Gait Recognition from 2D to 3D

报告人：刘鑫辰
京东探索研究院

Outline

01

Introduction

What is Gait Recognition ?

Identify the same person across multiple cameras with gait biometrics, i.e., the motion patterns of human walking [1].

Advantages:

1. Remotely accessible
2. Non-contacting
3. Hard to impersonate
4. Robust to lightening variation and cloth changing

Challenges:

1. Noise of silhouette segmentation
2. Cross-view variance
3. Uncertainty in temporal pattern
4. Ambiguity of 3D bodies projected into 2D images

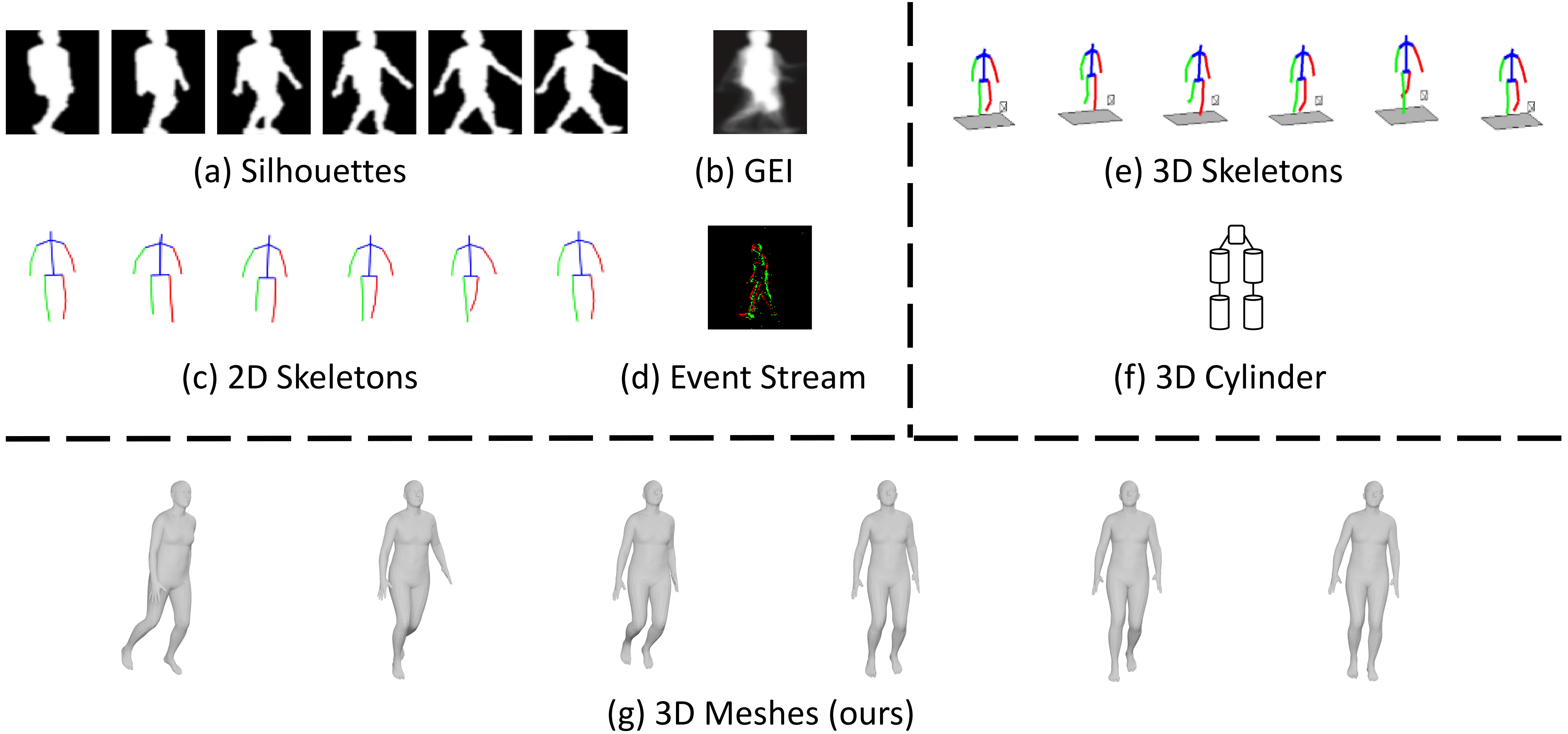
Applications:

1. ID authentication
2. Missing person search
3. Public security
4. …

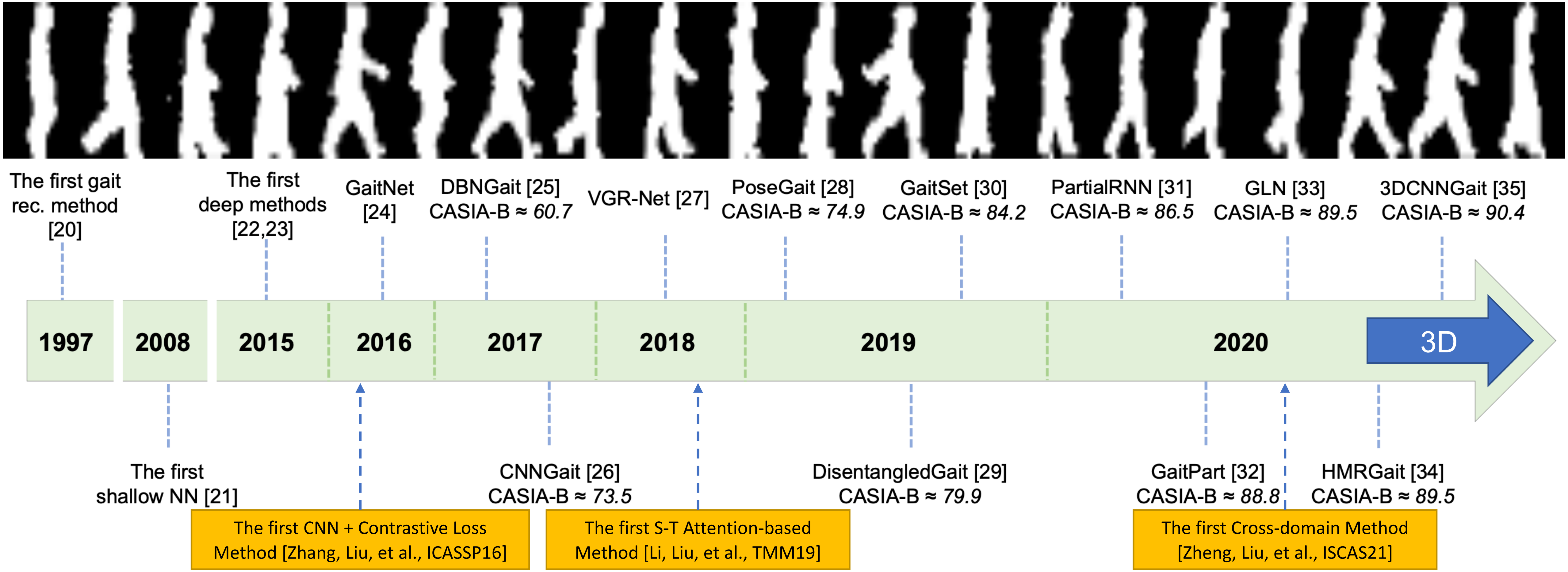
[1] Liang Wang, Tieniu Tan, Huazhong Ning, Weiming Hu: Silhouette Analysis-Based Gait Recognition for Human Identification. IEEE Trans. Pattern Anal. Mach. Intell. 25(12): 1505-1518 (2003)

Gait Representations

JDT 京东科技



Timeline of Gait Recognition



1. Alireza Sepas-Moghaddam, Ali Etemad: Deep Gait Recognition: A Survey. CoRR abs/2102.09546 (2021)
2. Cheng Zhang, Wu Liu, et al., : Siamese neural network based gait recognition for human identification. ICASSP 2016: 2832-2836
3. Shuangqun Li, Wu Liu, Huadong Ma: Attentive Spatial-Temporal Summary Networks for Feature Learning in Irregular Gait Recognition. IEEE Trans. Multim. 21(9): 2361-2375 (2019)
4. Jinkai Zheng, Xinchen Liu, et al. : TraND: Transferable Neighborhood Discovery for Unsupervised Cross-Domain Gait Recognition. ISCAS 2021: 1-5

02

Related Work

Existing Gait Datasets

JDT 京东科技

Dataset	Year	Subject #	Sequence #	Cam #	Data Type	Speed	Wild	3D-View
CASIA-A [17]	2003	20	240	3	RGB	✗	✗	✗
USF HumanID [12]	2005	122	1,870	2	RGB	✗	✗	✗
CASIA-B [20]	2006	124	13,680	11	RGB, Silh.	✗	✗	✗
CASIA-C [14]	2006	153	1,530	1	Infrared, Silh.	✓	✗	✗
OU-ISIR Speed [15]	2010	34	306	1	Silh.	✓	✗	✗
OU-ISIR MV [10]	2010	168	4,200	25	Silh.	✗	✗	✗
OU-LP [6]	2012	4007	31,368	4	Silh.	✗	✗	✗
OU-MVLP [13]	2018	10,307	259,013	14	Silh.	✗	✗	✗
OU-MVLP Pose [1]	2020	10,307	259,013	14	2D Pose	✗	✗	✗
GREW*[22]	2021	26,345	128,671	882	Silh., 2D/3D Pose, Flow	✗	✓	✗
Gait3D	-	**4,000**	**25,309**	**46**	**Silh., 2D/3D Pose, 3D Mesh&SMPL**	✓	✓	✓

[1] Sudeep Sarkar, P. Jonathon Phillips, Zongyi Liu, Isidro Robledo Vega, Patrick Grother, Kevin W. Bowyer: The HumanID Gait Challenge Problem: Data Sets, Performance, and Analysis. IEEE Trans. Pattern Anal. Mach. Intell. 27(2): 162-177 (2005)

[2] Shiqi Yu, Daoliang Tan, Tieniu Tan: A Framework for Evaluating the Effect of View Angle, Clothing and Carrying Condition on Gait Recognition. ICPR (4) 2006: 441-444

[3] Yasushi Makihara, Hidetoshi Mannami, Akira Tsuji, Md. Altab Hossain, Kazushige Sugiura, Atsushi Mori, Yasushi Yagi: The OU-ISIR Gait Database Comprising the Treadmill Dataset. IPSJ Trans. Comput. Vis. Appl. 4: 53-62 (2012)

[4] Yasushi Makihara, Atsuyuki Suzuki, Daigo Muramatsu, Xiang Li, Yasushi Yagi: Joint Intensity and Spatial Metric Learning for Robust Gait Recognition. CVPR 2017: 6786-6796

[5] Md. Zasim Uddin, Trung Ngo Thanh, Yasushi Makihara, Noriko Takemura, Xiang Li, Daigo Muramatsu, Yasushi Yagi: The OU-ISIR Large Population Gait Database with real-life carried object and its performance evaluation. IPSJ Trans. Comput. Vis. Appl. 10: 5 (2018)

[6] Noriko Takemura, Yasushi Makihara, Daigo Muramatsu, Tomio Echigo, Yasushi Yagi: Multi-view large population gait dataset and its performance evaluation for cross-view gait recognition. IPSJ Trans. Comput. Vis. Appl. 10: 4 (2018)

[7] Zheng Zhu, Xianda Guo, Tian Yang, Junjie Huang, Jiankang Deng, Guan Huang, Dalong Du, Jiwen Lu, Jie Zhou: Gait Recognition in the Wild: A Benchmark. ICCV, 2021.

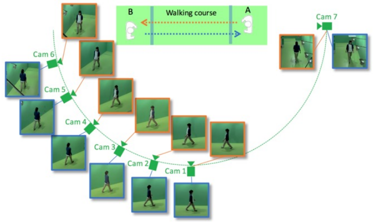
Existing Gait Datasets

JDT 京东科技

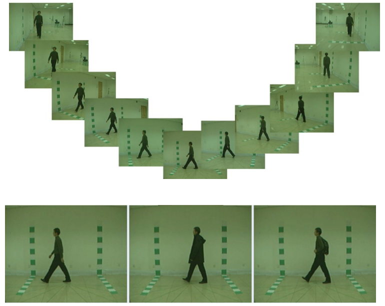
CASIA-A

OU-LP

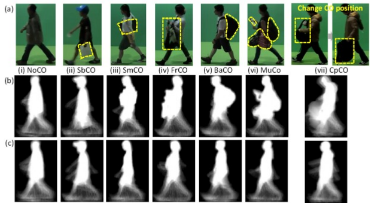
OU-MVLP



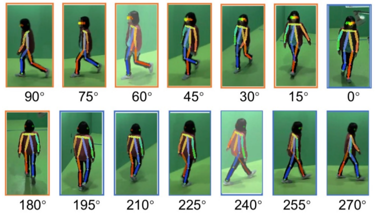
CASIA-B



OU-LP-Bag

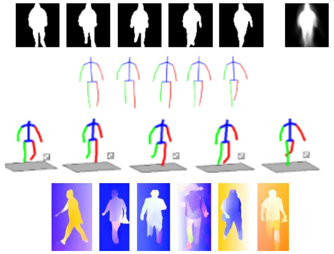


OU-MVLP-Pose

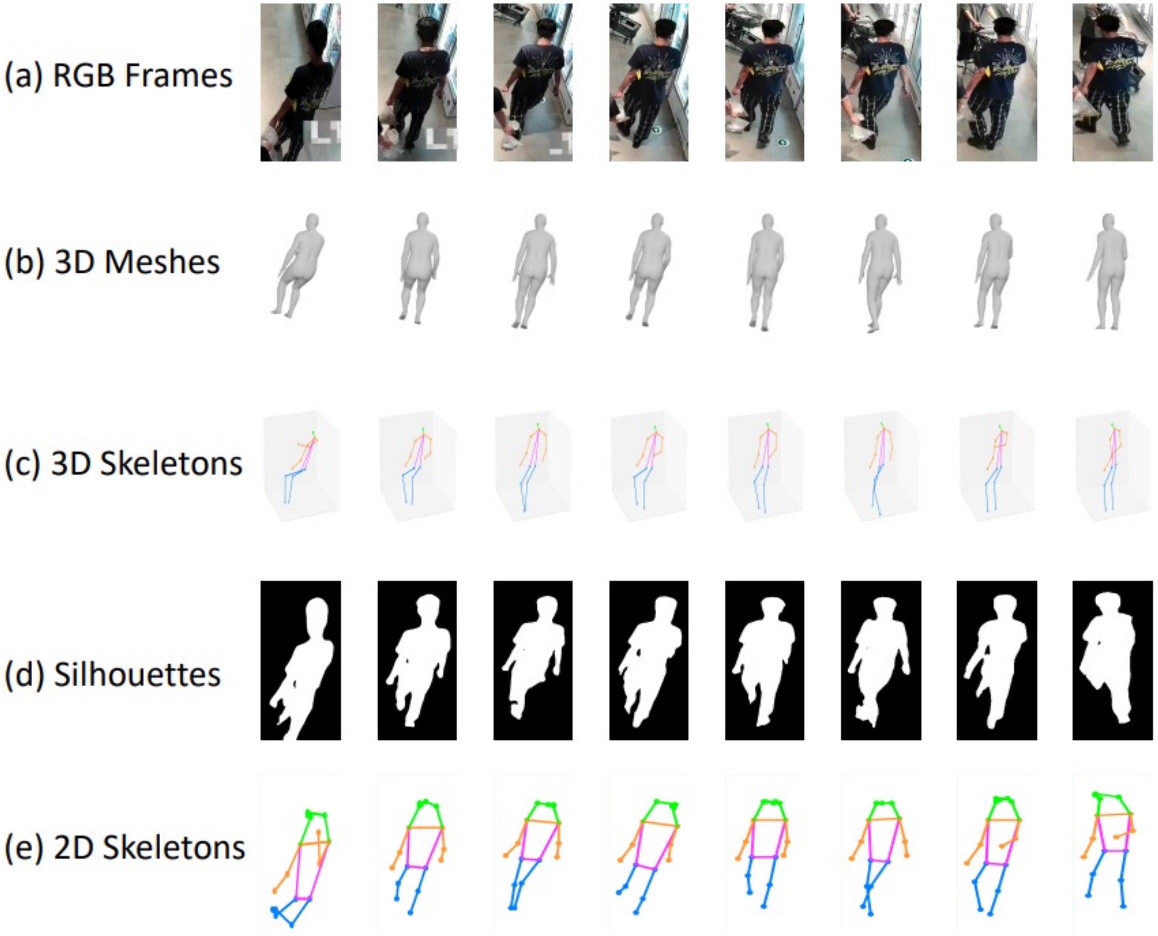


OU-LP-Age

GREW



Gait3D (ours)



Existing Gait Recognition Methods

JDT 京东科技

- Early Work (Before 2015):
  - Model-based:
    - 3D voxel-based [IJCB11]
  - Appearance-based:
    - **Gait Energy Image** [TPAMI06]
    - Chrono-Gait Image [TPAMI12]
- Recent Work:
  - GEI+CNN/SNN:
    - SiaNet [ICASSP16]
    - GEINet [ICB16]
  - Sequence/Set + RNN/LSTM/CNN:
    - ASTSN [TMM19]
    - GaitSet [AAAI19]
    - GaitPart [CVPR20]
    - GLN [ECCV20]
  - Pose + GCN
    - PoseGait [PR20]
    - GaitGraph [arXiv21]

Results on CASIA-B

Type	Method	Probe NM	Probe BG	Probe CL
appearance -based	GaitNet [1]	91.6	85.7	58.9
	GaitSet [2]	95.0	87.2	70.4
	GaitPart [3]	**96.2**	**91.5**	**78.7**
model -based	PoseGait [6]	60.5	39.6	29.8
	GaitGraph	**87.7**	**74.8**	**66.3**

Results on Gait3D

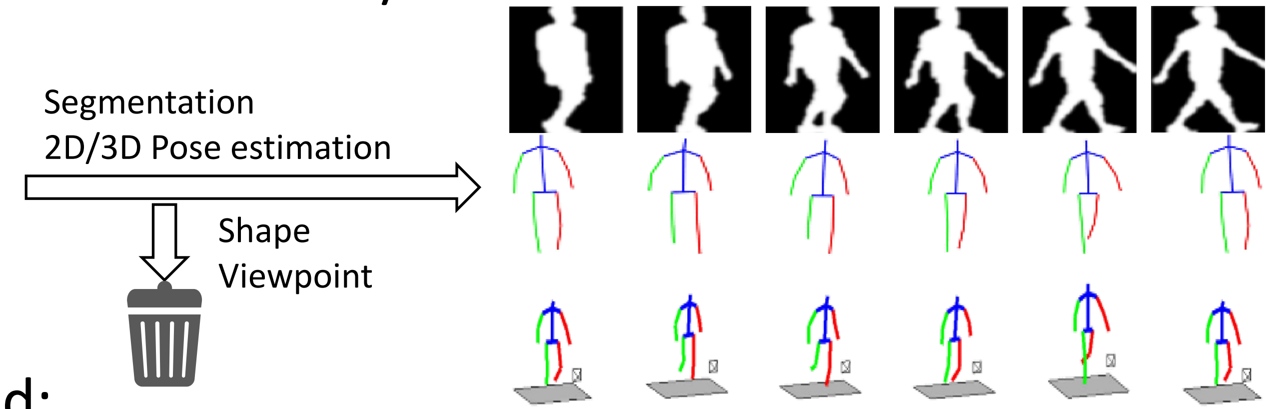
Input Size (W×H)		88×128			
Methods	Publication	R-1 (%)	R-5 (%)	mAP (%)	mINP
GEINet [32]	ICB 2016	7.00	16.30	6.05	3.77
GaitSet [5]	AAAI 2019	42.60	63.10	33.69	19.69
GaitPart [9]	CVPR 2020	29.90	50.60	23.34	13.15
GLN [15]	ECCV 2020	42.20	64.50	33.14	19.56
GaitGL [22]	ICCV 2021	23.50	38.50	16.40	9.20
CSTL [16]	ICCV 2021	12.20	21.70	6.44	3.28
PoseGait [21]	PR 2020	0.24	1.08	0.47	0.34
GaitGraph [38]	arXiv 2021	6.25	16.23	5.18	2.42

03

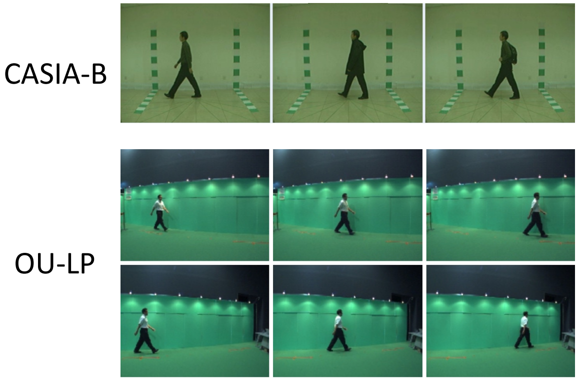
The Gait3D Benchmark

Why Gait Recognition in the Wild with 3D information?

- Problem of current gait recognition:
  - Information loss of 2D silhouette and 2D/3D skeleton

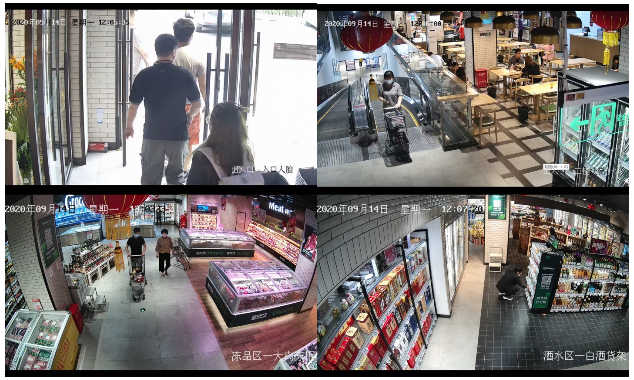


  - Gap between Lab and Wild:



Walking Style:
- Directed
- Straight routes
- Fixed speed

Environment Settings:
- Constrained scenes
- Pre-defined viewpoints
- Simple background



Supermarket

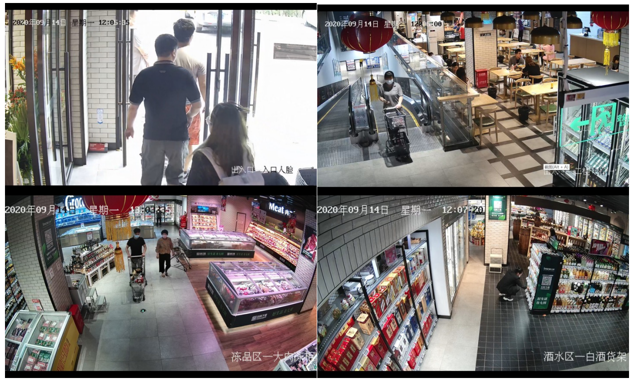
Walking Style:
- Natural Walking
- Irregular routes
- Varied speed

Environment Settings:
- Occlusions
- Arbitrary 3D viewpoint
- Background clutter

JDT 京东科技

What do we need and what are our challenges ?

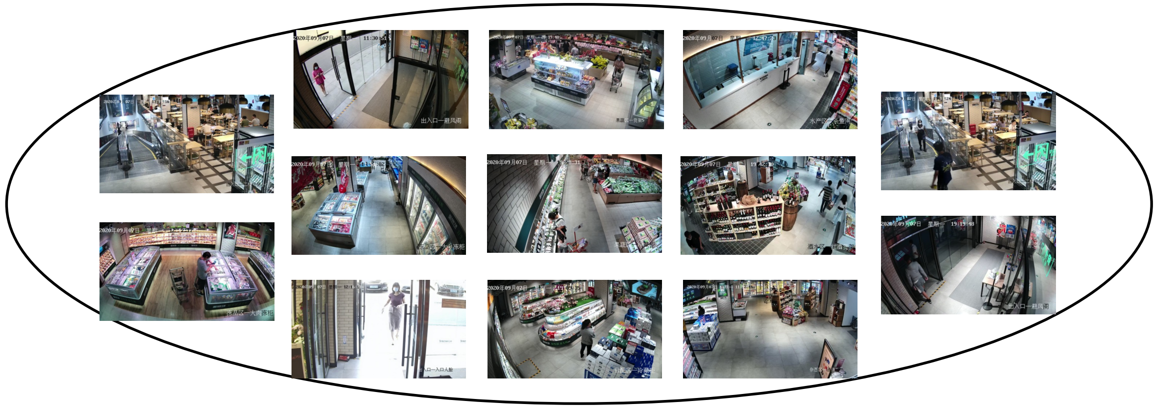
- In the wild



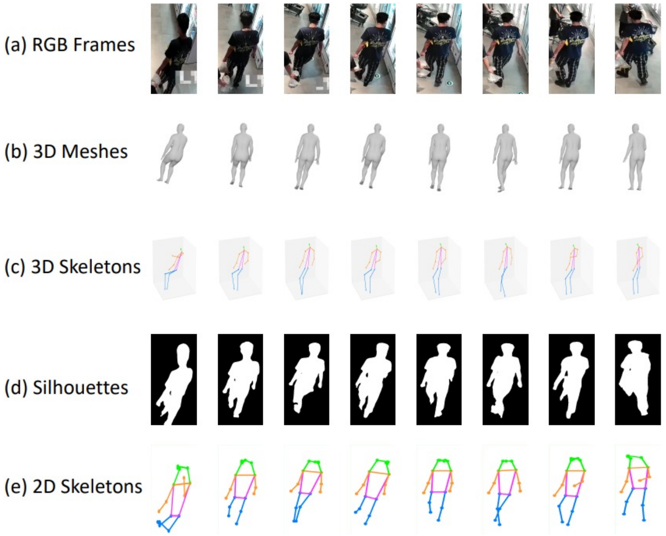
Walking Style:
- Natural Walking
- Irregular routes
- Varied speed

Environment Settings:
- Occlusions
- Arbitrary 3D viewpoint
- Background clutter

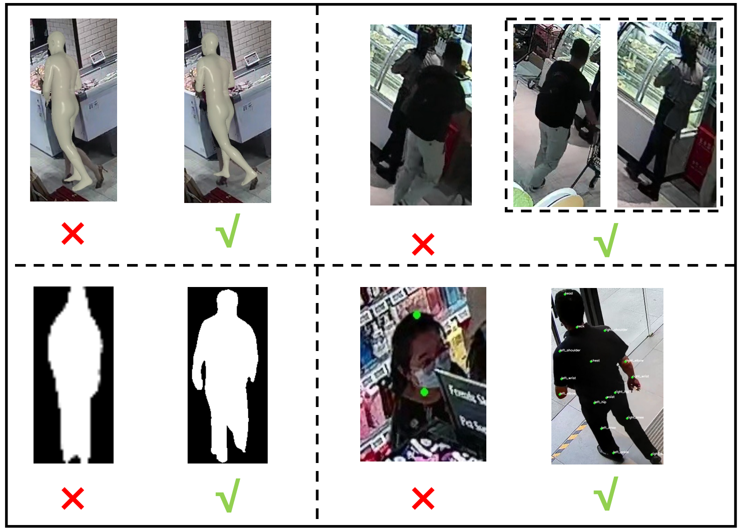
- Large scale and Diverse



- Multi-modal



- High-quality



Gait3D Construction

JDT 京东科技

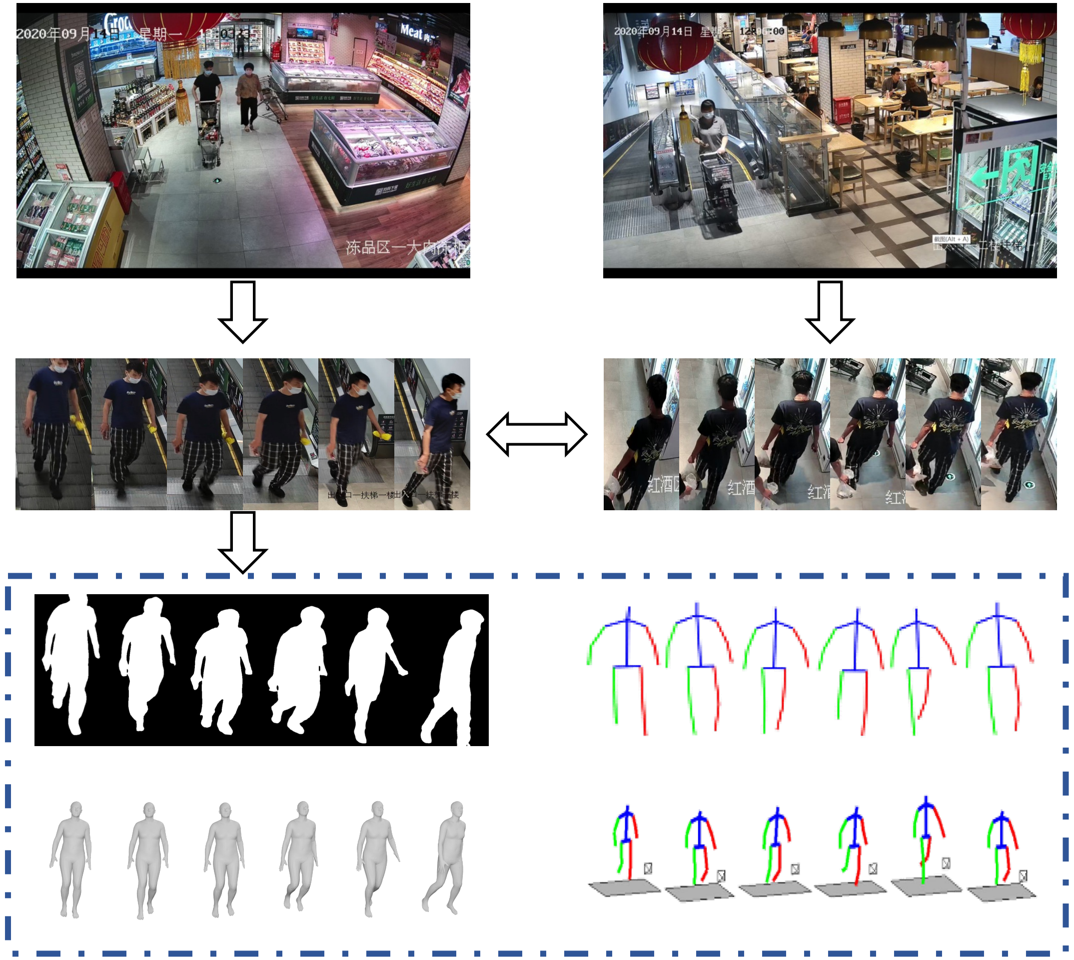
- A Large-scale In-the-Wild 3D Gait Recognition Benchmark
  - Construction:

1. Person detection and tracking from video frames

2. Cross-camera matching of the same ID using ReID feature

3. Manually filtering out the wrong sequences in individual IDs

4. Generation of gait representations from sequences:
   1) silhouettes using HRNet (finetuned)
   2) 2D pose using HRNet (finetuned)
   3) 3D SMPL & mesh using ROMP (finetuned)
   4) 3D pose using ROMP (finetuned)

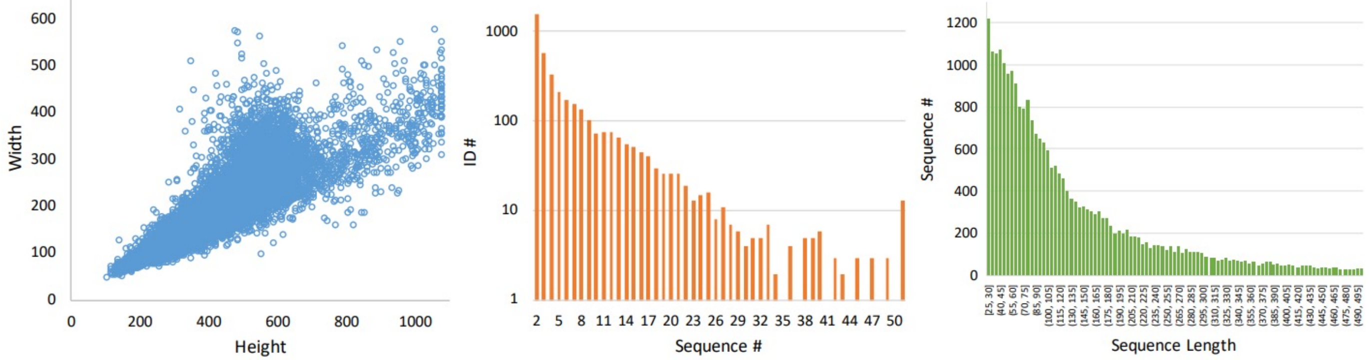


Jinkai Zheng, Xinchen Liu, Wu Liu, Lingxiao He, Chenggang Yan, Tao Mei: Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. . CVPR 2022: 20228-20237

Gait3D Dataset

- A Large-scale In-the-Wild 3D Gait Recognition Benchmark
  - Statistics:
    - **1,090-hour** videos of **7 days** from **unconstrained environment**, i.e., a large supermarket
    - **39** cameras with various viewpoints, **1,920x1,080** resolution, and **25** FPS
    - **4,000** IDs, **25,309** sequences, **3,279,239** bounding boxes
  - Protocol:
    - 3,000/1,000 IDs for training/testing
    - 1,000/5,369 sequences for query/gallery of testing set
    - Metrics: Rank-1, Rank5, mAP, mINP

Scan the QR code for the dataset!



(a) Statistics of frame sizes. (b) ID # over sequence #. (c) Sequence # over sequence lengths.

Jinkai Zheng, Xinchen Liu, Wu Liu, Lingxiao He, Chenggang Yan, Tao Mei: Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. . CVPR 2022: 20228-20237

Benchmarking of SOTA Methods

JDT 京东科技

Comparison of the state-of-the-art gait recognition methods on Gait3D.

Input Size (W×H)		88×128				44×64			
Methods	Publication	R-1 (%)	R-5 (%)	mAP (%)	mINP	R-1 (%)	R-5 (%)	mAP (%)	mINP
GEINet [32]	ICB 2016	7.00	16.30	6.05	3.77	5.40	14.20	5.06	3.14
GaitSet [5]	AAAI 2019	42.60	63.10	33.69	19.69	36.70	58.30	30.01	17.30
GaitPart [9]	CVPR 2020	29.90	50.60	23.34	13.15	28.20	47.60	21.58	12.36
GLN [15]	ECCV 2020	42.20	64.50	33.14	19.56	31.40	52.90	24.74	13.58
GaitGL [22]	ICCV 2021	23.50	38.50	16.40	9.20	29.70	48.50	22.29	13.26
CSTL [16]	ICCV 2021	12.20	21.70	6.44	3.28	11.70	19.20	5.59	2.59
PoseGait [21]	PR 2020	0.24	1.08	0.47	0.34	-	-	-	-
GaitGraph [38]	arXiv 2021	6.25	16.23	5.18	2.42	-	-	-	-

Results of cross-domain experiments.

Source	Target	R-1 (%)	R-5 (%)	mAP (%)
CASIA-B [56]	Gait3D	6.90	14.60	4.64
OU-LP [17]	Gait3D	6.10	12.40	4.42
GREW [63]	Gait3D	16.50	31.10	11.71
Gait3D	CASIA-B [56]	66.71	71.59	33.88
Gait3D	OU-LP [17]	97.84	99.38	68.06
Gait3D	GREW [63]	43.86	60.89	28.06

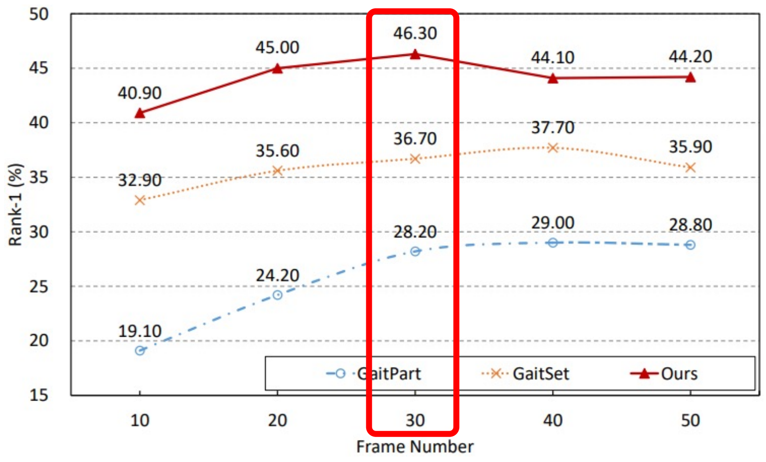
Insights:

1. Model-based approaches are greatly worse than model-free methods due to information loss.
2. There is a huge gap between the in-the-lab dataset and in-the-wild application. Our Gait3D enables the model to learn more generalized gait representations.

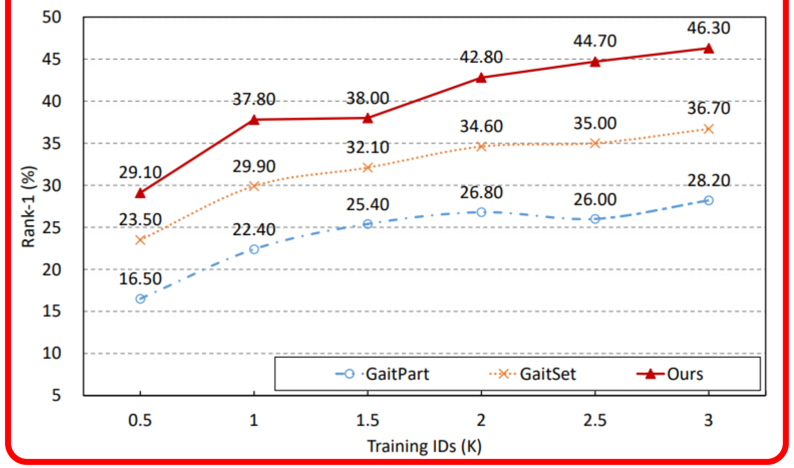
Jinkai Zheng, Xinchen Liu, Wu Liu, Lingxiao He, Chenggang Yan, Tao Mei: Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. . CVPR 2022: 20228-20237

Experiments

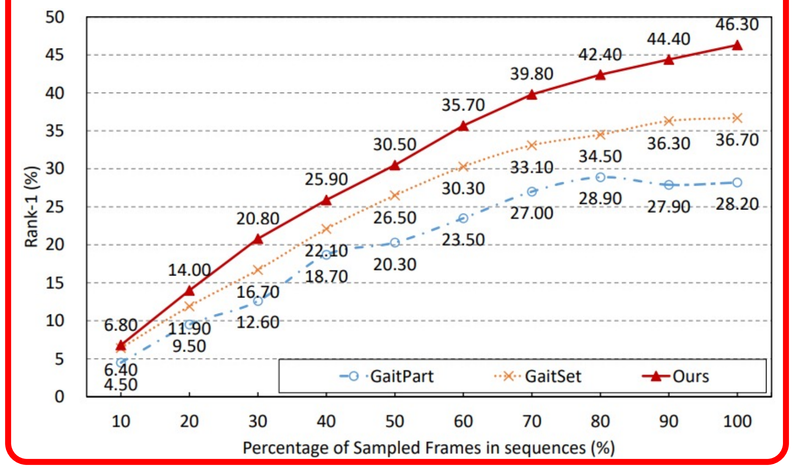
The effect of frame numbers in sequences.



The effect of different training ID numbers.



The effect of the lengths of test sequences.



Insights:

1. The best performance occurs around 30 frames per sequence in the training phase.
2. The performance grows stably with more training IDs and testing frames.

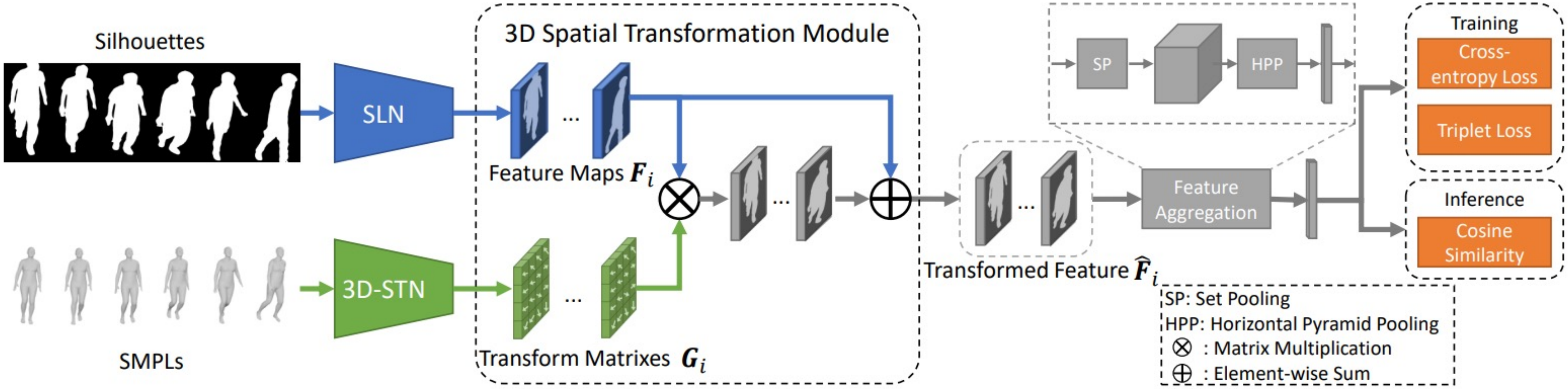
Jinkai Zheng, Xinchen Liu, Wu Liu, Lingxiao He, Chenggang Yan, Tao Mei: Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. . CVPR 2022: 20228-20237

04

Our Methods

SMPLGait framework

The architecture of the SMPLGait framework for 3D gait recognition in the wild.



The architecture of the Silhouette Learning Network (SLN).

SLN

Layers	Kernel #	Kernel Size	Stride	Padding
Conv1	64	5×5	1	2
LeakyReLU (0.01)	-	-	-	-
Conv2	64	3×3	1	1
LeakyReLU (0.01)	-	-	-	-
Max Pooling	-	2×2	2	0
Conv3	128	3×3	1	1
LeakyReLU (0.01)	-	-	-	-
Conv4	128	3×3	1	1
LeakyReLU (0.01)	-	-	-	-
Max Pooling	-	2×2	2	0
Conv5	256	3×3	1	1
LeakyReLU (0.01)	-	-	-	-
Conv6	256	3×3	1	1
LeakyReLU (0.01)	-	-	-	-

The architecture of the 3D Spatial Transformation Network (3D-STN).

3D-STN

Layers	Neuron #	Dropout Rate
FC1	128	0.0
BN1	-	-
ReLU	-	-
FC2	256	0.2
BN2	-	-
ReLU	-	-
FC3	$h \times w$	0.2
BN3	-	-
ReLU	-	-

Jinkai Zheng, Xinchen Liu, Wu Liu, Lingxiao He, Chenggang Yan, Tao Mei: Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. . CVPR 2022: 20228-20237

Experiments

JDT 京东科技

Comparison of the state-of-the-art gait recognition methods on Gait3D.

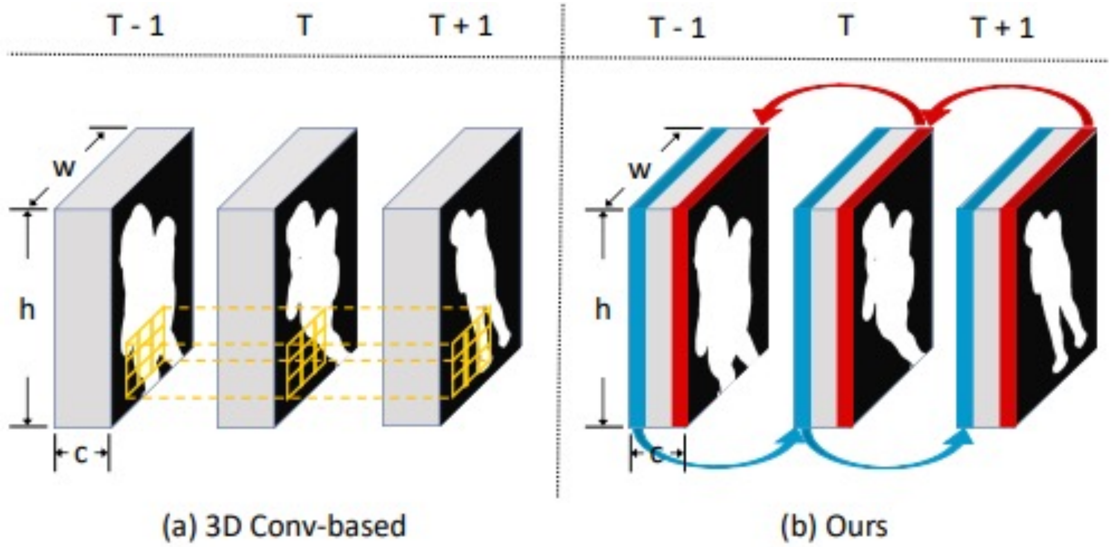
Methods	Input Size (W×H) Publication	88×128 R-1 (%)	R-5 (%)	mAP (%)	mINP	44×64 R-1 (%)	R-5 (%)	mAP (%)	mINP
GEINet [32]	ICB 2016	7.00	16.30	6.05	3.77	5.40	14.20	5.06	3.14
GaitSet [5]	AAAI 2019	42.60	63.10	33.69	19.69	36.70	58.30	30.01	17.30
GaitPart [9]	CVPR 2020	29.90	50.60	23.34	13.15	28.20	47.60	21.58	12.36
GLN [15]	ECCV 2020	42.20	64.50	33.14	19.56	31.40	52.90	24.74	13.58
GaitGL [22]	ICCV 2021	23.50	38.50	16.40	9.20	29.70	48.50	22.29	13.26
CSTL [16]	ICCV 2021	12.20	21.70	6.44	3.28	11.70	19.20	5.59	2.59
PoseGait [21]	PR 2020	0.24	1.08	0.47	0.34	-	-	-	-
GaitGraph [38]	arXiv 2021	6.25	16.23	5.18	2.42	-	-	-	-
SMPLGait w/o 3D	Ours	47.70	67.20	37.62	22.24	42.90	63.90	35.19	20.83
SMPLGait	Ours	**53.20**	**71.00**	**42.43**	**25.97**	**46.30**	**64.50**	**37.16**	**22.23**

Insights:

1. The performance is significantly improved with the addition of 3D mesh.

Jinkai Zheng, Xinchen Liu, Wu Liu, Lingxiao He, Chenggang Yan, Tao Mei: Gait Recognition in the Wild with Dense 3D Representations and A Benchmark. . CVPR 2022: 20228-20237

Temporal Modeling on Gait3D



✓ Motivation:

- 3D convolution-based methods have the problem of feature misalignment.
- Existing methods can hardly model varied temporal dynamics of gait sequences in unconstrained scenes.
- Existing methods based on 3D convolution for temporal modeling have the problem of large-scale model parameters and difficulty in training.

Jinkai Zheng, Xinchen Liu, Xiaoyan Gu, Yaoqi Sun, Chuang Gan, Jiyong Zhang, Wu Liu, Chenggang Yan: Gait Recognition in the Wild with Multi-hop Temporal Switch. ACM MM 2022

MTSGait

The architecture of the MTSGait framework for modeling spatial and multi-hop temporal gait features

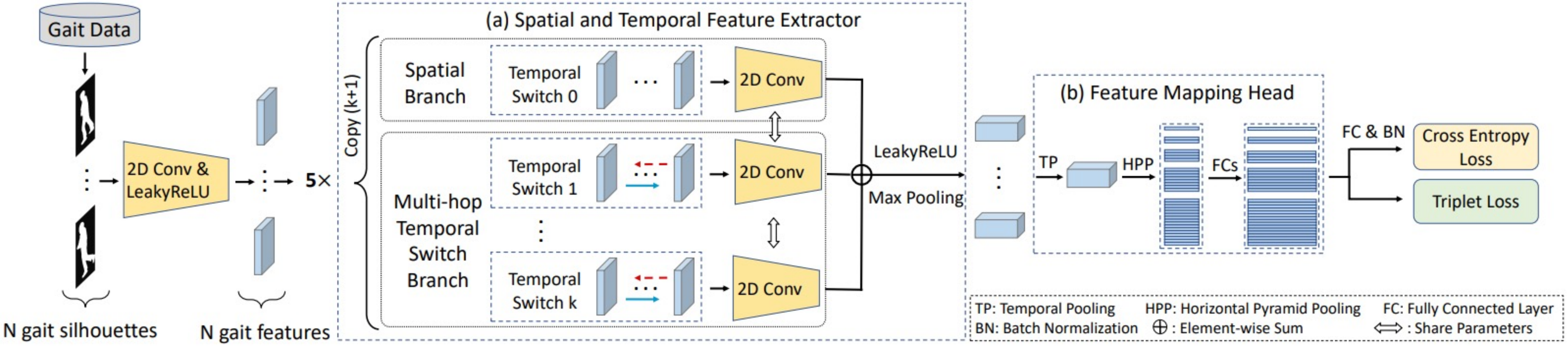
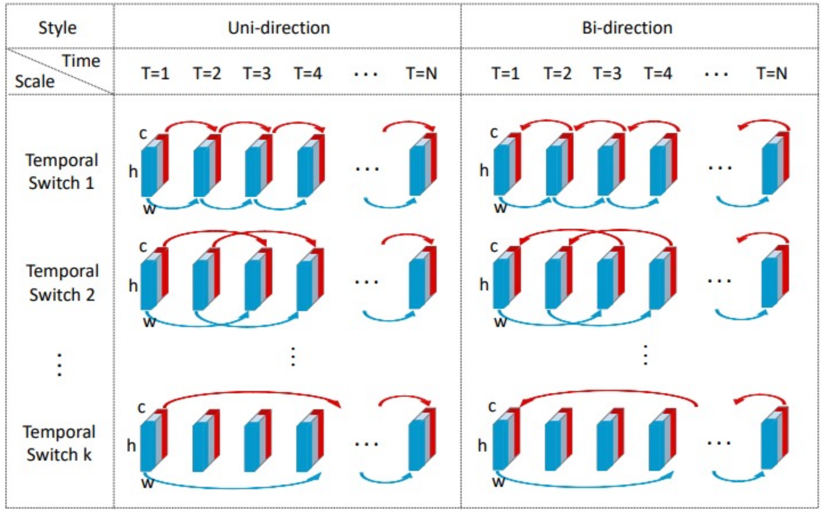


Illustration of the uni-direction, bi-direction, and multi-hop temporal switch operation



- ✓ MTSGait can learn spatial and multi-scale temporal information simultaneously.
- ✓ MTSGait avoids the problem of large-scale model parameters and difficulty in training 3D CNN.
- ✓ We propose a new sampling strategy, i.e., Non-cyclic continuous sampling, to learn more robust temporal features.

Jinkai Zheng, Xinchen Liu, Xiaoyan Gu, Yaoqi Sun, Chuang Gan, Jiyong Zhang, Wu Liu, Chenggang Yan: Gait Recognition in the Wild with Multi-hop Temporal Switch. ACM MM 2022

Experiments

JDT 京东科技

Results on Gait3D dataset

Methods	Rank-1	Rank-5	mAP	mINP
PoseGait [17]	0.24	1.08	0.47	0.34
GaitGraph [29]	6.25	16.23	5.18	2.42
GEINet [25]	5.40	14.20	5.06	3.14
GaitSet [2]	36.70	58.30	30.01	17.30
GaitPart [4]	28.20	47.60	21.58	12.36
GLN [9]	31.40	52.90	24.74	13.58
GaitGL [19]	29.70	48.50	22.29	13.26
CSTL [10]	11.70	19.20	5.59	2.59
SMPLGait [47]	46.30	64.50	37.16	**22.23**
Ours w/o MTS	42.90	63.90	35.19	20.83
Ours	**48.70**	**67.10**	**37.63**	21.92

Results on GREW dataset

Methods	Rank-1	Rank-5	Rank-10	Rank-20
PoseGait [17]	0.23	1.05	2.23	4.28
GaitGraph [29]	1.31	3.46	5.10	7.51
GEINet [25]	6.82	13.42	16.97	21.01
GaitSet [2]	46.28	63.58	70.26	76.82
GaitPart [4]	44.01	60.68	67.25	73.47
GaitGL [19]	47.28	63.56	69.32	74.18
Ours w/o MTS	50.42	67.89	74.28	79.38
Ours	**55.32**	**71.28**	**76.85**	**81.55**

Results:

1. Our method achieves superior performance on two public gait in-the-wild datasets, i.e., Gait3D and GREW, compared with state-of-the-art methods.
2. We can observe that our method is better than SMPLGait. This means that in addition to adding additional input data like 3D meshes, temporal modeling is also very important.

Jinkai Zheng, Xinchen Liu, Xiaoyan Gu, Yaoqi Sun, Chuang Gan, Jiyong Zhang, Wu Liu, Chenggang Yan: Gait Recognition in the Wild with Multi-hop Temporal Switch. ACM MM 2022

05

Conclusion

JDT 京东科技

Conclusion

- Contributions:
  - ✓ We believe that the next direction of gait recognition is gait recognition in the wild, especially in combination with dense 3D representations, such as 3D meshes.
  - ✓ There are many potential directions for this challenging task:
    - ➢ How to design a model for learning more discriminative features directly from 3D meshes.
    - ➢ How to learn the temporal information of gait representation, because the walking speed and route in Gait3D are irregular, it is significantly different from the datasets built in the lab.
    - ➢ How to fuse the multi-modal information like silhouette, 2D/3D skeleton, and 3D mesh for gait recognition in the wild.
- Discussion:
  - ✓ Dataset access:
    - ➢ case-by-case application via license
  - ✓ Ethic issues:
    - ➢ The involved subjects of the datasets are told to collect data for research purposes.
    - ➢ The dataset can only be used for research purposes.
    - ➢ We will not release any human cognizable data like original video files, original RGB video frames, and RGB bounding boxes of persons.

Scan the QR code for the Gait3D dataset!

Team & Collaborators

京东
JD Explore Academy


Xinchen Liu


Wu Liu


Lingxiao He


Tao Mei

杭州电子科技大学
HANGZHOU DIANZI UNIVERSITY


Jinkai Zheng


Yaoqi Sun


Jiyong Zhang


Chenggang Yan

中国科学院信息工程研究所
INSTITUTE OF INFORMATION ENGINEERING,CAS

MIT IBM

Toronto Metropolitan University


Xiaoyan Gu
IIE, CAS


Chuang Gan
MIT-IBM Watson AI Lab


Xiao-Ping Zhang
Ryerson University

CCIG 2022 中国图象图形大会 CHINESE CONGRESS ON IMAGE AND GRAPHICS

感谢观看

指导单位

中国科学技术协会
China Association for Science and Technology

主办单位

中国图象图形学学会
CHINA SOCIETY OF IMAGE AND GRAPHICS

承办单位

四川大学
SICHUAN UNIVERSITY

协办单位

电子科技大学
Universtiy of Electronic Svience and Technology of China

支持单位

成都市科技局

成都市博览局

成都市科学技术协会