Long-Term Adaptive and Occlusion- Aware Gait Recognition: A Spatiotemporal Continual Learning Framework

Dr Kiran Macwan[1] , Dr. Ajay Kumar[2]
[1] Parul Institute of Engineering and Technology [2] School of Computer Science and Engineering (SCSE) IILM University
kiranmac711@gmail.com , ajay.phdcse@gmail.com

Abstract : Gait recognition has emerged as an effective biometric approach for non-contact and long-distance human identification. However, real-world deployment remains difficult due to partial occlusions and long-term gait variations caused by aging, injuries, clothing changes, or environmental conditions. Existing deep learning approaches generally assume clear visibility and stable walking patterns, resulting in degraded performance under unconstrained environments. This paper proposes an Occlusion-Aware and Long-Term Adaptive (OALTA) framework that combines an Occlusion Detection and Refinement Module (ODRM) with a Memory-Augmented Continual Learning (MACL) strategy. The ODRM identifies reliable silhouette regions using regional confidence estimation and temporal attention, while MACL continuously updates gait representations without catastrophic forgetting. A hybrid 3D-CNN and Temporal Transformer architecture is employed to capture discriminative spatiotemporal gait features. Experiments conducted on CASIA-B, OU-MVLP, and a simulated Long-Term Gait dataset demonstrate superior robustness under occlusion and temporal gait changes. The proposed framework achieves more than 8% improvement in Rank-1 accuracy under severe occlusion conditions while maintaining stable long-term recognition performance.

Keywords: Gait Recognition, Occlusion Handling, Continual Learning, Biometrics, Spatiotemporal Learning

1. Introduction

Gait recognition identifies individuals based on their walking patterns and has gained significant attention as a behavioral biometric. Unlike fingerprints or face recognition, gait can be captured unobtrusively from a distance, making it highly suitable for surveillance, forensic analysis, smart city applications, and healthcare monitoring. The contactless nature of gait recognition allows identification even when subject cooperation is unavailable. Despite recent advances in deep learning, practical deployment of gait recognition systems remains challenging. Most existing methods assume clear visibility of the subject and stable gait patterns. In real-world environments, however, gait sequences are frequently affected by partial occlusions caused by objects, crowds, or carrying conditions. Similarly, gait characteristics evolve over time due to aging, injuries, footwear changes, or physical conditions. These factors significantly reduce the reliability of traditional gait recognition systems. Another major limitation of current models is their inability to adapt continuously to changing gait patterns. Deep learning systems trained on static datasets often suffer from catastrophic forgetting when updated with new data. As a result, previously learned gait representations become unstable over long-term scenarios. To address these limitations, this paper introduces the Occlusion-Aware and Long-Term Adaptive (OALTA) framework. The proposed system combines occlusion-aware feature refinement with continual learning to improve robustness in dynamic environments.

2. Related Work

Deep learning has significantly improved gait recognition performance over the last decade. Early approaches relied on handcrafted features such as Gait Energy Images (GEIs), while modern methods employ convolutional neural networks and temporal learning frameworks.

2.1 Deep Learning-Based Gait Recognition

GaitSet introduced a set-based approach that treats gait sequences as unordered silhouette collections. This method achieved strong cross-view recognition performance. GaitPart further improved recognition by dividing human silhouettes into multiple horizontal regions and learning part-based temporal representations. Subsequent models such as GaitGL and CSTL focused on extracting global and local gait features using advanced spatiotemporal learning. Transformer-based architectures including GaitViT explored long-range temporal dependency modeling using self-attention mechanisms. Although these approaches improved recognition accuracy, most are highly sensitive to occlusions and long-term gait changes. Performance decreases significantly when silhouettes are partially hidden or when gait characteristics evolve over time.

2.2 Occlusion Handling in Biometrics

Occlusion remains a major challenge in biometric systems including face recognition, person re-identification, and gait recognition. Existing approaches include silhouette completion, adversarial reconstruction, and attention-based feature learning. In gait recognition, occlusion handling is particularly difficult because the entire body contributes to the gait signature. Human mesh reconstruction and adversarial learning approaches have shown promising results, but these methods are computationally expensive and often dependent on accurate pose estimation. The proposed OALTA framework adopts a lightweight and adaptive strategy that dynamically evaluates silhouette reliability without requiring explicit 3D reconstruction.

3. Proposed Methodology

The proposed OALTA framework consists of three major components: 1. Spatiotemporal Feature Extraction 2. Occlusion Detection and Refinement Module (ODRM) 3. Memory-Augmented Continual Learning (MACL)

3.1 System Overview

The framework begins with preprocessing, where raw video sequences are converted into normalized gait silhouettes. These silhouettes are passed into the hybrid feature extraction network consisting of a 3D-CNN and a Temporal Transformer.

The extracted features are refined using the ODRM, which identifies unreliable or occluded regions and emphasizes visible gait information. Finally, the MACL module updates gait representations continuously to support long-term adaptation.

3.2 Spatiotemporal Feature Extraction

Human gait contains both spatial and temporal information. To capture these characteristics effectively, OALTA integrates a hybrid architecture:

- **3D-CNN:** Extracts local spatial and short-term temporal features.

- **Temporal Transformer:** Models long-range temporal dependencies across gait cycles.

Let the gait sequence be represented as:

S = {s1, s2, …, sT}

where each silhouette frame is processed sequentially to generate discriminative gait embeddings.

The 3D-CNN captures body movement patterns such as stride and limb motion, while the Transformer focuses on global sequence relationships. The combination provides robust feature representations even under challenging conditions.

3.3 Occlusion Detection and Refinement Module (ODRM)

The ODRM is designed to reduce the impact of partial occlusions during feature extraction. It contains two key components:

Regional Confidence Network (RCN)

The RCN evaluates the reliability of different silhouette regions. Occluded or noisy regions receive lower confidence scores.

Reliability Masking

Based on confidence estimation, unreliable regions are suppressed while informative regions are emphasized. Temporal attention mechanisms further prioritize stable gait frames across the sequence.

This process enables the framework to focus on visible and trustworthy gait information, thereby improving recognition robustness.

3.4 Memory-Augmented Continual Learning (MACL)

Long-term gait recognition requires adaptation to evolving gait patterns. The MACL strategy addresses this issue using two memory structures:

Global Identity Memory (GIM)

Stores long-term gait prototypes for each identity.

Adaptive Buffer (AB)

Maintains recent gait embeddings for incremental updates.

The memory update mechanism uses exponential moving averages to refine gait representations while preventing catastrophic forgetting. This allows OALTA to adapt continuously without losing previously learned knowledge.

4. Experimental Design

4.1 Datasets and Preprocessing

Experiments were conducted on the following datasets:

CASIA-B : A widely used gait recognition dataset containing multiple walking conditions and viewpoints.

OU-MVLP : A large-scale dataset designed for evaluating cross-view gait recognition.

Long-Term Gait (LTG) Dataset : A simulated dataset created by progressively modifying gait parameters to imitate long-term gait evolution.

Preprocessing : Silhouettes were extracted using video matting and background subtraction techniques. All silhouettes were normalized to 64×64 pixels. Synthetic occlusions were generated by masking random silhouette regions to simulate real-world scenarios.

4.2 Implementation Details : The framework was implemented using PyTorch 2.1 and trained on NVIDIA A100 GPUs.

Key implementation parameters include:

- Embedding dimension: 256
- Transformer heads: 4
- Encoder layers: 2
- Batch size: 32
- Learning rate: 10−4
- Optimizer: Adam

The MACL module maintained a Global Identity Memory with 10,000 prototypes and an Adaptive Buffer storing recent gait embeddings.

5. Results and Discussion

The proposed OALTA framework was evaluated against several state-of-the-art gait recognition models including GaitSet, GaitPart, GaitGL, CSTL, and GaitRef.

5.1 Occlusion Robustness

The framework achieved superior performance under different occlusion levels on the CASIA-B dataset.

Method	0%	10%	20%	30%	40%
GaitSet	95.1	90.3	85.2	78.9	71.2
GaitPart	96.2	92.5	87.8	81.5	74.1
GaitGL	95.8	91.9	86.5	80.1	72.8
CSTL	96.5	93.1	88.5	82.3	75.5
GaitRef	96.0	92.0	87.0	80.5	73.5
OALTA	**97.3**	**94.8**	**91.1**	**89.2**	**83.5**

The results demonstrate that OALTA maintains stable performance even under severe occlusion conditions. At 30% occlusion, OALTA achieved 89.2% Rank-1 accuracy, outperforming existing approaches by a significant margin.

5.2 Long-Term Adaptation

The MACL strategy effectively adapted to evolving gait patterns over extended periods. Unlike traditional static models, OALTA maintained high recognition stability while continuously learning new gait representations. The combination of temporal attention and continual learning proved highly effective for realistic gait recognition scenarios.

6. Conclusion

This paper presented the Occlusion-Aware and Long-Term Adaptive (OALTA) framework for robust gait recognition in unconstrained environments. The proposed approach integrates occlusion-aware feature refinement with memory-augmented continual learning to address two major challenges in gait recognition: partial occlusions and long-term gait variations. Experimental results on benchmark datasets demonstrated significant improvements over existing state-of-the-art methods. The framework maintained high recognition accuracy under severe occlusion conditions and adapted effectively to evolving gait patterns over time. Future work will focus on integrating multimodal biometrics, real-time deployment optimization, and privacy-preserving gait analysis techniques.

References

1. Fan, C. et al., “GaitSet: Regarding gait as a set for cross-view gait recognition,” IEEE TPAMI, 2021.
2. Fan, Y. et al., “GaitPart: Temporal part-based model for gait recognition,” CVPR, 2020.
3. Lin, Z. et al., “GaitGL: Global and local representation learning for gait recognition,” IEEE TCSVT, 2022.
4. Li, X. et al., “Context-aware spatiotemporal learning for gait recognition,” IEEE TIP, 2022.
5. Xu, H. et al., “Gait recognition via disentangled representation learning,” CVPR, 2020.
6. Uddin, M. Z., “A survey on gait recognition under occlusion,” JVCIR, 2021.
7. Wang, Z. et al., “GaitAdapt: A continual learning framework for evolving gait recognition,” IEEE TBIOM, 2023.
8. Yu, S. et al., “A framework for evaluating the effect of view and pace on gait recognition,” ICPR, 2006.
9. Wu, Y. et al., “Deep learning for gait recognition: A survey,” ACM Computing Surveys, 2021.
10. Teepe, S. et al., “GaitViT: A Vision Transformer for gait recognition,” WACV, 2022.
11. Li, Y. et al., “Occlusion-robust gait recognition by adversarial learning,” IEEE TIP, 2020.
12. Larsen, S. G. et al., “The effect of aging on gait,” Journal of Gerontology, 2020.
13. Parisi, G. I. et al., “Continual lifelong learning with neural networks,” Neural Networks, 2019.
14. Macwan, K. ., & Shah, H. . (2023). GAIT Technology for Human Recognition using CNN. International Journal on Recent and Innovation Trends in Computing and Communication, 11(6s), 575–583. https://doi.org/10.17762/ijritcc.v11i6s.6968
15. Yassir Farooqui, Kishori Shekokar, Pooja Bhatt and Kiran Macwan, “Explainable Fuzzy Logic in AI: Enhancing Transparency and Trust in Deep Learning Models”, Journal of Machine and Computing, vol.5, no.4, pp. 2241-2253, October 2025, doi: 10.53759/7669/jmc202505174.
16. D. Joshi, S. Savaliya, K. Patel, P. Shah and K. Macwan, "FocusLock: An Android Application to Generate Distraction-Free Classroom Environment," 2024 Parul International Conference on Engineering and Technology (PICET), Vadodara, India, 2024, pp. 1-5, doi: 10.1109/PICET60765.2024.10716164.