Progress in Gait Recognition:

Beyond Constrained Datasets to Large Vision Models

Talker: Chao Fan (樊超), third-year Ph.D. student

E-mail: 12131100@mail.sustech.edu.cn

Github: https://github.com/ShiqiYu/OpenGait

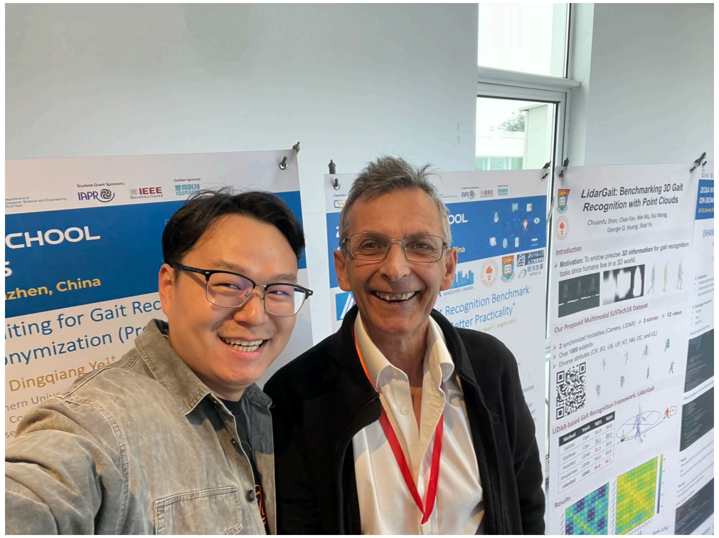
Homepage: https://chaofan996.github.io

南方科技大学
SOUTHERN UNIVERSITY OF SCIENCE AND TECHNOLOGY

Outlines

- What’s Gait Recognition?
- Why We Need Gait Recognition?
- Challenges and Our Solutions

What’s Gait Recognition?



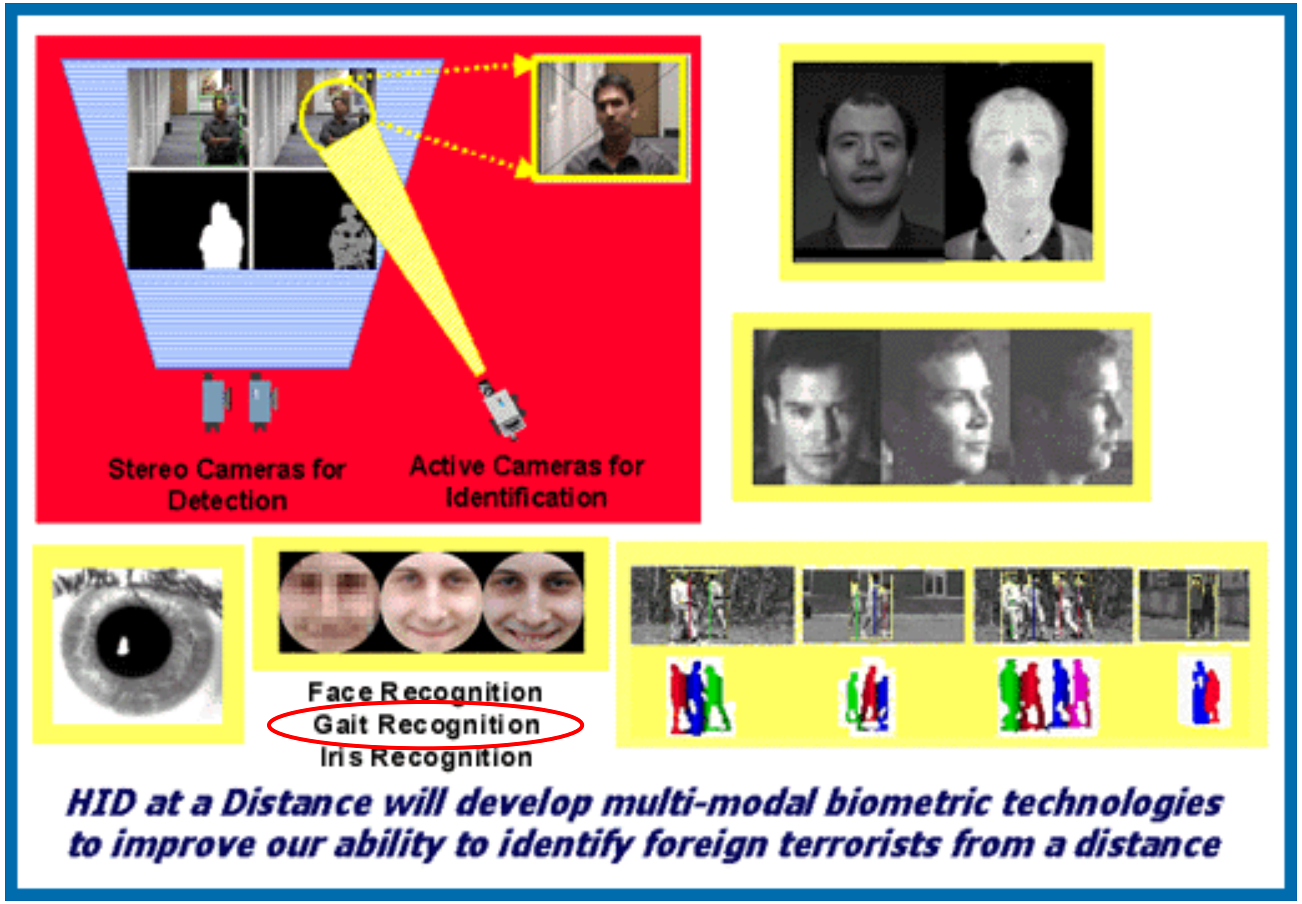
Me with **Prof. Mark Nixon**, a pioneer researcher in Gait Recognition (or two, ha ha, hopefully)

Mission Impossible 5

Gait is a kind of **behavioral biometric feature,** which raw data are video sequences presenting walking people

The goal of gait recognition is to identify people based on their walking patterns

Why We Need Gait Recognition?



The HID Program in 2000s:

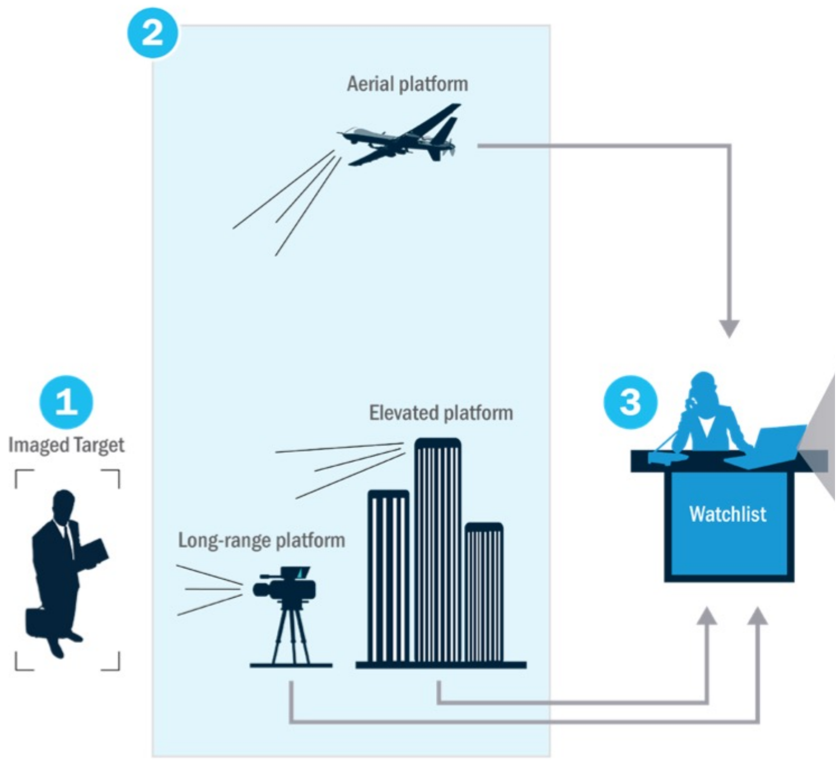
(Funded by Defense Advanced Research Projects Agency)

Improve Our Ability to Identify Foreign Terrorists from a Distance

https://vrphobia.com/research/darpa-human-identification-at-a-distance-hid/

Why We Need Gait Recognition?

BRIAR - Biometric Recognition and Identification at Altitud



KEYWORDS

- Biometrics
- Atmospheric turbulence
- Long-range
- Unmanned aerial vehicles
- Machine learning
- Algorithms
- Face recognition
- Whole-body recognition
- Gait recognition

The BRIAR Program in 2020s:

(Funded by Intelligence Advanced Research Projects Activity)

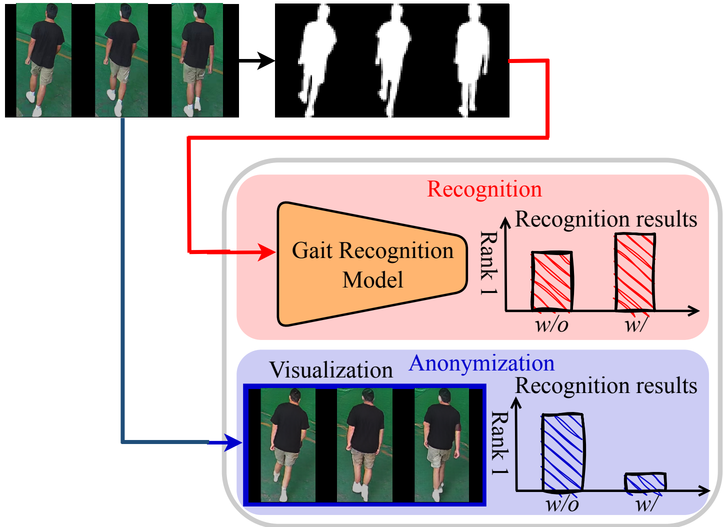
Biometric Recognition and Identification at Altitude and Range

https://www.iarpa.gov/images/OA-Slicksheets/BRIAR_SlickSheet_03192022.pdf

Why We Need Gait Recognition?

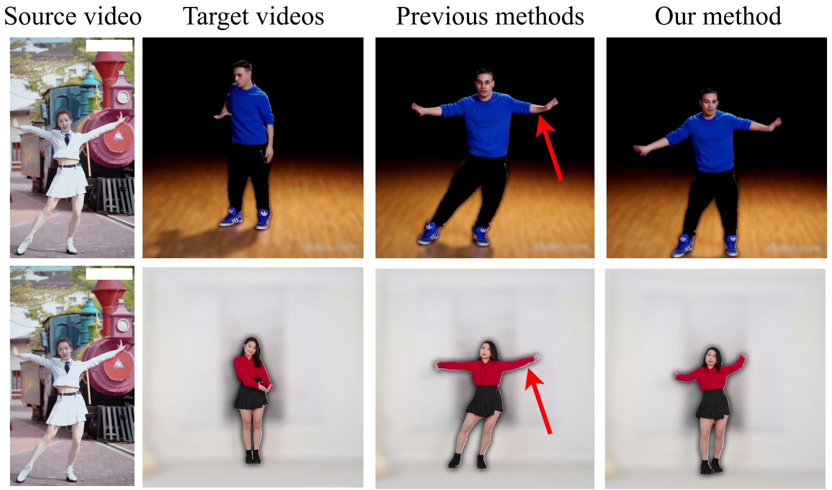
Applications Beyond Human Identification

[1] Anonymization



People may identify you through Internet videos even when you show no face.
But take it easy, we can protect your walking patterns by Gait Anonymization.

[2] Individuation



Gait Individuation can preserve your uniqueness in the human-centered AIGC applications, specifically for the popular video-based ones.

[3] Medical Purposes

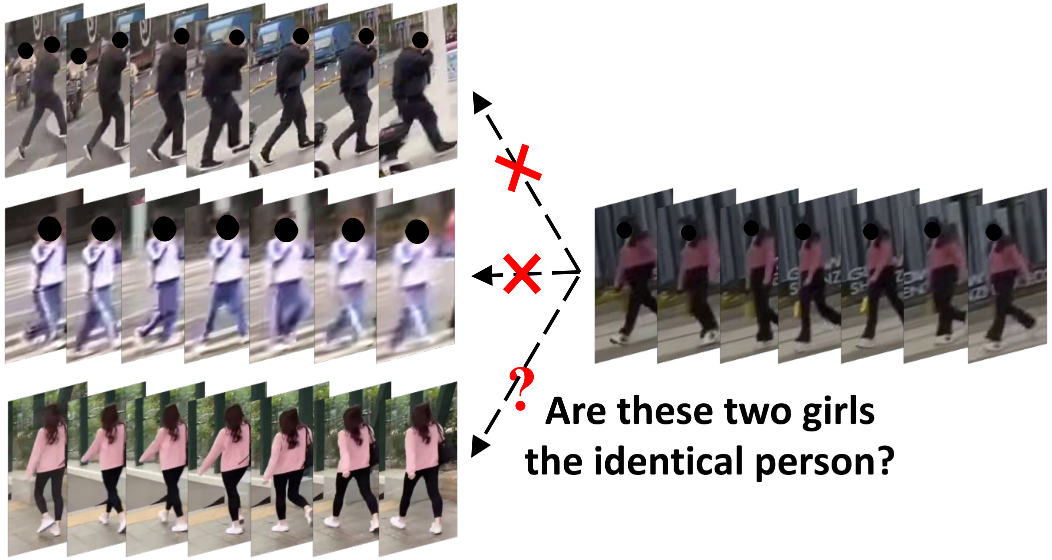
(e.g., population-scale scoliosis screening)

The human gait, as a kind of biomarker, can reflect various diseases related to muscle stiffness and posture instability.
Its camera-based applications are well-suitable for population-scale early diagnosis.

[1] J. Ma*, D. Ye*, C. Fan*, and S. Yu, Pedestrian Attribute Editing for Gait Recognition and Anonymization, Arxiv 2024.
[2] J. Ma, X. Zhang, and S. Yu, An Identity-Preserved Framework for Human Motion Transfer, TIFS 2024.
[3] Z. Zhou, J. Liang, Z. Peng, C. Fan, F. An, and S. Yu, Gait Patterns as Biomarkers: A Video-Based Approach for Classifying Scoliosis, Arxiv 2024.

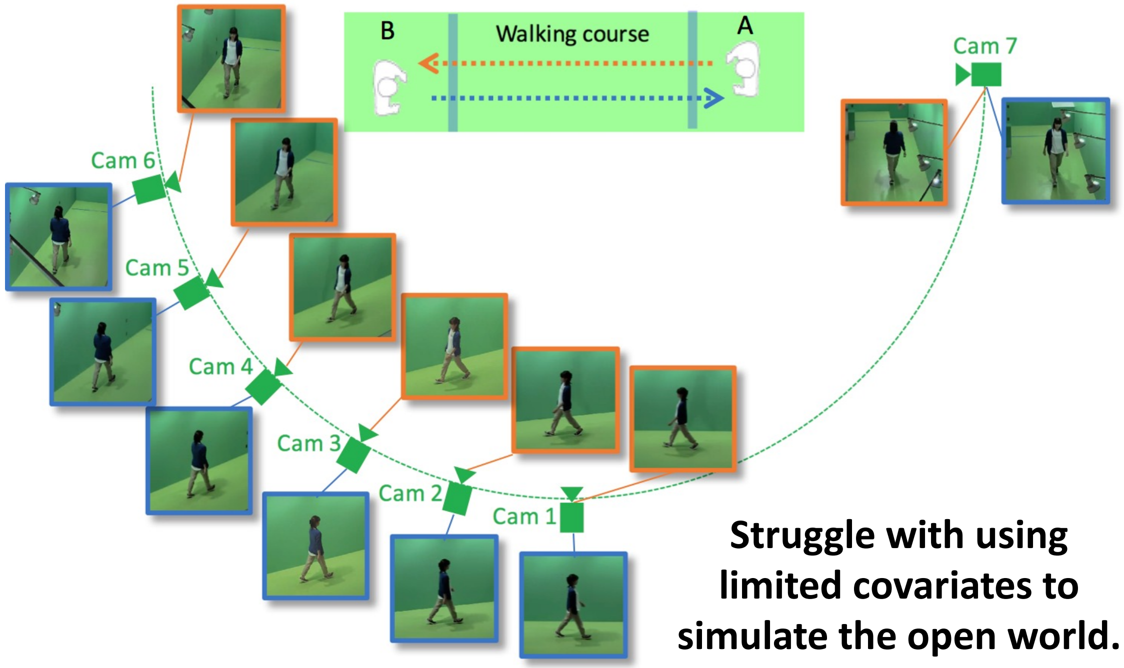
Challenge#1: Large-scale Gait Data Collection

[1] Gait Data Collection In the Wild



If there are no other apparent soft-biometric features or auxiliary information, such as gender, age, dress, location, background, and so on, it would be almost impossible for an annotator to identify strangers from massive unlabelled videos.

[2] Gait Data Collection In the Lab



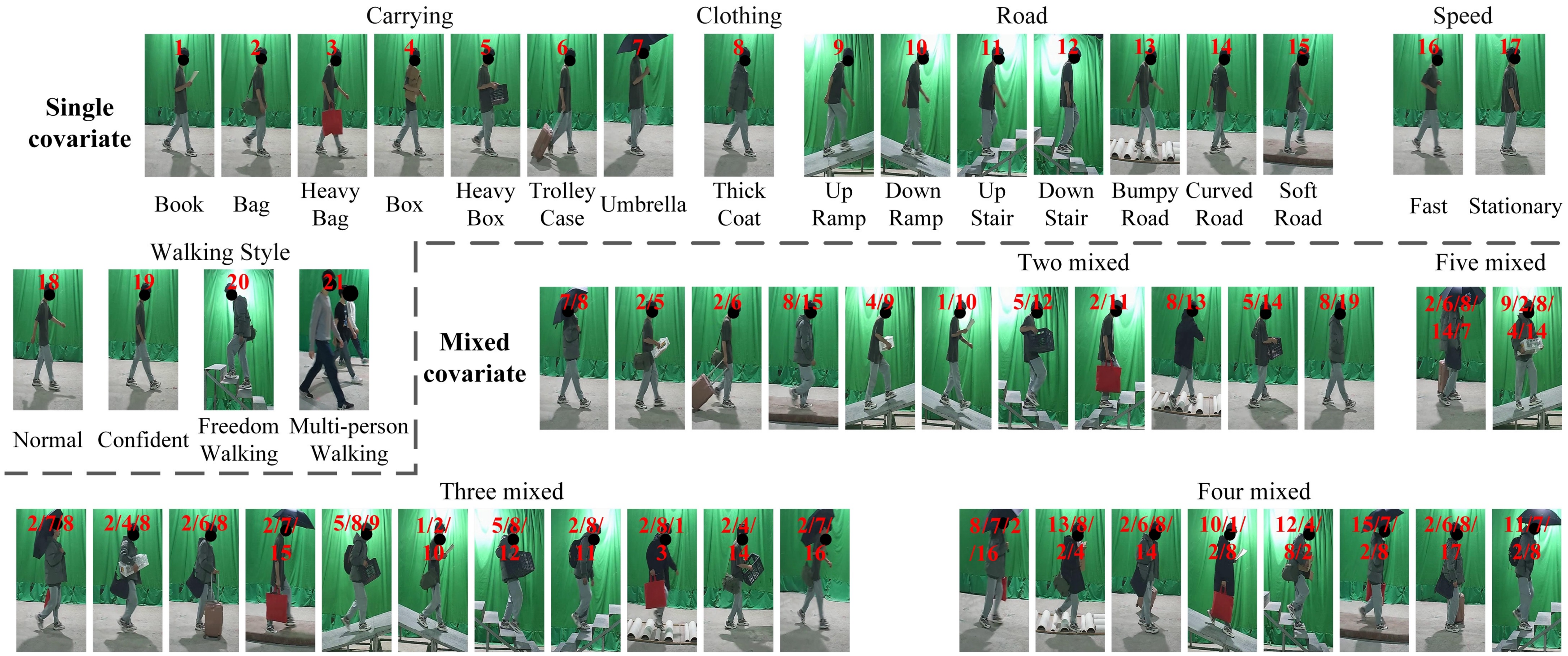
Struggle with using limited covariates to simulate the open world.

Require participants walking around the fixed course repeatedly.

[1] C. Fan, S. Hou, J. Wang, Y. Huang, and S. Yu, Learning gait representation from massive unlabelled walking videos: A benchmark, T-PAMI2023.
[2] N. Takemura, Y. Makihara, D. Muramatsu, T. Echigo, and Y. Yagi, Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition, IPSJ T-CVA 2018.

Challenge#1: Large-scale Gait Data Collection

Solution#1.1: Collecting Gait Data with Exhaustively Rich Covariates



✅ The first million-level gait recognition dataset
✅ $53 \times 33 = 1{,}749$ walking conditions

We can carefully evaluate Gait Recognition under various complex scenarios.

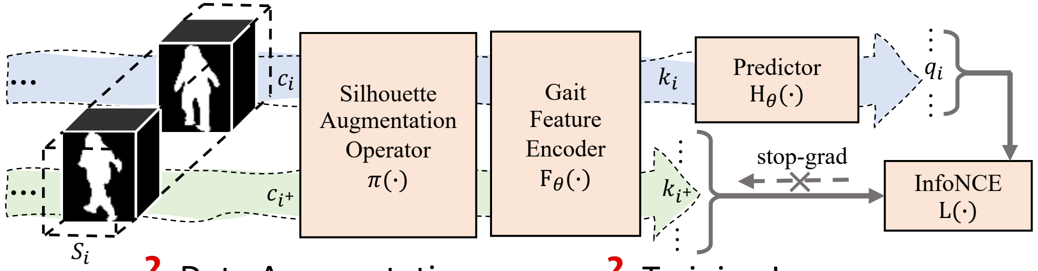
[1] S. Zou, C. Fan, J. Xiong, C. Shen, S. Yu, and J. Tang, Cross-covariate Gait Recognition: A Benchmark, AAAI2024.

Challenge#1: Large-scale Gait Data Collection

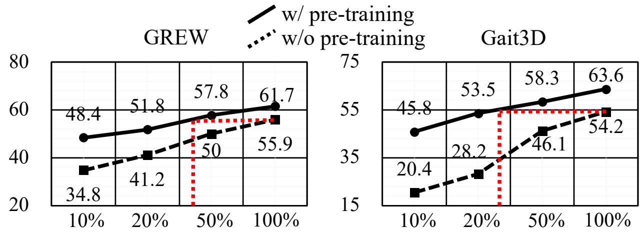
Solution#1.2: Learning Gait Priors from Massive Unlabelled Walking Videos

GaitLU-1M: the First Large-scale Unlabelled Gait Dataset

GaitSSB: A Self-Supervised Baseline for Gait Pretraining



? Data Augmentation
? Gait Encoder
? Training Loss
? Theoretical Analysis



✅ Within-domain Improvements 🚀

Method	Source	Cross-domain: Source→Target GREW→Gait3D	Cross-domain: Source→Target Gait3D→GREW
GaitSet	AAAI2019	19.0	19.2
GaitPart	CVPR2020	19.3	14.2
GaitGL	ICCV2021	15.6	14.3
GaitSSB	w/o pre-training	20.1	20.6
	w/ pre-training	**27.2**	**27.6**

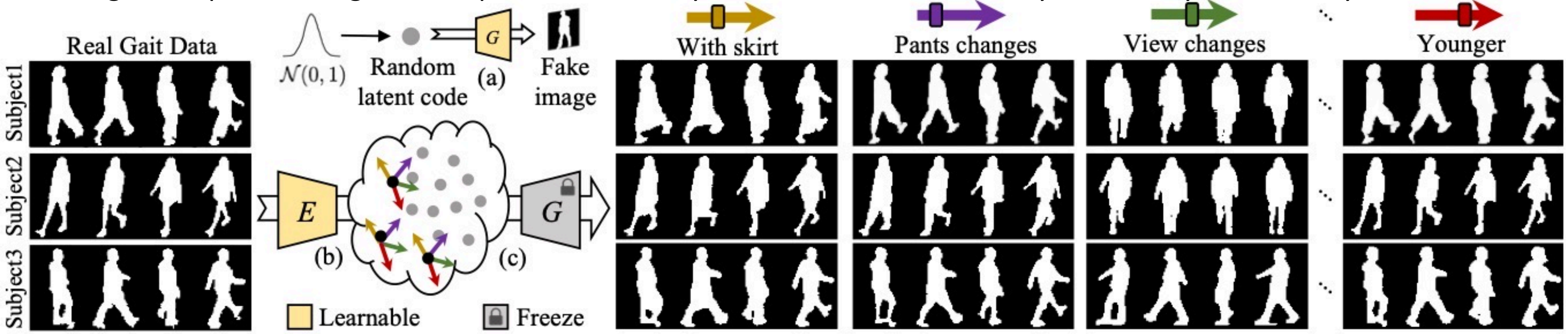
✅ Cross-domain Improvements 🚀

[1] C. Fan, S. Hou, J. Wang, Y. Huang, and S. Yu, Learning gait representation from massive unlabelled walking videos: A benchmark, T-PAMI2023

Challenge#1: Large-scale Gait Data Collection

Solution#1.3: Gait Attribute Editing for Gait Data Augmentation

❓ Large Viewpoint Change ❓ Temporal Consistency ❓ Attribute Consistency ❓ Identity Consistency



GaitEditor: the First Gait Attribute Editing Framework

✅ Improvements Verified on Various Datasets 🚀

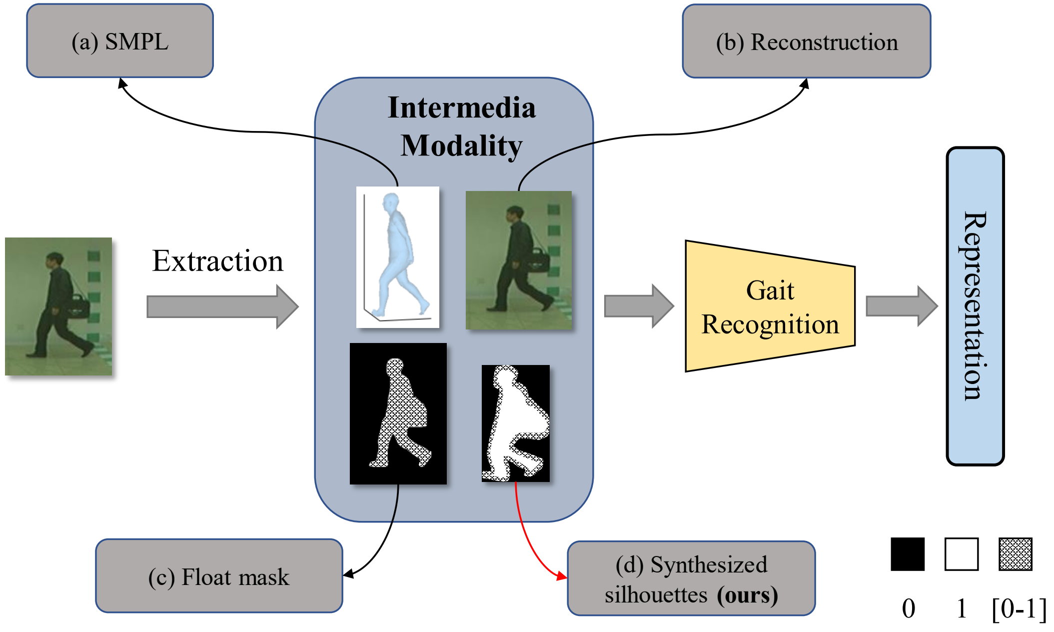
Datasets	GaitEditor	Rank-1 accuracy (%) Different conditions	Mean
OU-MVLP	✘	NM: 89.85	89.85
	✔	**NM: 90.27**	**90.27**
CCPG	✘	CL: 71.56, UP: 75.00, DN: 76.85, BG: 78.52	75.48
	✔	**CL: 73.21, UP: 76.98, DN: 77.81, BG: 80.32**	**77.08**
Gait3D	✘	64.60	64.60
	✔	**65.40**	**65.40**

✅ Improvements Verified by Various Gait Models 🚀

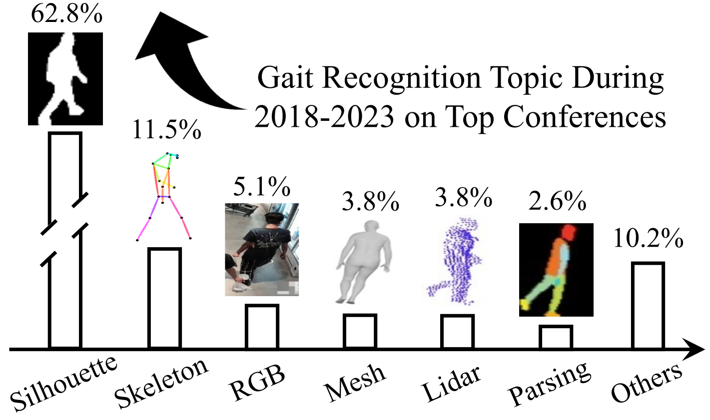
Gait models	GaitEditor	Rank-1 accuracy (%) CL	UP	DN	BG	Mean
GaitSet	✘	60.18	65.22	65.11	68.53	64.76
	✔	**61.30**	**66.09**	**66.21**	**69.45**	**65.76**
GaitPart	✘	64.25	67.76	68.58	71.68	68.07
	✔	**65.28**	**70.15**	**71.36**	**73.98**	**70.19**
GaitBase	✘	71.56	75.00	76.85	78.52	75.48
	✔	**73.21**	**76.98**	**77.81**	**80.32**	**77.08**
DeepGaitv2	✘	78.37	84.67	80.95	89.45	83.36
	✔	**79.76**	**85.70**	**82.30**	**90.83**	**84.65**

[1] J. Ma*, D. Ye*, C. Fan*, and S. Yu, Pedestrian Attribute Editing for Gait Recognition and Anonymization, Arxiv 2024.

Challenge#2: Good Gait Representation Exploration



In most cases, gait models tend to utilize the intermedia representations extracted from RGB videos rather than the videos themselves as inputs.

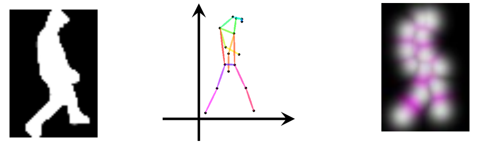


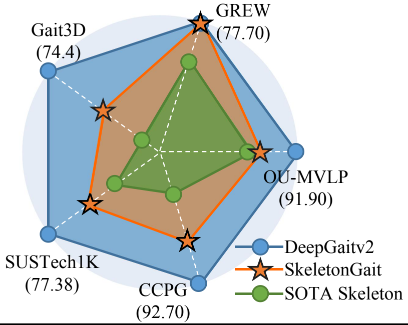
- On the one hand, we want gait representations contain more discriminative characteristics;
- On the other hand, we also want them to be as clean as possible, i.e., extremely excluding gait-unrelated factors such as color and texture elements.

[1] J. Liang*, C. Fan*, S. Hou, C. Shen, Y. Huang, and S. Yu, GaitEdge: Beyond Plain End-to-end Gait Recognition for Better Practicality, ECCV2022.

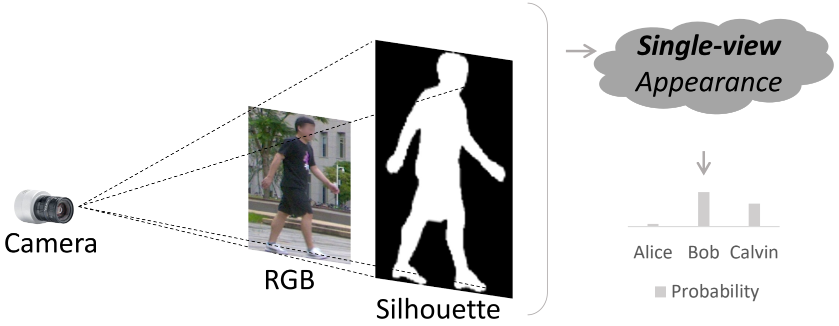
Challenge#2: Good Gait Representation Exploration

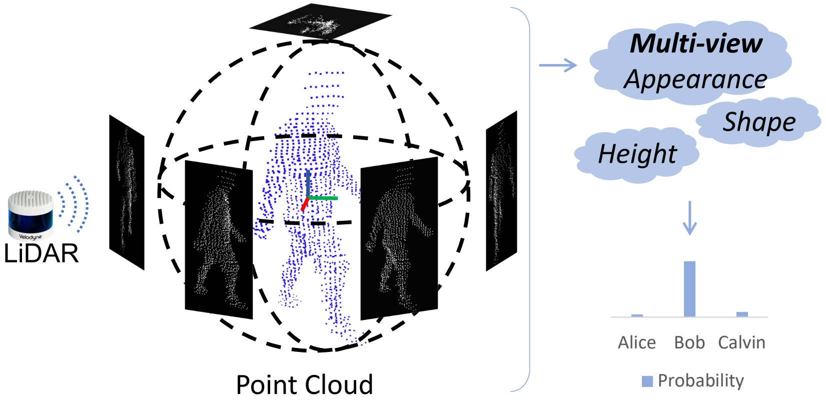
Solution#2.1: Exploring More



Representation	Binary Image	Coordinates	Skeleton Map
Format	Image	Graph	Image
Architecture	CNN	GCN	CNN
Shape Features	√	×	×
Structural Features	√	√	√
Model	DeepGaitV2	SOTA Skeleton	SkeletonGait
Main Results			

SkeletonGait++: A SOTA Multi-modal Gait Framework



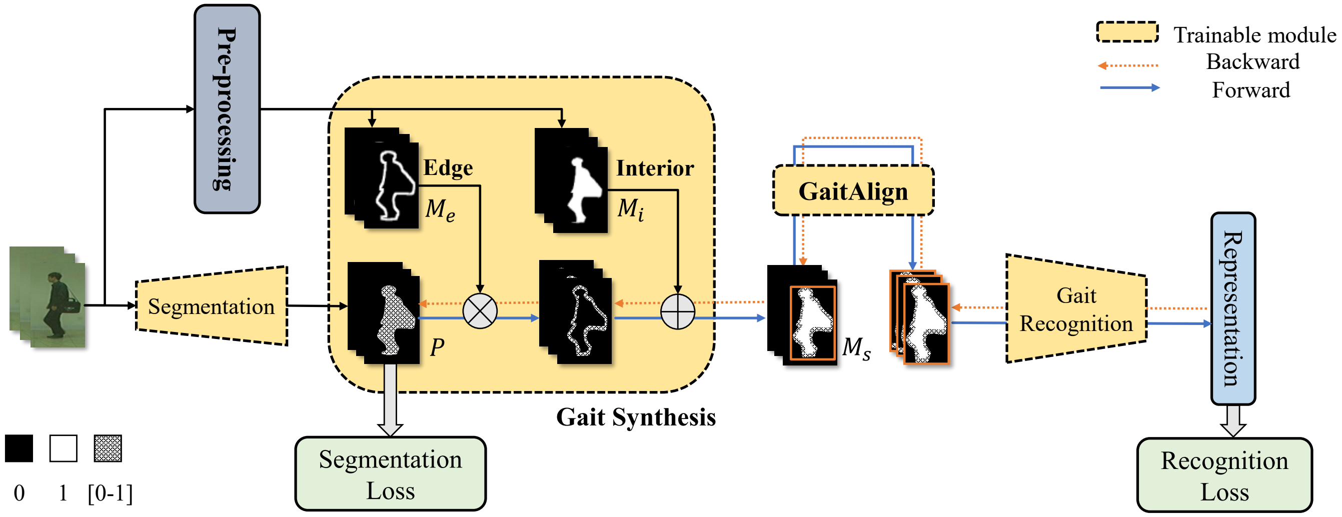


LidarGait: the First Large-scale Lidar-based Gait Benchmark

[1] C. Fan, J. Ma, D. Jin, C. Shen, and S. Yu, SkeletonGait: Gait Recognition Using Skeleton Maps, AAAI2024.
[2] C. Shen, C. Fan, W. Wu, R. Wang, G. Q.Huang, and S. Yu, LidarGait: Benchmarking 3D Gait Recognition with Point Clouds, CVPR2023.

Challenge#2: Good Gait Representation Exploration

Solution#2.2: End-to-end with Making Gait Edge Trainable



✅ Within-domain Improvements 🚀

✅ Cross-domain Improvements 🚀

Training Set	Method		Test Set: CASIA-B* NM	BG	CL	Mean	TTG-200
			NM	BG	CL	Mean	-
CASIA-B*	Two-step	GaitSet [4]	92.30	86.10	73.36	83.92	40.26
		GaitPart [5]	93.14	85.99	75.05	84.72	42.23
		GaitGL [16]	94.15	89.98	81.42	88.52	**(48.74)**
	End2end	GaitGL-E2E	99.06	98.24	89.45	**95.58**	37.18
		GaitEdge	97.94	96.06	86.36	**93.45**	**(49.12)**
TTG-200	Two-step	GaitSet [4]	41.32	35.15	21.59	32.69	77.62
		GaitPart [5]	45.21	38.75	25.92	36.62	80.24
		GaitGL [16]	50.47	45.29	40.34	**(45.37)**	80.46
	End2end	GaitGL-E2E	51.24	45.93	27.18	41.45	**90.37**
		GaitEdge	54.76	49.85	38.16	**(47.59)**	**88.66**

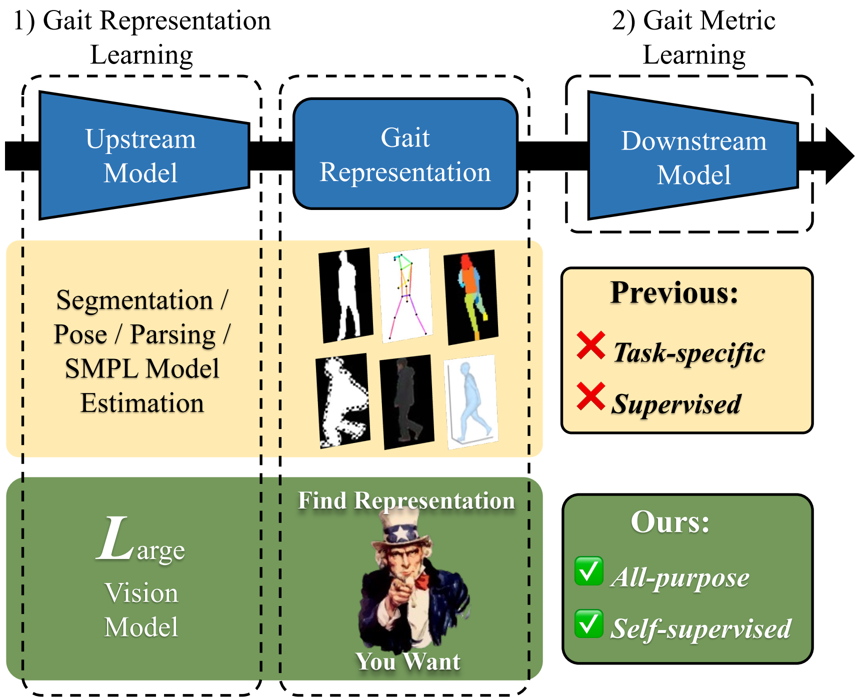
GaitEdge:
- ➢ *Making Edge Regions Differentiable*
- ➢ *Making Non-edge Regions Nondifferentiable*

Core Motivations:
- ➢ *the RGB-informed noises are mainly distributed in the non-edge regions*
- ➢ *the edge region plays a vital role in describing the shape of the human body*

[1] J. Liang*, C. Fan*, S. Hou, C. Shen, Y. Huang, and S. Yu, GaitEdge: Beyond Plain End-to-end Gait Recognition for Better Practicality, ECCV2022.

Challenge#2: Good Gait Representation Exploration

Solution#2.3: Learning Gait Representation You Wanted by Large Vision Models



The idea of BigGait

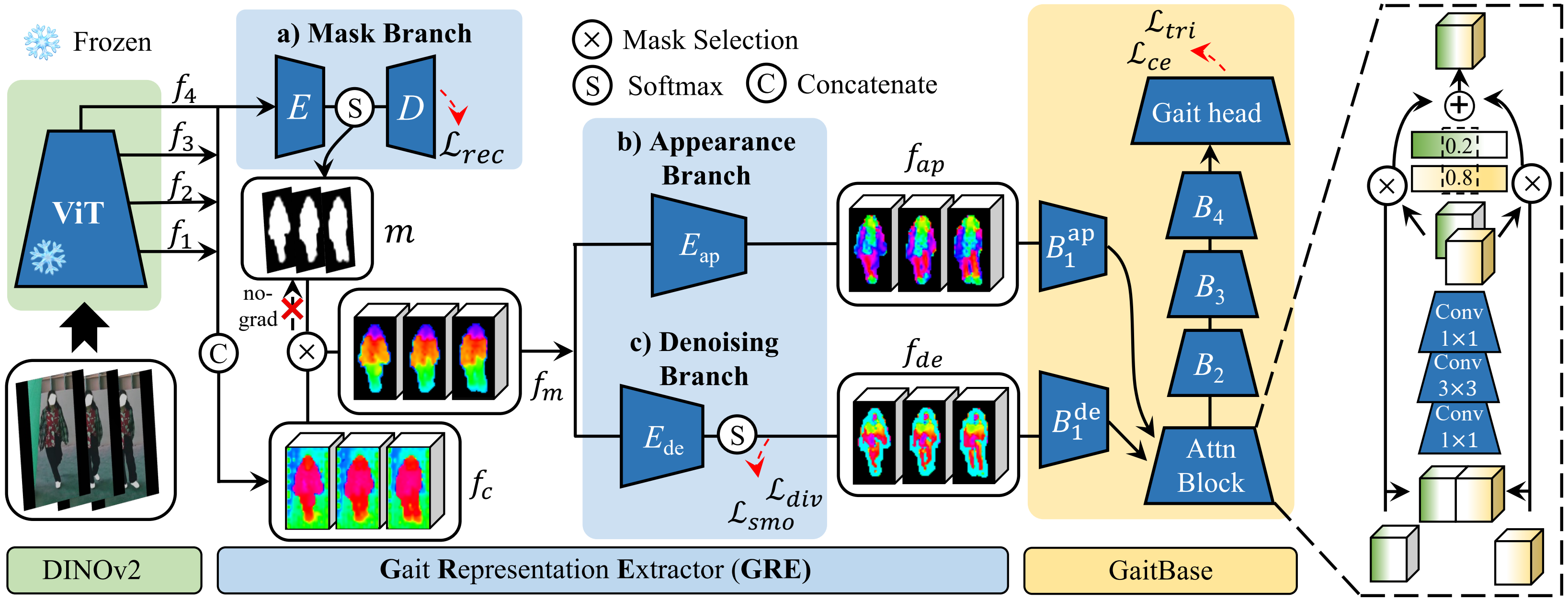
Escaping from the reliance on upstream task-specific models, we make a pioneering effort to acquire desired gait features from LVMs:

- *The* ***discriminability and generalization*** *of all-purpose features produced by LVMs;*
- *Self-supervised pre-training of LVMs* ***obviates the need for annotating*** *elements such as the silhouette, skeleton, and more, on a large scale;*
- *We can* ***avoid cumulative errors*** *imposed by specific upstream tasks to a large extent.*

[1] D. Ye*, C. Fan*, J. Ma, X. Liu, and S. Yu, BigGait: Learning Gait Representation You Want by Large Vision Models, CVPR2024.

Challenge#2: Good Gait Representation Exploration

Solution#2.3: Learning Gait Representation You Wanted by Large Vision Models



❓ How to exclude gait-unrelated elements within all-purpose features?

- *Mask Branch: Remove the Background on the Whole*
- *Appearance Branch: Maintain the Most Discriminative Features*
- *Denoising Branch: Reduce High-frequency Textural Characteristics*

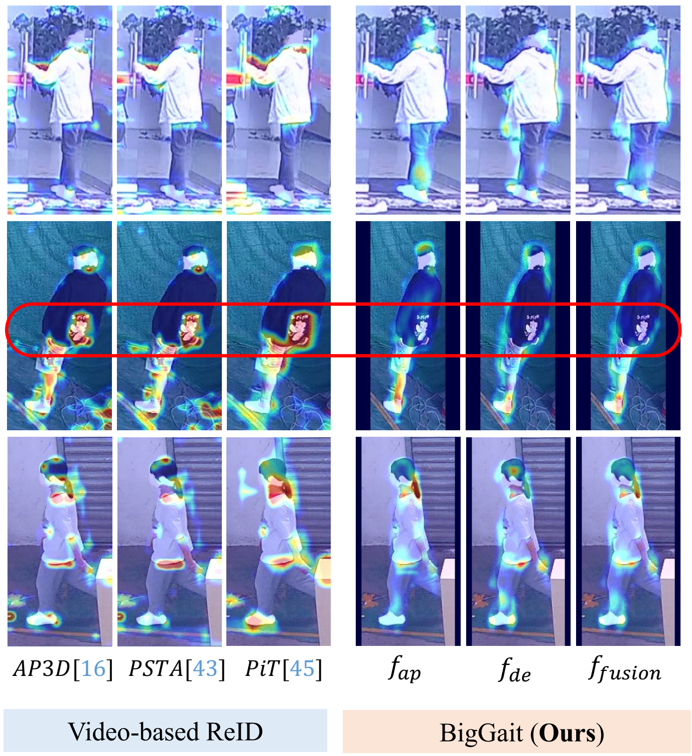
[1] D. Ye*, C. Fan*, J. Ma, X. Liu, and S. Yu, BigGait: Learning Gait Representation You Want by Large Vision Models, CVPR2024.

Challenge#2: Good Gait Representation Exploration

Solution#2.3: Learning Gait Representation You Wanted by Large Vision Models

Input	Model	Venue	CL	UP	DN	BG	Mean
Skeleton	GaitGraph2 [43]	CVPRW'22	5.0	5.3	5.8	6.2	5.6
	Gait-TR [48]	ES'23	15.7	18.3	18.5	17.5	17.5
	MSGG [35]	MTA'23	29.0	34.5	37.1	33.3	33.5
Sils	GaitSet [3]	TPAMI'22	60.2	65.2	65.1	68.5	64.8
	GaitPart [10]	CVPR'20	64.3	67.8	68.6	71.7	68.1
	AUG-OGBase [25]	CVPR'23	52.1	57.3	60.1	63.3	58.2
	GaitBase [13]	CVPR'23	71.6	75.0	76.8	78.6	75.5
	DeepGaitV2 [11]	Arxiv	78.6	84.8	80.7	89.2	83.3
Parsing	GaitBase^{p}	CVPR'23	59.1	62.1	66.8	68.1	64.0
Parsing+Sils	GaitBase^{p+s}	CVPR'23	73.6	76.2	79.1	79.2	77.0
Skeleton+Sils	SkeletonGait++ [14]	AAAI'24	79.1	83.9	81.7	89.9	83.7
RGB+Sils	GaitEdge [28]	ECCV'22	66.9	74.0	70.6	77.1	72.2
RGB	AP3D [18]	ECCV'20	53.4	57.3	69.7	91.4	67.8
	PSTA [45]	ICCV'21	42.2	52.2	60.3	84.5	59.8
	PiT [47]	TII'22	41.0	47.6	64.3	91.0	61.0
	BigGait	Ours	76.0	79.1	84.2	93.0	83.1
	BigGait-L	Ours	79.0	82.3	86.7	**94.5**	85.6
	†BigGait[30]	Ours	**82.6**	**85.9**	**87.1**	93.1	**87.2**

† *BigGait[30] is officially recommended for citation.*



BigGait can learn discriminative yet clean gait representations!

- Within-domain Improvements 🚀
- Cross-clothing Improvements 🚀
- Cross-domain Improvements 🚀
- Activation Visualization ✅
- Background Inclusion Damages Performance ✅

[1] D. Ye*, C. Fan*, J. Ma, X. Liu, and S. Yu, BigGait: Learning Gait Representation You Want by Large Vision Models, CVPR2024.

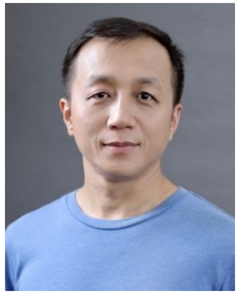
Future Works

- Cross-X Gait Recognition
- Large Gait Models
- Open-set Gait Recognition

Main Co-authors

Many Thanks to ALL!!!

Supervisor



Prof. Shiqi Yu
SUSTech

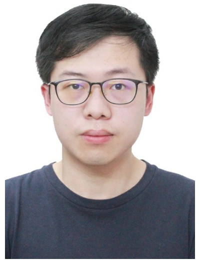
Ph.D. Candidates

Chuanfu Shen
SUSTech & HKU

Jingzhe Ma
SUSTech

Chao Fan (ME!)
SUSTech

Master Candidates



Zirui Zhou
SUSTech

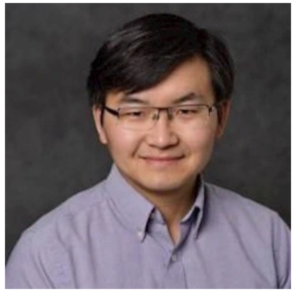
Junhao Liang
SUSTech

Dingqiang Ye
SUSTech

Rui Wang
SUSTech

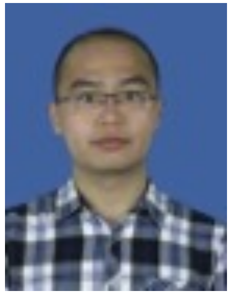
Dongyang Jin
SUSTech

More Advisors



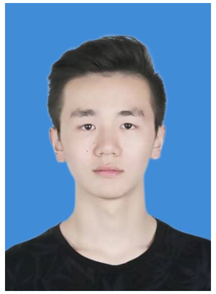
Prof. Xiaoming Liu
MSU

Prof. Yongzhen Huang
BNU



Prof. Saihui Hou
BNU

More Guys (Ph.D. Candidates)



Jilong Wang
USTC

Shinan Zou
CSU

Thanks for Attention!!!

[1] OpenGait Codebase @ GitHub

https://github.com/ShiqiYu/OpenGait

OpenGait

The Source Code of Our Works are Released on This Repo.

[2] HID Competition @ CodaLab

https://hid2024.iapr-tc4.org

Awards

Our sponsor, Watrix Technology, will provide 6 awards (19,000 CNY in total, ~2,6000 USD) to the top 6 teams from the second phase.

- First Prize (1 team): 10,000 CNY (~1,500 USD)
- Second Prize (2 teams): 3,000 CNY (~450 USD)
- Third Prize (3 teams): 1,000 CNY (~150 USD)

where CNY stands for Chinese Yuan.

Many Thanks to Our Sponsor! 银河水滴 WATRIX. AI

[1] C. Fan, J. Liang, S. Shen, S. Hou, Y. Huang, and S. Yu, OpenGait: Revisiting Gait Recognition Toward Better Practicality, CVPR2023 (Highlight Paper)
[2] S. Yu, C. Wang, Y. Zhao, L. Wang, Ming. Wang, Q. Li, W. Li, R. Wang, Y. Huang, L. Wang, Y. Makihara, and M. A. R. Ahad, Human Identification at a Distance: Challenges, Methods and Results on HID 2023, IJCB2023.