitSet and multimodal fusion," 2025.10.1109/ACCESS.2025.3533571.
- [20] L. Zhang, Z. Men, W. J. S. Xie, Image, and V. Processing, "Gaitts: indoor gait recognition with multi-scale temporal-spatial information aggregation," vol. 19, no. 1, pp. 1-12, 2025. <https://doi.org/10.1007/s11760-024-03611-5>.
- [21] J. Kim, B. Kim, and H. J. I. A. Lee, "Temporally deformable convolution for gait recognition," 2025.10.1109/ACCESS.2025.3526886.
- [22] N. Aman, M. R. Islam, M. F. Ahamed, and M. J. T. Ahsan, "Performance Evaluation of Various Deep Learning Models in Gait Recognition Using the CASIA-B Dataset," vol. 12, no. 12, p. 264, 2024. <https://doi.org/10.3390/technologies12120264>.